



IntechOpen

Finite Element Analysis

Edited by David Moratal



Finite Element Analysis

edited by
David Moratal

Finite Element Analysis

<http://dx.doi.org/10.5772/281>

Edited by David Moratal

© The Editor(s) and the Author(s) 2010

The moral rights of the and the author(s) have been asserted.

All rights to the book as a whole are reserved by INTECH. The book as a whole (compilation) cannot be reproduced, distributed or used for commercial or non-commercial purposes without INTECH's written permission.

Enquiries concerning the use of the book should be directed to INTECH rights and permissions department (permissions@intechopen.com).

Violations are liable to prosecution under the governing Copyright Law.



Individual chapters of this publication are distributed under the terms of the Creative Commons Attribution 3.0 Unported License which permits commercial use, distribution and reproduction of the individual chapters, provided the original author(s) and source publication are appropriately acknowledged. If so indicated, certain images may not be included under the Creative Commons license. In such cases users will need to obtain permission from the license holder to reproduce the material. More details and guidelines concerning content reuse and adaptation can be found at <http://www.intechopen.com/copyright-policy.html>.

Notice

Statements and opinions expressed in the chapters are these of the individual contributors and not necessarily those of the editors or publisher. No responsibility is accepted for the accuracy of information contained in the published chapters. The publisher assumes no responsibility for any damage or injury to persons or property arising out of the use of any materials, instructions, methods or ideas contained in the book.

First published in Croatia, 2010 by INTECH d.o.o.

eBook (PDF) Published by IN TECH d.o.o.

Place and year of publication of eBook (PDF): Rijeka, 2019.

IntechOpen is the global imprint of IN TECH d.o.o.

Printed in Croatia

Legal deposit, Croatia: National and University Library in Zagreb

Additional hard and PDF copies can be obtained from orders@intechopen.com

Finite Element Analysis

Edited by David Moratal

p. cm.

ISBN 978-953-307-123-7

eBook (PDF) ISBN 978-953-51-5957-5

We are IntechOpen, the first native scientific publisher of Open Access books

3,250+

Open access books available

106,000+

International authors and editors

112M+

Downloads

151

Countries delivered to

Our authors are among the
Top 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Contents

Preface XI

- Chapter 1 **Finite element analysis on strains of viscoelastic human skull and duramater 1**
Xianfang YUE, Li WANG, Ruonan WANG and Feng ZHOU
- Chapter 2 **Application of finite element analysis in dentistry 43**
Ming-Lun Hsu and Chih-Ling Chang
- Chapter 3 **Finite element analysis for dental prosthetic design 61**
Akikazu Shinya and Daiichiro Yokoyama
- Chapter 4 **Application of finite element analysis in root canal therapy 103**
Tao Hu, Ran Cheng, Meiying Shao, Hui Yang, Ru Zhang, Qianhua Gao and Liyang Guo
- Chapter 5 **Finite element simulation. Applications in Orthopaedic Surgery and Traumatology 125**
Antonio Herrera, Luis Gracia, Elena Ibarz, Juan J. Panisello, José Cegoñino, Jesús Mateo, Javier Rodríguez-Vela and Sergio Puértolas
- Chapter 6 **Finite Element Analysis in Orthopaedic Biomechanics 155**
Daniel Kluess, Jan Wieding, Robert Souffrant, Wolfram Mittelmeier and Rainer Bader
- Chapter 7 **Orthopaedic Biomechanics: A Practical Approach to Combining Mechanical Testing and Finite Element Analysis 175**
Rad Zdero, Ph.D. and Habiba Bougherara, Ph.D.
- Chapter 8 **Finite element modeling for a morphometric and mechanical characterization of trabecular bone from high resolution magnetic resonance imaging 199**
Angel Alberich-Bayarri, Luis Martí-Bonmatí, M. Ángeles Pérez, Juan José Lerma and David Moratal
- Chapter 9 **Finite Element Modeling of the Human Lumbar Spine 213**
Marta Kurutz

- Chapter 10 **Analysis of human pressure ulcer and cushion pads for its prevention** 241
Masataka Akimoto M.D.,Ph.D.
- Chapter 11 **Microfinite Element Modelling for Evaluating Polymer Scaffolds Architecture and their Mechanical Properties from Microcomputed Tomography** 255
Angel Alberich-Bayarri, Manuel Salmerón Sánchez,
M. Ángeles Pérez and David Moratal
- Chapter 12 **Computational Modelling of Auxetics** 269
Bogdan Maruszewski, Tomasz Strek, Artur A. Pozniak
and Krzysztof W. Wojciechowski
- Chapter 13 **Modelling of polymeric fibre-composites and finite element simulation of mechanical properties** 289
Robert A. Shanks
- Chapter 14 **Finite Element Analysis of Nonlinear Isotropic Hyperelastic and Viscoelastic Materials for Thermoforming Applications** 319
Erchiqui Fouad
- Chapter 15 **Finite element analysis and Fracture in viscoelastic materials by M0v integral-Part I: crack initiation** 339
Rostand Moutou Pitti, Frédéric Dubois and Mustapha Taazount
- Chapter 16 **Finite Element modelling of Elastic-Plastic Contact of Rough Surfaces** 365
Jamil Abdo, M. Danish Haneef and Abdullah M. Al-Shabibi
- Chapter 17 **Numerical study of backward extrusion process using finite element method** 385
K.Abrinia and S.Orangi
- Chapter 18 **Finite element analysis on v-die bending process** 411
Sutasn Thipprakmas
- Chapter 19 **Analysis of Welding Residual Stresses and Its Applications** 433
Byeong-Choon Goo, Jung-Won Seo and Seung-Yong Yang
- Chapter 20 **Dynamic Finite Element Analysis on Underlay Microstructure of Cu/low-k Wafer during Wirebonding** 457
Hsiang-Chen Hsu, Chin-Yuan Hu, Wei-Yao Chang,
Chang-Lin Yeh and Yi-Shao Lai

- Chapter 21 **Finite element analysis of deformation and fracture of cylindrical tubes under internal moving pressures** 483
Majid Mirzaei
- Chapter 22 **FE Analysis of Evolution of Defects during Rolling** 505
Dr. YU Hai-liang
- Chapter 23 **Finite element analysis of strip and rolling mills** 565
Jian-guo CAO, Jie ZHANG, Ning KONG and Kai-fu MI
- Chapter 24 **Strain variations on rolling condition in accumulative roll-bonding by finite element analysis** 593
Tadanobu INOUE
- Chapter 25 **Finite element analysis of wall deflection and ground movements caused by braced excavations** 615
Gordon Tung-Chin Kung
- Chapter 26 **Vehicle-Bridge Dynamic Interaction Using Finite Element Modelling** 641
Arturo González
- Chapter 27 **Finite Element Modelling of Sound Transmission Loss in Reflective Pipe** 667
Tomasz Strek

Preface

Finite element analysis (FEA) is an engineering method for the numerical analysis of complex structures based on their material properties to determine the distribution of stress and strain when such structures are subjected to forces. Its development began in the 1940s in the field of structural engineering. Since then, the numerical method of FEA has been applied to a wide range of different areas in engineering and mathematical physics.

This book collects original and innovative research studies on recent applications in Finite Element Analysis, exhibiting various investigation directions and providing a bird's eye view on this very broad matter. The 27 chapters composing it have been grouped into four major domains: Biomedical Engineering, Materials Engineering, Process Analysis and Civil Engineering. It is meant to provide a small but valuable sample of contemporary research activities around the world in this field and it is expected to be useful to a large number of researchers. Through its 27 chapters the reader will have access to works related to dental prosthetic design, orthopaedics, polymer scaffolds architecture, viscoelastic materials, vehicle-bridge dynamic interactions and several other exciting topics.

The text is addressed not only to researchers, but also to professional engineers, engineering lecturers and students seeking to gain a better understanding of where Finite Element Analysis stands today. It has been written at a level suitable for the use in a graduate course on applications of finite element modelling and analysis (mechanical, civil and biomedical engineering studies, for instance).

I am very honoured to be editing such a valuable book, which contains contributions of a selected group of researchers describing the best of their work. I would like to express my sincere gratitude to all of them for their outstanding chapters.

I also wish to acknowledge the Sciyo editorial staff, in particular Iva Lipovic, for indispensable technical assistance in book preparation and publishing.

Editor

David Moratal

*Center for Biomaterials and Tissue Engineering
Universitat Politècnica de València
Valencia, Spain <http://dmoratal.webs.upv.es>*

Finite element analysis on strains of viscoelastic human skull and duramater

Xianfang YUE*, Li WANG, Ruonan WANG and Feng ZHOU

*Mechanical Engineering School,
University of Science and Technology Beijing, Beijing 100083
China*

1. Fundamental Theory of Finite Element

There are many mechanics and field problems in the domain of engineering technology, such as the analyses of stress-strain field, displacement field, temperature field in the heat conduction, flow field in the hydrodynamics, electromagnetic field in the electromagnetism, etc, which can be considered as the problems to solve the basic differential equations with a certain boundary conditions. Although the basic equations and boundary conditions have been established, very few simple questions can be solved to obtain the analytical solution. For the questions with complex mathematical equations or irregular physical boundary, it is insurmountably difficult to solve the mathematical problems by analytical method. There are usually two ways to solve such problems: one is the analytic method, another is the numerical method. After introducing the simplifying assumptions, the analytic method can be used to get the approximate solution. But this approach is not always feasible to the actual problems, so the incorrect or wrong answers will always be led to. To meet the needs of the complex project problems, the numerical solution can be adopted to finish a variety of effective numerical calculation. At present, numerical solution methods in engineering practical applications commonly include the limited element method, finite difference method, the boundary element method, weighted residual method, etc. With the rapid development and extensive application of computer, the finite element method, a relatively new and very effective numerical method, has gradually been used to solve the complex engineering problems.

1.1 Finite Element Method, FEM

Finite Element Method (FEM) is an integrated product of many disciplines, including mechanics, mathematical physics, computational methods, computer technology, and so on. The three research methods, theoretical analysis, scientific experiments and scientific

* Corresponding author. Address: School of Mechanical Engineering, University of Science and Technology Beijing, Beijing 100083, China

Tel.: +86-01-62332743; Fax: +86-10-62332743.

E-mail address: yuexf@me.ustb.edu.cn

computing, have been applied to study the nature problems. Because of the limitations of scientific theory and experiments, the scientific computing becomes one of the most important research tools. As one of ways to carry out the scientific computing, the finite-element method can almost analyze any complex engineering structure so as to obtain various mechanical properties in most engineering fields.

The analyses of various mechanical problems can be concluded to the following two forms : the analytical method and the numeric method. If a problem given, the answers can be got by some derivation, which is called the analytical method. Due to the complexity of the actual structure, it is difficult to analytically solve the majority of scientific researches and engineering problems except for a few very simple questions. Therefore, the numerical methods, such as the finite element method, finite difference method, boundary element method, have become the irreplaceable and widely used methods. Rapidly developed in the mid-20th century, the finite-element method can flexibly analyze and solve various complex problems for the precise mathematical logic, clear physical concept and wide utilization. By adopting the matrix, the finite-element method can conveniently make computer programs to express the basic formula, which makes it vitality.

The essence of the finite-element method is that a complex continuum is divided into a limited number of simple cell body, which transforms the infinite freedom degrees to the limited freedom degree, and the (partial) differential equation to the algebraic equations of finite number of parameters. While analyzing the problems of engineering structure with the finite-element method, after the discretization of an ideal discrete body, it is one of main discussing contents of the finite-element theory how to ensure the convergence and stability of numerical solution. The convergence of numerical solution is related to the element division and shape. During the solving process, as the basic variable, the displacement is usually solved through the virtual displacement or minimum potential energy principle.

The finite-element method is initially applied in the fields of engineering science and technology. As a mathematical and physical method, the finite-element method is used to simulate and solve the problems of engineering mechanics, thermology, electromagnetism and other physical subjects. Moreover, as a numerical analytical method with theoretical foundation and extensive applicable effect, it can solve the problems can't be dealt with by the analytic method. Especially for the complex issues of the irregular boundary conditions or structural shapes, the finite-element method is an effective modern analytical tool. In recent years, the application and development of the finite-element method in the field of bio-mechanics show as follows:

1.1.1 Generation of the Finite-element method

The basic ideas of the finite-element method are usually regarded to begin in the 20th century. In fact, as early as the 3rd century AD, it is the basic expression when Hui Liu, a Chinese mathematician, suggested the method to solve the circumference length by the segmentation unit. According to the thought solving the internal force of structural frame in the classical mechanics, the structural frame is treated as the finite pole elements connected at the junction points in the displacement method, in which each pole element is firstly studied and then finally comprehensively analyze the combination of pole elements. The basic idea of finite-element method is early discretization and later integration.

In 1942, Hrenikoff first proposed the framework method solving mechanical problems, which is only limited to construct the discrete model with the plot series of structure. In 1943,

Courant [1] published a paper to solve the reverse problem by using the polynomial function in the triangular region, in which the flexure function is originally assumed as a simple linear function on each element in a collection of triangular element collection. It is the first time to deal with the continuum problem by the finite element method.

In 1950', due to the rapid development of aviation industry, more accurate design and calculation needed be used in the aircraft structure. In 1955, the professor JH Argyris in the University of Stuttgart, Germany, published a set of papers about energy principles and matrix, which laid the theoretical basis of the finite element method. In 1956, Turner, Clough, Martin, Top, etc, extended the displacement method to the plane problems of elasticity, and use it in the aircraft structural analysis and design. By systematically study the stiffness expression of discrete bar, beam, triangular elements, the correct answers are be obtained for the plane stress problems. Thus their work began the new phase to solve complex elasticity problems using computers.

In 1960, Clough firstly proposed and used the name of "finite element method" while dealing with the section elastic problems, which made people better understand the characteristics and effectiveness of this approach. Since then, a large number of scholars and experts begun to use the discrete approach to deal with the complex issues of structural analysis, fluid analysis, thermal conductivity, electromagnetism, etc. From 1963 to 1964, Besseling, Pian T.H. and other researchers proved that the finite-element method is actually one of means of Rayleigh - Ritz method in the elasticity variational principles, which in theory lays a mathematical foundation for the finite-element method. Compared with the variational principle, the finite-element method is more flexible, adaptable and precise. The results also greatly promoted the research and development of variational principle, and a series of new finite-element models have been emerged based on the variational principle, such as the mixed finite element, non-conforming element, generalized conforming element, etc. In 1967, Zienkiewicz and Cheung published the first book on the finite-element analysis. After the 1970', the finite-element method entered into high-speed period with the development of computer technology and software technology. At that time, the in-depth research was carried out on the finite element method, including the theories in the fields of mathematics and mechanics, principles of element division, the selection of shape functions, numerical methods, error analysis, convergence, stability, computer software development, nonlinear problems, large deformation problems, etc. In 1972, Oden published his first book dealing with non-linear continuum. In the process of development of finite-element method, chinese scientists have also made the outstanding contributions, for example the generalized variational principle proposed by Changhai Hu, the relationship between Lagrange multiplier method and generalized variational principles first studied by Weichang Qian, the accuracy and convergence of finite element method made by Kang Feng, the complementary energy principle by Lingxi Qian, and so on.

With the continuous development and improvement, the finite-element method has become a mature discipline, has been extended to other research fields, and become a powerful tool to solve practical problems for the science and technology researchers.

1.1.2 Basic Theory of the Finite-element method

The finite-element method is a general numerical one dealing with continuum problems. The basic idea of finite-element method is to split up, and then accumulate zero into the whole, which is artificially divided a continuum into finite elements. That is, a structure is

regarded as an integral made up of elements joined by the number of nodes. After the element analyses, the assembled elements is overall analyzed to represent the original structure. From the view of mathematical point, the finite element method is one that turns a partial differential equation into an algebraic equation, and then solved by computer. Using a matrix algorithm, the results can be quickly calculated through the computer.

The basic theory of finite element method is firstly to discretize the continuous computational region into a finite number of unit integration linked together by a certain forms. As the unit can be combined by different connected ways and the unit itself can have different shapes, it can be simulated to solving small region of different geometric shapes, then the units (small area) are carried out on the mechanical analysis, and finally is the overall analysis. That is the basic idea which breaks up the whole into parts and collect parts for the whole. Another important feature of the finite-element method is to express piecewise the unknown field function in the sub-domains by using the hypothetical approximate function in each unit. The approximate function is usually expressed by unknown field function or the data and difference of derivative on each node. Thus, the unknown field function or derivative in each node become the new unknown amounts (the degrees of freedom) while analyzing the problems using the finite-element analysis. So that the issues of continuous unlimited freedom become those of the discrete limited freedom. After solving these unknown variables, the approximation of field function in each element can be calculated by interpolation, and then the approximate solution can be obtained in the entire solution region. With the increasing unit numbers (decreasing unit size), or with the rising element freedom degree and improved accuracy of interpolation function, the solutions approximation is continuously improved. As long as each unit meets the convergence requirements, the approximate solution will eventually converge to the exact solutions.

1.1.3 The characteristics of finite-element method

(1) The adaptability for complex geometry components

As the finite element method of cell division can be one-dimensional, two-dimensional, three-dimensional in space, and can have different shapes, such as two-dimensional elements can be triangular, quadrilateral, 3D element can be tetrahedral, five-sided, six sided, etc. At the same time, the modules can also have different connection forms. Therefore, any complex structure or construction in the practical application can all be separated into a collection made up of finite of modules.

(2) The adaptability for a variety of configuration problems

The finite element method has been developed from the first pole structure to the current plastic, viscoelastic-plastic and power issues of problems, which can be used to solve various complex non-linear problems in the fluid mechanics, thermodynamics, electromagnetism, aerodynamics.

(3) The reliability of theoretical basis

The theory of finite element method is based on variational and energy-conservation principles, which have been reliably proven in mathematics and physics. As soon as the mathematical model is appropriately set up and the algorithm of finite element equation is stable convergence, the solution obtained are authentic.

(4) The credibility of accuracy

As long as the research question itself is solvable, the accuracy of finite element method will

be continually increased as the cell number increases in the same conditions, then the approximate solutions will keep close to the exact ones.

(5) The efficiency of calculation

As each step of the finite-element analysis can be expressed in matrix form, the final solutions come down to solve the standard matrix algebra problem, which particularly suitable for computer programming calculations by turning a number of complex differential, partial differential equations into solving algebraic equations.

1.2 The analysis process of finite-element method

1.2.1 The discretization of structure

The first step of a finite element analysis is to decentralize the structure, that is, to divide the entire structure into a finite element based on different accuracy requirements, performance requirements and other factors. The positions between elements or element and boundary are connected by nodes.

During the discretization, we must pay more attention on the following three points: the choice of element types, including element shape, node number, and node number of freedom, a certain regularity of element partition, in order to compute the network of automatic generation, and then conducive to encrypt the network. And the same element should be composed of the same material.

1.2.2 The element analysis

The element analysis regards each discretized element as a research object. The relationship between node displacement and the nodal force includes the following two aspects:

(1) Determine the displacement mode of element

For the displacement finite-element method, the displacement model of element is to calculate the displacement at any point with the node displacement of element, in which the element displacement can be expressed as a function of node displacement. It directly affects the accuracy, efficiency and reliability of finite element analysis whether or not a reasonable assumption of displacement function.

(2) Analyze the element characteristics

After establishing the displacement function of element, according to the relationship among stress, strain and displacement, the relationship between the element rod end force and displacement of rod end and then the element stiffness matrix are obtained by using the virtual displacement and minimum potential energy principles. Here the load on the element must be equivalent to the node load. Thus the element analysis is in fact the process to set up the element stiffness and equivalent nodal load matrixes.

1.2.3 The overall analysis

After determining the stiffness equation of each element, you can set the various units of a whole structure, the node balance equations of entire structure are established, that is the overall stiffness equation, through analyzing the whole structure gathered by each element. After the introduction of structural boundary conditions, the equations are solved to reach the node displacement, and then the internal force and deformation of each element.

1.3 Steps to solve the finite element method

When any continuum are solved with the finite element method, the solution continuous

area should be divided into a finite number of elements, and a finite number of nodes are specified in each element. Generally, the adjacent elements can be considered to constitute a group of element by in the nodes, and then form the collection which is used to simulate or approximate the solving region. At the same time, the node data are selected in the field function, such as the node displacement regarded as the basic unknown quantity. By the block approximation to each element, a simple function (called interpolation function) can approximately express the assuming the displacement distribution. Using the variational principle or other methods, a set of algebraic equations can be obtained in which the nodal displacements is regarded as the unknowns by establishing the mechanical properties relationship between force and displacement of element nodes. Thereby the node displacement components can be solved. The approximate function in the element is usually expressed by the unknown field function or each node data of field functions and their interpolation function. Thus, in the finite element analysis, the unknown field function or derivative values of each node in the element becomes a new unknown quantity (ie degrees of freedom). Then a continuous problem of infinite freedom degrees becomes the discretized problem of limited freedom degrees. Upon solving these unknowns, the approximation of field function in each element can be calculated by interpolation function, and the approximate solution in the overall solving domain can be obtained. Obviously, if the element meet the convergence demands of problem, the approximate level of solution will be improved with the shrinking element size, increasing element numbers in the solving region, rising element freedom degree or improving accuracy of interpolation function, and finally the approximate solution will eventually converge to exact solutions.

While solving the mechanics problems by using the finite element method, the basic unknowns mustn't be given with the node displacement but the node force. Therefore, for the different taken basic unknown quantity, there are the so-called displacement, hybridization and mixing method, in which the displacement method is the most common. The ANSYS, well-known large-scale commercial software, adopts the finite element displacement method [2,3].

1.4 The basic steps of the finite element method [4]

(1) Discretization of objects and selected element type

The core of the finite element method is discretization. For each specific issue, the specific content of discretization is how to choose the appropriate element type to determine element size, number, layout and sequential manner of nodes. Element size is small enough to ensure the accuracy of calculation, and element size is large enough to reduce the computation workload. In theory, the finer elements, the more nodes can be arranged, and more accurate the results are. Today, the computer's capacity and speed are not the principal contradiction. However, it shows that the grid encrypt is not effective for improving the calculation accuracy after the numbers of nodes and elements reach a certain value. The general principle divided into elements: the element nodes should be those of intersection, turning, support and sectional mutations points in the structure.

(2) The analysis of element

The element analysis is to solve the relationship between the nodal displacement of basic unknown quantity and the corresponding nodal forces. For the element, the Nodal forces are the external forces on the elements through the nodes, which determines not only the displacement of this node, but also the effects of other nodes in this element. After

determining the element displacement, the stress and strain of elements can be easily obtained by the geometric and physical equations. In general, the stiffness matrix can express the element properties, and then the relationship between the element nodal force and displacement can be determined.

(3) The overall analysis

The overall analyses include the combination of whole stiffness matrix and the establishment of balance equations. The whole stiffness matrix is made up of the element stiffness matrix, and each item contains all corresponding information of relevant elements of nodes.

(4) The introduction of support conditions and total load of nodes

The support conditions are constraint ones. Suspended structures without constraint can't support the loads. The total loads of nodes include ones acted on nodes and the non-node ones of equivalent transplant. The node load vectors are consistent with node force vectors of overall stiffness equations. After changing the node force vector of overall stiffness equations to the node load vectors, the node balance equations are established and the nodal displacements as unknowns.

(5) The nodal displacement by solving the finite-element equations

All unknown nodal displacements can be solved for solving algebraic equations with the finite-element software.

(6) Return to the results of element calculation

The element strain can be solved by the nodal displacements, and the element stress can be reversed through the physical equations.

1.5 The finite-element method is the synthesis of a variety of methods [4]

At first, the finite element method is a kind of model one. The model is a kind of simplified description on a simplified understanding of the objects' for a particular purpose. By studying the prototype model, it reveals the morphology, characteristics and nature of objects. An important feature of the finite element method is to express the unknown function to be calculated in the overall solution domain by using the approximate function supposed in each element. Then there is an approximation problem between approximate and real solutions. Obviously, the approximation degree of solution is higher with the increasing element numbers, that is the shrinking element size, or with the rising element freedom degree, that is the improved interpolation function. Secondly, the finite-element method is the result about the integrated application of the analysis and comprehensive methods. The first part work of finite-element method is to discretize the study object into many simple shapes and easy description. According to the special circumstances, the object can be discretized to the two-dimensional or three-dimensional finite element collections, in which the discretized elements have simple and regular geometry easy to calculate. For the two-dimensional elements, there are normally triangular or rectangular. For three-dimensional elements, there are commonly tetrahedral or parallelepiped. For the same shape of element, there are different nodes, so many different kinds of elements. According to various analysis object and purpose, the elements selected are different.

1.6 The application of finite-element method in bio-mechanics

It has made great achievements that the finite-element method is applied the life science to carry out the quantitative research. In particular, its great superiority is demonstrated in the

research on human biomechanics. After a long evolution of human labor, the human skeleton has almost formed a perfect mechanical structure. However, the mechanics experiment is almost impossible to directly develop while the mechanical structure of human body is studied. At that time, it is an effective means that the finite element numerically simulates the mechanical experiments.

In 1960's, the finite element method was initially applied in the cardiovascular system study on mechanical problems. From the 1970's, the orthopedic biomechanics research began to be initially applied to spine. After 80 years, the applicable range is gradually extended to craniofacial bone, mandible bone, femur[5], teeth[6], joint[7, 8], cervical [9], lumbar [10] and its subsidiary structure in the bio-mechanics fields.

(1) Improvement and optimization design of equipment

The finite element analysis can improve the mechanical properties of medical devices and optimize the design of equipment. The mechanical properties of medical equipment often determine the effectiveness of the clinical application. Therefore it is more important to evaluate the mechanical properties of appraisal instrument. In addition to the experimental methods, the simulative experiment to the mechanical equipment with the finite-element method has the advantages of short time, cost less, handling complex conditions, comprehensive mechanical properties and good repetitiveness, etc. In addition, the finite-element method can also optimize the design, guide the design and improvement of medical devices, and obtain better clinical effects.

(2) The mechanical simulation experiment by finite element model

The powerful modeling capabilities and interface tool if the finite element software can distinctly build the models of three-dimensional human bone, muscle, blood vessels and other organs, and be able to give their biomechanical properties. In the simulated experiments, the model simulation to experimental conditions and mechanical tests to simulate the stretching, bending, torsion, anti-fatigue can solve the deformation, stress, strain distribution, the internal energy change and the ultimate destruction situation on different experimental conditions [11].

(3) Nowadays, the finite element method has been widely used in the country, and has made a lot of successes. Particularly, there are more certain guidances in the clinical application. In bio-mechanics the finite-element application, there are a large number of spaces to research the shape and structure of finite element model.

2. Mechanical Analysis of cranial cavity deformation

2.1 Mechanical Analysis of deformation of the skull as a whole

There are two aspects of effects on external forces on the objects. One can make objects produce the acceleration, another is make objects deform. In discussing the external force effect, the objects are assumed to be a rigid body not compressed. But in fact, all objects will deform under loading, but different with the degrees. Here, we will mechanically analyze the overall deformation of cranial cavity under the external force.

(1) Two basic assumptions

To simplify the analysis of deformation of human skull, we assume:

1) Uniformly-continuous materials

The human skull is presumed to be everywhere uniform, and the sclerotin is no gap in bone of cranial cavity.

2) The isotropy

The human skull is supposed to have the same mechanical properties in all directions. The thickness and curvature of human skull vary here and there. The external and internal boards are all compact bones, in which external board is thicker than internal board but the radius of external board is smaller than that of the internal board. The diploe is the cancellous bone between the external board and the internal board, which consists of the marrow and diploe vein. The parietal bone is the transversely isotropic material⁸, namely, it has the mechanical property of rotational symmetry in any axially vertical planes of skull⁹. The tensile and compressive abilities of compact bone are strong. The important mechanical characteristics of cancellous bone are viscoelasticity¹⁰, which is generally considered as the construction of semi-closed honeycomb composed of bone trabecula reticulation. The main composition of cerebral duramater, a thick and tough bilayer membrane, is collagenous fiber¹¹, which is viscoelastic material¹². And the thickness of duramater is obviously variable with the changing ICP¹³. The mechanical performance of skull is isotropic along tangential direction on the skull surface¹⁴, in which the performance of compact bone in the external board is basically the same as that in the internal board¹⁵, thus both cancellous bone and duramater can be regarded as isotropic materials¹⁶. And the elastic modulus of fresh duramater is variable with delay time¹⁷. And there are a number of sutures in the cranial cavity. But while a partial skull is studied on the local deformation, we can regard each partial skull as quasi-homogeneous and quasi-isotropic. In addition, the sutures of cranial cavity are also the continuous integration with ages.

(2) Two concepts

(1) stress (σ)

Stress is the internal force per unit area. The calculation unit is kg/cm^2 or kg/mm^2 .

2) Transverse deformation coefficient

When the objects are under tension, not only the length is drawn out from l to l_1 , but also the width is reduced from b to b_1 . This shows that there are the horizontal compressive stresses in the objects. Similarly, while the object is compressed, not only its length shortens but its width increases. It indicates that the horizontal tensile stress distributes in the objects.

The horizontal absolute deformation is noted as $\Delta b = b - b_1$, and the transverse strain is $\varepsilon_0 = \Delta b/b$.

In mechanics of materials, the transverse strain ε_0 is proportional to the longitudinal strain ε of the same material within the scope of the Hooke theorems' application. The ratio of its absolute value $\mu = |\varepsilon_0|/|\varepsilon|$ is a constant. μ is known as the coefficient of lateral deformation, or Poisson ratio. The Poisson ratio of any objects can be detected by the experiment.

While vertically compressed, the objects simultaneously have a horizontal tension. Therefore, when the head attacked in opposite directions force, the entire human skull will take place the longitudinal compression and transverse tension with the same direction of force. Then the longitudinal compressive stress and the horizontal tensile stress will be generated in the sclerotin. Thus, the stress of arbitrary section along radial direction in shell

is just equal to the tangential pulling force along direction perpendicular to the normal vertical when ICP is raised.

2.2 The finite-element analysis of strains by ignoring the viscoelasticity of cranial cavity

The geometric shape of human skull is irregular and variable with the position, age, gender and individual. So the cranial cavity system is very complex. Moreover, the cranial cavity is a kind of viscoelastic solid with large elastic modulus, and the brain tissue is also a viscoelastic fluid with great bulk modulus. It is now almost impossible to accurately analyze the brain system. Only by some simplification and assumptions, the complex issues can be made. Considering the special structure of cranial cavity composed of skull, duramater, encephalic substance, etc, here we simplify the model of cranial cavity as a regular geometry spheroid of about 200mm external diameter, which is an hollow equal-thickness thin-wall shell.

The craniospinal cavity may be considered as a balloon. For the purpose of our analysis, we adopted the model of hollow sphere. We presented the development and validation of a 3D finite-element model intended to better understand the deformation mechanisms of human skull corresponding to the ICP change. The skull is a layered sphere constructed in a specially designed form with a Tabula externa, Tabula interna, and a porous Diploe sandwiched in between. Based on the established knowledge of cranial cavity importantly composed of skull and dura mater (Fig2.1), a thin-walled structure was simulated by the composite shell elements of the finite-element software [18]. The thickness of skull is 6mm, that of duramater is 0.4mm, that of external compact bone is 2.0mm, that of cancellous bone is 2.8mm, and that of internal compact bone is 1.2mm.

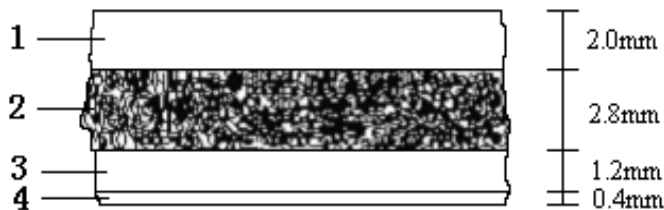


Fig. 2.1 Sketch of layered sphere. The thin-walled structure of cranial cavity is mainly composed of Tabula externa, Diploe, Tabula interna and dura mater.

Above all, we should prove the theoretical feasibility of the strain-electrometric method to monitor ICP. We simplify the theoretical calculation by ignoring the viscoelasticity of cancellous bone and dura mater. And then we make the analysis of the actual deformation of cranial cavity by considering the viscoelasticity of human skull-dura mater system with the finite-element software. At the same time, we can determine how the viscoelasticity of human skull and dura mater influences the strains of human skull respectively by ignoring and considering the viscoelasticity of human skull and dura mater.

2.2.1 The stress and strain analysis of discretized elements of cranial cavity

In order to obtain the numerical solution of the skull strain, the continuous solution region of cranial cavity divided into a finite number of elements, and a group element collection

glued on the adjacent node points. Then the large number of cohesive collection can be simulated the overall cranial cavity to carry out the strain analysis in the solving region. Based on the block approximation ideas, a simple interpolation function can approximately express the distribution law of displacement in each element. The node data of the selected field function, the relationship between the nodal force and displacement is established, and the algebraic equations of regarding the nodal displacements as unknowns can be formed, thus the nodal displacement components can be solved. Then the field function in the element collection can be determined by using the interpolation function. If the elements meet the convergence requirements, with the element numbers increase in the solving region with the shrinking element size, and the approximate solution will converge to exact solutions [12].

The solving steps for the strains of cranial cavity with the ICP changes are shown in Fig2.2. The specific numerical solution process is:

1) The discretized cranial cavity

The three-dimensional hollow sphere of cranial cavity is divided into a finite number of elements. By setting the nodes in the element body, an element collection can replace the structure of cranial cavity after the parameters of adjacent elements has a certain continuity.

2) The selection of displacement mode

To make the nodal displacement express the displacement, strain and stress of element body, the displacement distribution in the elements are assumed to be the polynomial interpolation function of coordinates. The items of polynomial number are equal to the freedom degrees number of elements, that is, the number of independent displacement of element node. The orders of polynomial contain the constant term and linear terms.

According to the selected displacement mode, the nodal displacement is derived to express the displacement relationship of any point in the elements. Its matrix form is:

$$\{f\} = [N]\{\delta\}^e \quad (2.1)$$

Where: $\{f\}$ - The displacement array of any point within the element;

$[N]$ - The shape function matrix, its elements is a function of location coordinates;

$\{\delta\}^e$ - The nodal displacement array of element.

The block approximation is adopted to solve the displacement of cranial cavity in the entire solving region, and an approximate displacement function is selected in an element, where need only consider the continuity of displacement between elements, not the boundary conditions of displacement. Considering the special material properties of the middle cancellous and duramater, the approximate displacement function can adopt the piecewise function.

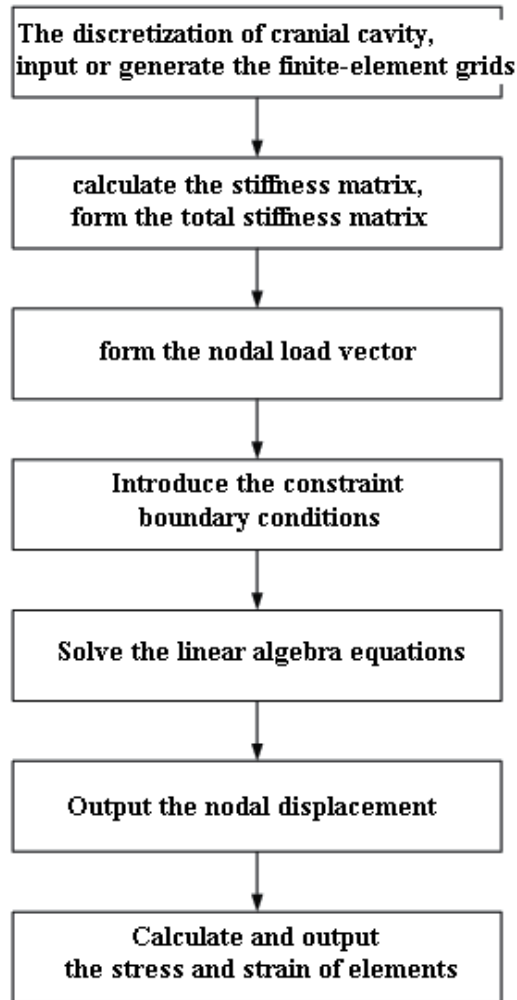


Fig. 2.2 Block diagram of numerical solution steps of cranial cavity with the finite-element method

3) Analyze the mechanical properties of elements, and derive the element stiffness matrix

a. Using the following strain equations, the relationship of element strain (2.2) is expressed by the nodal displacements derived from the displacement equation (2.1):

$$\text{Strain equations} \left\{ \begin{array}{l} \varepsilon_x = \frac{\partial u}{\partial x}, \gamma_{xy} = \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \\ \varepsilon_y = \frac{\partial u}{\partial y}, \gamma_{yz} = \frac{\partial u}{\partial z} + \frac{\partial w}{\partial y} \\ \varepsilon_z = \frac{\partial u}{\partial z}, \gamma_{zx} = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial z} \end{array} \right.$$

$$\{\varepsilon\} = [B]\{\delta\}^e \quad (2.2)$$

Where : $[B]$ — The strain matrix of elements ;

$\{\varepsilon\}$ — The strain array at any points within the elements.

b. The constitutive equation reflecting the physical characteristics of material is $\{\sigma\} = [D]\{\varepsilon\}$, so the stress relationship of stress can be expressed with the nodal displacements derived from the strain formula (2.2):

$$\{\sigma\} = [D][B]\{\delta\}^e \quad (2.3)$$

Where: $\{\sigma\}$ - The stress array of any points in the elements;

$[D]$ - The elastic matrix related to the element material.

c. Using the variational principle, the relationship between force and displacement of the element nodes is established:

$$\{F\}^e = [k]^e \{\delta\}^e \quad (2.4)$$

Where: $[k]^e$ - Element stiffness matrix, $[k]^e = \iiint [B]^T [D] [B] dx dy dz$;

$\{F\}^e$ - Equivalent nodal force array, $\{F\}^e = \{F_V\}^e + \{F_S\}^e + \{F_C\}^e$;

$\{F_V\}^e$ - Equivalent nodal force on the nodes transplanted from the element volume force P_V , $\{F_V\}^e = \iiint_V [N]^T \{P_V\} dV$;

$\{F_S\}^e$ - Equivalent nodal force on the nodes transplanted from the element surface force, $\{F_A\}^e = \iint_A [N]^T \{P_S\} dA$;

$\{F_C\}^e$ - Concentration force of nodes.

4) Collecting all relationship between force and displacement, and establish the relationship between force and displacement of cranial cavity

According to the displacement equal principle of the public nodes in all adjacent elements, the relationship between force and displacement of overall cranial cavity collected from the element stiffness matrix:

$$\{F\} = [K]\{\delta\} \quad (2.5)$$

Where: $\{F\}$ - Load array;

$[K]$ - The overall stiffness matrix;

$\{\delta\}$ - The nodal displacement array of the entire cranial cavity.

5) Solve the nodal displacement

After the formula (2.1) ~ (2.5) eliminating the stiffness displacement of geometric boundary conditions, the nodal displacement can be solved from the gathered relationship groups between force and displacement.

6) By classifying the nodal displacement solved from the formula (2.2) and (2.3), the strain

and stress in each element can be calculated.

In this paper, the studied cranial cavity is a hollow three-dimensional sphere, its external radius $R = 100 \text{ mm}$, the curvature of hollow shell is 0.01 rad/mm , the thickness of shell wall is 6 mm , so the element body of hollow spherical can be treated as the regular hexahedron. The following is the stress and strain analyses in the three-dimensional elements in the cranial space. The 8-node hexahedral element (Fig1.3) is used to be the master element. The origin is set up as the local coordinate system (ξ, η, ζ) in the element. Trough the transformation between rectangular coordinates and local coordinates, the space 8-node isoparametric centroid element can be obtained. The relationship of coordinate transformation is:

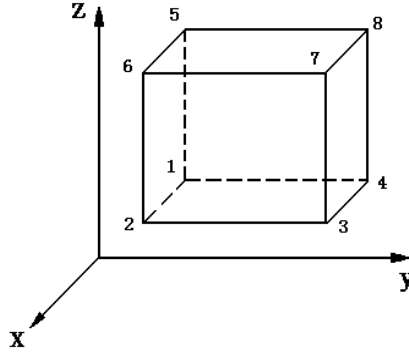


Fig. 2.3 The space 8-node isoparametric centroid element

$$\begin{cases} x = \sum_{i=1}^8 N_i(\xi, \eta, \zeta) x_i \\ y = \sum_{i=1}^8 N_i(\xi, \eta, \zeta) y_i \\ z = \sum_{i=1}^8 N_i(\xi, \eta, \zeta) z_i \end{cases} \quad (2.6)$$

Then the displacement function of element is:

$$\begin{cases} u = \sum_{i=1}^8 N_i(\xi, \eta, \zeta) u_i \\ v = \sum_{i=1}^8 N_i(\xi, \eta, \zeta) v_i \\ w = \sum_{i=1}^8 N_i(\xi, \eta, \zeta) w_i \end{cases} \quad (2.7)$$

Where: x_i 、 y_i 、 z_i and u_i 、 v_i 、 w_i are respectively the coordinate values and actual displacement of nodes.

The element displacement function with matrix is expressed as:

$$\{\delta\} = \begin{Bmatrix} u \\ v \\ w \end{Bmatrix} = \sum_{i=1}^8 \begin{bmatrix} N_i & 0 & 0 \\ 0 & N_i & 0 \\ 0 & 0 & N_i \end{bmatrix} \begin{Bmatrix} u_i \\ v_i \\ w_i \end{Bmatrix} = \sum_{i=1}^8 [N_i] \{\delta_i\} = [N] \{\delta\}^e \quad (2.8)$$

Where: $\{\delta_i\}$ - Nodal displacement array, $\{\delta_i\} = [u_i \quad v_i \quad w_i]^T$ ($i=1,2,\dots,8$);

$\{\delta\}^e$ - The nodal displacement array of entire element,
 $\{\delta\}^e = [\{\delta_1\} \quad \{\delta_2\} \quad \dots \quad \{\delta_8\}]^T$;

N_i - The uniform shape function of 8 nodes, which can be expressed as:

$$N_i = \frac{1}{8} (1 + \xi_i \xi) (1 + \eta_i \eta) (1 + \zeta_i \zeta) \quad (i=1,2,\dots,8) \quad (2.9)$$

Where: ξ_i 、 η_i 、 ζ_i is the coordinates of node i in the local coordinate system (ξ, η, ζ) .

The derivative of composite function to local coordinates is:

$$\begin{cases} \frac{\partial N_i}{\partial \xi} = \frac{1}{8} \xi_i (1 + \eta_i \eta) (1 + \zeta_i \zeta) \\ \frac{\partial N_i}{\partial \eta} = \frac{1}{8} \eta_i (1 + \xi_i \xi) (1 + \zeta_i \zeta) \\ \frac{\partial N_i}{\partial \zeta} = \frac{1}{8} \zeta_i (1 + \xi_i \xi) (1 + \eta_i \eta) \end{cases} \quad (2.10)$$

The strain relationship of space elements is:

$$\{\varepsilon\} = [B] \{\delta\}^e = \sum_{i=1}^8 [B_i] \{\delta_i\} \quad (2.11)$$

The strain matrix $[B]$ of space element:

$$[B_i] = \begin{bmatrix} \frac{\partial N_i}{\partial x} & 0 & 0 \\ 0 & \frac{\partial N_i}{\partial y} & 0 \\ 0 & 0 & \frac{\partial N_i}{\partial z} \\ \frac{\partial N_i}{\partial y} & \frac{\partial N_i}{\partial x} & 0 \\ 0 & \frac{\partial N_i}{\partial z} & \frac{\partial N_i}{\partial y} \\ \frac{\partial N_i}{\partial z} & 0 & \frac{\partial N_i}{\partial x} \end{bmatrix} \quad (2.12)$$

The shape function was derivative to be:

$$\begin{Bmatrix} \frac{\partial N_i}{\partial x} \\ \frac{\partial N_i}{\partial y} \\ \frac{\partial N_i}{\partial z} \end{Bmatrix} = [J]^{-1} \begin{Bmatrix} \frac{\partial N_i}{\partial \xi} \\ \frac{\partial N_i}{\partial \eta} \\ \frac{\partial N_i}{\partial \zeta} \end{Bmatrix} \quad (2.13)$$

The matrix $[J]$ is the three-dimensional Yake ratio matrix of coordinate transformation:

$$[J] = \begin{Bmatrix} \frac{\partial x}{\partial \xi} & \frac{\partial y}{\partial \xi} & \frac{\partial z}{\partial \xi} \\ \frac{\partial x}{\partial \eta} & \frac{\partial y}{\partial \eta} & \frac{\partial z}{\partial \eta} \\ \frac{\partial x}{\partial \zeta} & \frac{\partial y}{\partial \zeta} & \frac{\partial z}{\partial \zeta} \end{Bmatrix} = \begin{Bmatrix} \sum_{i=1}^8 \frac{\partial N_i}{\partial \xi} x_i & \sum_{i=1}^8 \frac{\partial N_i}{\partial \xi} y_i & \sum_{i=1}^8 \frac{\partial N_i}{\partial \xi} z_i \\ \sum_{i=1}^8 \frac{\partial N_i}{\partial \eta} x_i & \sum_{i=1}^8 \frac{\partial N_i}{\partial \eta} y_i & \sum_{i=1}^8 \frac{\partial N_i}{\partial \eta} z_i \\ \sum_{i=1}^8 \frac{\partial N_i}{\partial \zeta} x_i & \sum_{i=1}^8 \frac{\partial N_i}{\partial \zeta} y_i & \sum_{i=1}^8 \frac{\partial N_i}{\partial \zeta} z_i \end{Bmatrix} \quad (2.14)$$

The stress-strain relationship of space elements is:

$$\{\sigma\} = [D]\{\varepsilon\} = [D][B]\{\delta\}^e \quad (2.15)$$

The elasticity matrix $[D]$ is:

$$[D] = \frac{E(1-\mu)}{(1+\mu)(1-2\mu)} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ \frac{\mu}{1-\mu} & 1 & 0 & 0 & 0 & 0 \\ \frac{\mu}{1-\mu} & \frac{\mu}{1-\mu} & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1-2\mu}{2(1-\mu)} & \frac{\mu}{1-\mu} & \frac{\mu}{1-\mu} \\ 0 & 0 & 0 & 0 & \frac{1-2\mu}{2(1-\mu)} & \frac{\mu}{1-\mu} \\ 0 & 0 & 0 & 0 & 0 & \frac{1-2\mu}{2(1-\mu)} \end{bmatrix} \quad (2.16)$$

The element stiffness matrix from the principle of virtual work is:

$$[k]^e = \iiint_V [B]^T [D][B] dx dy dz = \begin{bmatrix} k_{11} & k_{12} & \cdots & k_{18} \\ k_{21} & k_{22} & \cdots & k_{28} \\ & & \cdots & \\ k_{81} & k_{82} & \cdots & k_{88} \end{bmatrix} \quad (2.17)$$

Where:

$$[k_{ij}] = \iiint_V [B]^T [D][B] dx dy dz = \int_{-1}^1 \int_{-1}^1 \int_{-1}^1 [B]^T [D][B] J d\xi d\eta d\zeta \quad (2.18)$$

The equivalent nodal forces acting on the space element nodes are:

$$\{F\}^e = [k]^e \{\delta\}^e \quad (2.19)$$

Because the internal pressure in the cranial cavity is surface force, the equivalent load for the pressure acting on the element nodes is:

$$\{F_S\}^e = \iint_S [N]^T \{P_S\} dS \quad (2.20)$$

The relationship between force and displacement in the entire cranial cavity is:

$$\{F\} = [K]\{\delta\} \quad (2.21)$$

Then after obtaining the nodal displacement, the stress and strain in each element can be calculated by combining the formula (2.11) and (2.15).

2.2.2 The stress and strain analysis for complex structure of cranial cavity deformation

Cranial cavity is the hollow sphere formed by the skull and the duramater. From the Fig.1.2, there are obvious interfaces among the various parts of outer compact bone, middle cancellous bone, inner compact bone and duramater, which is consistent with the characteristics of composite materials [13]. Four layered composite structure of cranial cavity is almost lamelleted distribution. Therefore, the lamelleted structure is adopted to establish and analyze the finite-element model of cranial cavity, and the laminated shell element is used to describe the thin cranial cavity made up of skull and duramater. Here the cranial cavity deformation of laminated structure is analyzed as follows:

(1) The stress and strain analysis for the single layer of cranial deformation

Each layers of cranial cavity are all thin flat film. The skulls are transversely isotropic material. The thickness of Tabula externa, diploe, Tabula interna, duramater is all very small. So compared with the components in the surface, the stress components are very small along the normal direction, and can be neglected. So the deformation analysis to single-layer cranial cavity can be simplified to be the stress problems of two-dimensional generalized plane.

The stress-strain relationship of each single-layer structure in the cranial cavity:

$$\{\sigma\} = [Q]\{\varepsilon\} \quad (2.22)$$

Namely:

$$\begin{Bmatrix} \sigma_1 \\ \sigma_2 \\ \tau_{12} \end{Bmatrix} = \begin{bmatrix} Q_{11} & Q_{12} & 0 \\ Q_{21} & Q_{22} & 0 \\ 0 & 0 & Q_{66} \end{bmatrix} \begin{Bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \mu_{12} \end{Bmatrix} \quad (1.23)$$

Where, 1,2 - The main direction of elasticity in the plane;

$[Q]$ - Stiffness matrix, $Q_{11} = mE_1$, $Q_{12} = m\mu_{12}E_2$, $Q_{22} = mE_2$, $Q_{66} = G_{12}$;

$$m = \left(1 - \frac{\mu_{12}^2 E_2}{E_1} \right)^{-1};$$

E_1 、 E_2 - The elastic modulus of four independent surfaces in each layer structure;

G_{12} - Shear modulus;

μ_{12} - Poisson's ratio of transverse strain along the 2 direction that the stress acts on the 1 direction.

(2) The stress and strain analysis for the laminated deformation of cranial cavity

The cranial cavity is as a whole formed by the four-layer structures. So the material, thickness and elastic main direction are all different. The overall performance of cranial cavity is anisotropic, macroscopic non-uniformity along the thickness direction and non-continuity of mechanical properties. Thus, the assumptions need to be made before analyzing the overall deformation of cranial cavity [14]:

1) The same deformation in each layer

Each single layer is strong glued. There are the same deformation, and no relative displacement;

2) No change of direct normal

The straight line vertical to the middle surface in each layer before the deformation remains still the same after the deformation, and the length of this line remains unchanged whether before or after deformation;

3) $\sigma_z = 0$

The positive stress along the direction of thickness is small compared to other stress, and can be ignored;

4) The plane stress state in each single layer

Each single-layer structure is similar to be assumed in plane stress state.

From the four-layer laminated structure composed of Tabula externa, diploe, Tabula interna, duramater, the force of each single-layer structure is indicated in Fig2.4. The middle surface in the laminated structure of cranial cavity is the xy coordinate plane. z axis is perpendicular to the middle surface in the plate. Along the z axis, each layer in turn will be compiled as layer 1, 2, 3, and 4. The corresponding thickness is respectively t_1 , t_2 , t_3 , t_4 .

As a overall laminated structure, the thickness of cranial cavity is h , shown in Fig2.5.

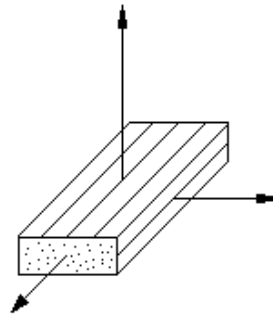


Fig. 2.4 The orientation relationship in each single-layer structure of cranial cavity

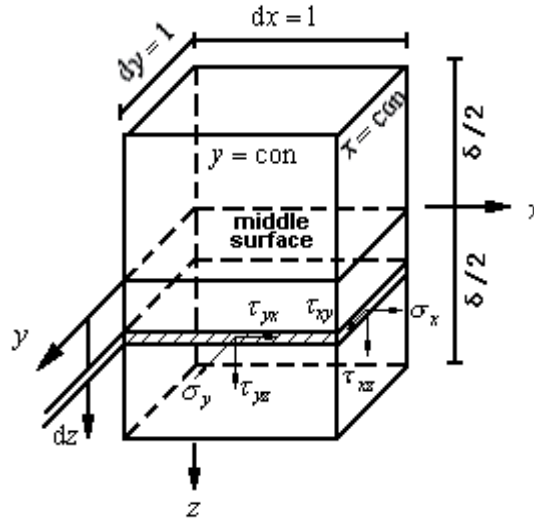


Fig. 2.5 The sketch of four-layered laminated structure of cranial cavity

Then

$$h = \sum_{k=1}^4 t_k \quad (2.24)$$

The z coordinates is respectively z_{k-1} and z_k , then $z_0 = -h/2$ and $z_4 = h/2$.

The displacement components at any point within the laminated structure of cranial cavity:

$$\left. \begin{aligned} u &= u(x, y, z) \\ v &= v(x, y, z) \\ w &= w(x, y, z) \end{aligned} \right\} \quad (2.25)$$

The strain is:

$$\left\{ \begin{aligned} \varepsilon_x &= \frac{\partial u}{\partial x} = \frac{\partial u_0}{\partial x} - z \frac{\partial^2 w}{\partial x^2} \\ \varepsilon_y &= \frac{\partial v}{\partial y} = \frac{\partial v_0}{\partial y} - z \frac{\partial^2 w}{\partial y^2} \\ \gamma_{xy} &= \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} = \frac{\partial u_0}{\partial y} + \frac{\partial v_0}{\partial x} - 2z \frac{\partial^2 w}{\partial x \partial y} \end{aligned} \right. \quad (2.26)$$

Where: $u = u(x, y, z)$, $v = v(x, y, z)$, $w = w(x, y, z)$ - The displacement components at any point within the cranial cavity;

$u_0(x, y)$, $v_0(x, y)$ - The displacement components in the middle surface;

$w(x, y)$ - Deflection function, the deflection function of each layer is the same.

Fomular (2.26) can be expressed to be in matrix form:

$$\{\varepsilon\} = \{\varepsilon_0\} + z\{k\} \quad (2.27)$$

Where: ε_0 - Strain array in the plane, $\{\varepsilon_0\} = \left[\frac{\partial u_0}{\partial x}, \frac{\partial v_0}{\partial y}, \left(\frac{\partial u_0}{\partial y} + \frac{\partial v_0}{\partial x} \right) \right]^T$;

k - Strain array of bending in the surface, $\{k\} = \left[-\frac{\partial^2 w}{\partial x^2}, -\frac{\partial^2 w}{\partial y^2}, -2\frac{\partial^2 w}{\partial x \partial y} \right]^T$.

The mean internal force and torque acting on the laminated structure of cranial cavity in the unit width is:

$$\begin{cases} N_i = \int_{-h/2}^{h/2} \sigma_i dz \\ M_i = \int_{-h/2}^{h/2} \sigma_i z dz \end{cases} \quad (i = x, y, xy) \quad (2.28)$$

Taking into account the discontinuous stress caused by the discontinuity along the direction of laminated structure in the cranial cavity, the formula (2.28) can be rewritten in matrix form:

$$\begin{cases} \{N\} = \sum_{k=1}^n \int_{z_{k-1}}^{z_k} \{\sigma\} dz \\ \{M\} = \sum_{k=1}^n \int_{z_{k-1}}^{z_k} \{\sigma\} z dz \end{cases} \quad (2.29)$$

After substituting the formula (2.22) and (2.27) into equation (2.29), the average internal force and internal moment of the laminated structure in the cranial cavity is:

$$\begin{Bmatrix} N \\ M \end{Bmatrix} = \begin{bmatrix} \sum \int [Q] dz & \sum \int [Q] z dz \\ \sum \int [Q] z dz & \sum \int [Q] z^2 dz \end{bmatrix} \begin{Bmatrix} \varepsilon_0 \\ k \end{Bmatrix} = \begin{bmatrix} A & B \\ B & D \end{bmatrix} \begin{Bmatrix} \varepsilon_0 \\ k \end{Bmatrix} \quad (2.30)$$

Where: $[A]$ - The stiffness matrix in the plane, $A_{ij} = \sum_{k=1}^n Q_{ij}^{(k)} (z_k - z_{k-1})$;

$[B]$ - Coupling stiffness matrix, $B_{ij} = \frac{1}{2} \sum_{k=1}^n Q_{ij}^{(k)} (z_k^2 - z_{k-1}^2)$;

$[D]$ - Bending stiffness matrix, $D_{ij} = \frac{1}{3} \sum_{k=1}^n Q_{ij}^{(k)} (z_k^3 - z_{k-1}^3)$, $(i, j = 1, 2, 6)$.

Then the flexibility matrix of laminated structure in the cranial cavity is:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} A & B \\ B & D \end{bmatrix}^{-1} \quad (2.31)$$

Where: $[a] = [A]^{-1} + [A]^{-1} [B] ([D] - [B] [A]^{-1} [B])^{-1} [B] [A]^{-1}$

$[b] = -[A]^{-1} [B] ([D] - [B] [A]^{-1} [B])^{-1}$

$$\begin{aligned} [c] &= -([D] - [B][A]^{-1}[B])^{-1} [B][A]^{-1} = [b]^T \\ [d] &= ([D] - [B][A]^{-1}[B])^{-1} \end{aligned}$$

Thus, the stress-strain relationship of laminated structure in the cranial cavity is:

$$\begin{Bmatrix} \varepsilon_0 \\ k \end{Bmatrix} = \begin{bmatrix} a & b \\ b^T & d \end{bmatrix} \begin{Bmatrix} N \\ M \end{Bmatrix} \quad (2.32)$$

With the changing ICP, to determine how the viscoelasticity of human skull and duramater influences the strains of human skull respectively by ignoring and considering the viscoelasticity of human skull and duramater, we make the analysis of the actual deformation of cranial cavity by considering the viscoelasticity of human skull-duramater system with the finite-element software MSC_PATRAN/NASTRAN and ANSYS. Considering the complexity to calculate the viscoelasticity of human skull and duramater, we can simplify the calculation while on-line analysis only considering the elasticity but ignoring the viscoelasticity of human skull and duramater after obtaining the regularity how the viscoelasticity influence the deformation of cranial cavity.

2.2.3 The finite-element analysis of strains by ignoring the viscoelasticity of human skull and duramater

The craniospinal cavity may be considered as a balloon. For the purpose of our analysis, we adopted the model of hollow sphere (Fig2.6). We presented the development and validation of a 3D finite-element model intended to better understand the deformation mechanisms of human skull corresponding to the ICP change. The skull is a layered sphere constructed in a specially designed form with a Tabula externa, Tabula interna, and a porous Diploe sandwiched in between. Based on the established knowledge of cranial cavity importantly composed of skull and duramater, a thin-walled structure was simulated by the composite shell elements of the finite-element software [15].

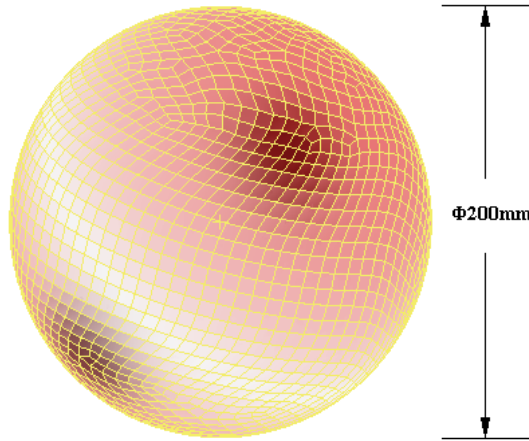


Fig. 6. The sketch of 3D cranial cavity and grid division

Of course, the structure, dimension and characteristic parameter of human skull must be given before the calculation. The thickness of calvaria [16] varies with the position, age, gender and individual, so does dura mater [17]. Tabula externa and interna are all compact bones and the thickness of Tabula externa is more than that of Tabula interna. Diploe is the cancellous bone between Tabula externa and Tabula interna [18]. The parietal bone is the transversely isotropic material, namely it has the mechanical property of rotational symmetry in the axially vertical planes of skull [19]. The important mechanical characteristic of cancellous bone is viscoelasticity, which is generally considered as the semi-closed honeycomb structure composed of bone trabecula reticulation. The main composition of cerebral dura mater, a thick and tough bilayer membrane, is the collagenous fiber which has the characteristic of linear viscoelasticity [20]. And the thickness of dura mater obviously varies with the changing ICP [21]. The mechanical performance of skull is isotropic along the tangential direction on the surface of skull bone [22], in which the performance of compact bone in the Tabula externa is basically the same as that in the Tabula interna [23]. Thus both cancellous bone and dura mater can be regarded as isotropic materials. And the elastic modulus of fresh dura mater varies with the delay time [24].

Next we need determine the fluctuant scope of human ICP. ICP is not a static state, but one that influenced by several factors. It can rise sharply with coughing and sneezing, up to 50 or 60mmHg to settle down to normal values in a short time. It also varies according to the activity the person is involved with. For these reasons single measurement of ICP is not a true representation. ICP needs to be measured over a period. Measured by means of a lumbar puncture, the normal ICP in adults is 8 mmHg to 18 mmHg. But so far there are almost no records of the actual human being's ICP in clinic. The geometry and structure of monkey's skull, mandible and cervical muscle are closer to those of human beings than other animals'. So the ICP of monkeys [25] can be taken as the reference to that of human beings'. The brain appears to be mild injury when ICP variation is about 2.5 kPa, moderate injury when ICP variation is about 3.5 kPa and severe injury when ICP variation is about or more than 5 kPa. Therefore, we carried out the following theoretical analysis with the ICP scope from 1.5 kPa to 5 kPa.

In this paper, the finite-element software MSC_PATRAN/NASTRAN and ANSYS are applied to theoretically analyze the deformation of human skull with the changing ICP. The external diameter of cranial cavity is about 200 mm. The thickness of shell is the mean thickness of calvarias. The average thickness of adult's calvaria is 6.0 mm, that of Tabula externa is 2.0 mm, diploe is 2.8 mm, Tabula interna is 1.2 mm and, dura mater in the parietal position is 0.4 mm.

Considering the characteristic of compact bone, cancellous bone and dura mater, we adopt their elastic modulus and Poisson ratios as 1.5×10^4 MPa, 4.5×10^3 MPa [26], 1.3×10^2 MPa [27] and 0.21, 0.01, 0.23 respectively.

After ignoring the viscoelasticity of human skull and dura mater, the strains of cranial cavity are shown in Table 1 with the finite-element software MSC_PATRAN/NASTRAN as ICP changing from 1.5 kPa to 5.0 kPa (Fig2.7). There is the measurable correspondence between skull strains and ICP variation. The strains of human skull can reflect the ICP change. When ICP variation is raised up to 2.5 kPa, the stress and strain graphs of skull bone are shown in Fig2.8~Fig2.13.

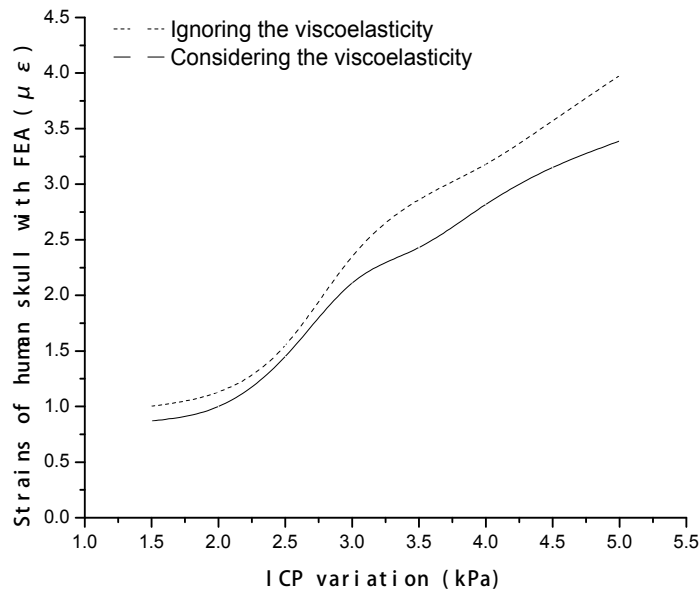


Fig. 2.7 The strain curves of finite-element simulation under the conditions of ignoring and considering the viscoelasticity of human skull and duramater with the changing ICP from 1.5kPa to 5kPa

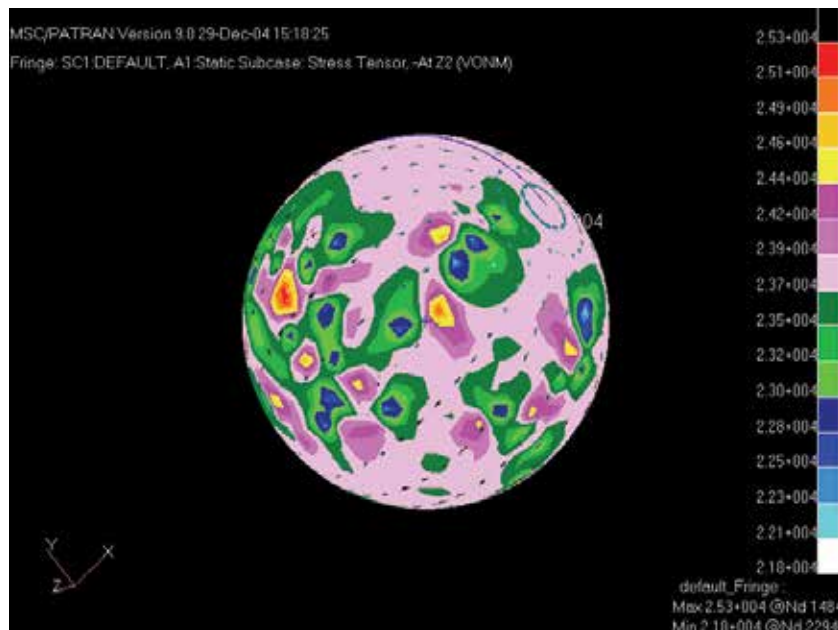


Fig. 2.8 Stress distribution

The scope of stress change on the outside surface is from 22.1 kPa to 25.3 kPa when ICP variation is raised up to 2.5 kPa by ignoring the viscoelasticity of human skull and duramater.

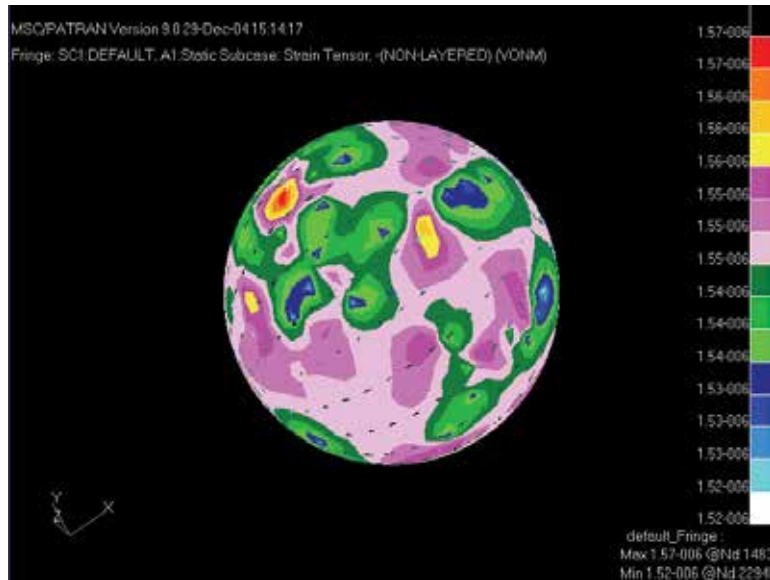


Fig. 2.9 Strain distribution

The scope of strain change on the outside surface is from $1.52\mu\epsilon$ to $1.57\mu\epsilon$ when ICP variation is raised up to 2.5 kPa by ignoring the viscoelasticity of human skull and dura mater.

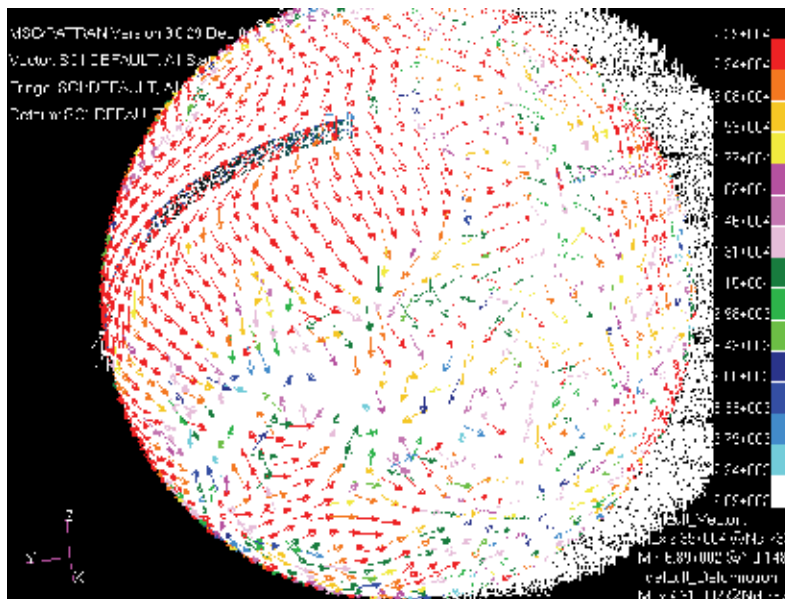


Fig. 2.10 The maximal stress vector distribution

The maximal main stress is about 22.4kPa when ICP variation is raised up to 2.5 kPa by ignoring the viscoelasticity of human skull and dura mater.

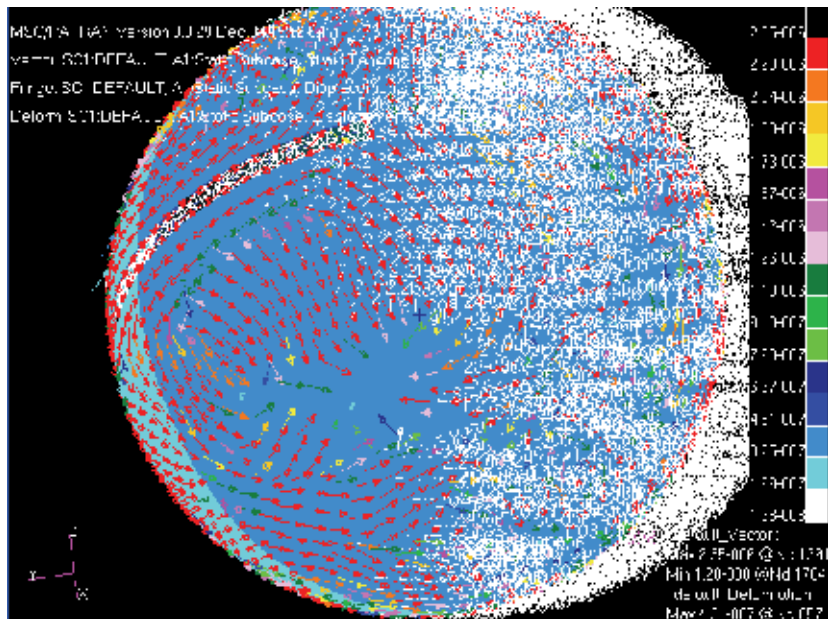


Fig. 2.11 The maximal strain vector distribution

The maximal main strain is about $2.2\mu\epsilon$ when ICP variation is raised up to 2.5 kPa by ignoring the viscoelasticity of human skull and dura mater.

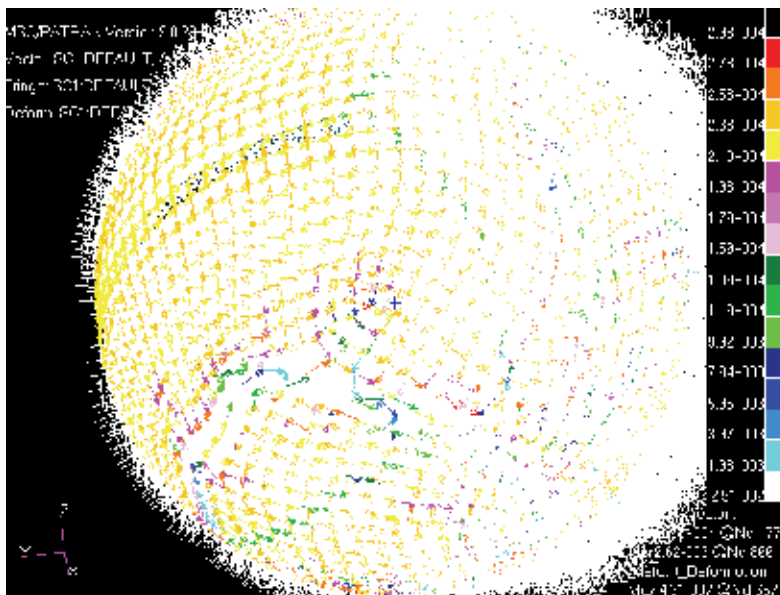


Fig. 2.12 Stress vector distribution

The main stress vector is about 21.8kPa when ICP variation is raised up to 2.5 kPa by ignoring the viscoelasticity of human skull and dura mater.

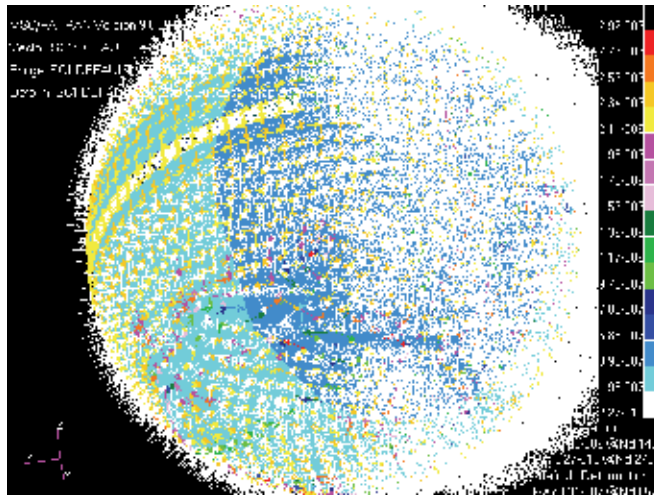


Fig. 2.13 Strain vector distribution

The main strain vector is about $2.14\mu\epsilon$ when ICP variation is raised up to 2.5 kPa by ignoring the viscoelasticity of human skull and dura mater.

2.2.4 The finite-element analysis of strains by considering the viscoelasticity of human skull and dura mater

Human skull has the viscoelastic material [28]. Considering the viscoelasticity of human skull and dura mater, we use the viscoelastic option of the ANSYS finite-element program to analysis the strains on the exterior surface of human skull as ICP changing. According to the symmetry of 3D model of human skull, the preprocessor of the ANSYS finite-element program is used to construct a 1/8 finite-element model of human skull and dura mater consisting of 25224 nodes and 24150 three-dimensional 8-node isoparametric solid elements, shown in Fig2.10.

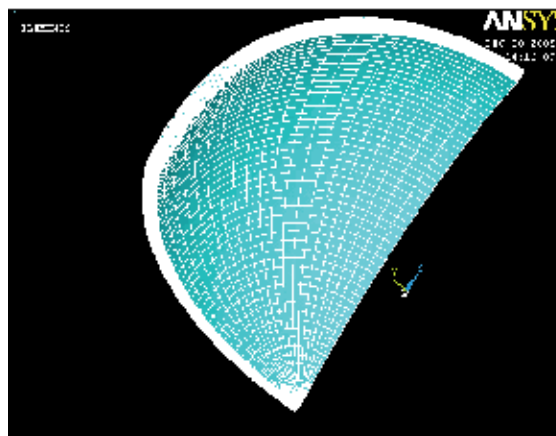


Fig. 2.14 Finite element model of 1/8 cranial cavity shell

The three-dimensional stress-strain relationships for a linear isotropic viscoelastic material are given by:

$$\sigma_{ij} = \int_0^t \left[2G(t-\tau) \frac{\partial e_{ij}(\tau)}{\partial \tau} + \delta_{ij} K(t-\tau) \frac{\partial \theta(\tau)}{\partial \tau} \right] d\tau; \quad (i, j = 1, 2, 3) \quad (2.33)$$

Here, σ_{ij} — the Cauchy stress tensor;

e_{ij} — the deviatoric strain tensor;

δ_{ij} — the Kronecker delta;

$G(t)$ — the shear relaxation function;

$K(t)$ — the bulk relaxation function;

$\theta(t)$ — the volumetric strain;

t — the present time;

τ — the past time.

Before the theoretical analysis of the minitraumatic strain-electrometric method, we need to set up the viscoelastic models to describe the relevant mechanical properties of human skull and dura mater.

(1) Viscoelastic model of human skull

Under the constant action of stress, the strain of ideal elastic solid is invariable and that of ideal viscous fluid keeps on growing at the equal ratio with time. However, the strain of actual material increases with time, namely so-called creep. Generally, Maxwell and Kelvin models are the basic models to describe the performance of viscoelastic materials. Maxwell model represents in essence the liquid. Despite the representative of solid, Kelvin model can't describe stress relaxation but only stress creep (Fig2.11). So the combined models made up of the primary elements are usually adopted to describe the viscoelastic performance of actual materials. The creep of linear viscoelastic solid can be simulated by the Kelvin model of three parameters or the generalized Kelvin model.

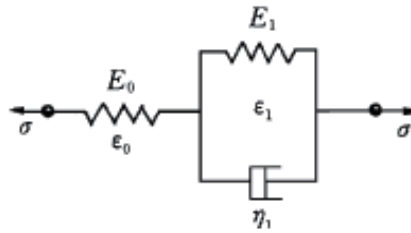


Fig. 2.15 Three parameters Kelvin model of human skull.

Kelvin model of three parameters is shown in Fig2.12 (a). Fig2.12 (b) is the relaxation curves of human skull and Kelvin model of three parameters in the compressive experiment. Fig2.10 (c) is the creep curves of human skull and Kelvin model of three parameters. It shows that the theoretical Kelvin model of three parameters can well simulate the mechanical properties of human skull in the tensile experiments. Thus the Kelvin model of three parameters is adopted to describe the viscoelasticity of human skull in this paper.

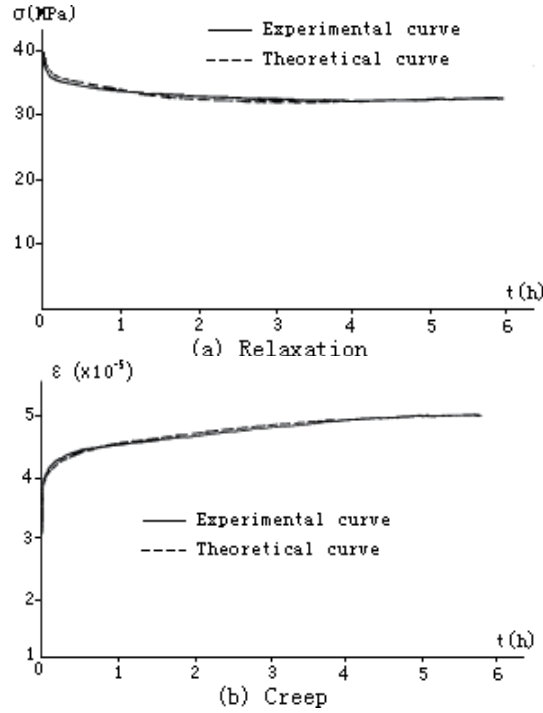


Fig. 2.16 The relaxation and creep train-time curves between experiment and three parameters Kelvin theoretical model of human skull.

For the Kelvin model of three parameters, the stress and strain of human skull are shown in equation (2.34),

$$\begin{cases} \varepsilon = \varepsilon_0 + \varepsilon_1 \\ \sigma = E_1 \varepsilon_1 + \eta \dot{\varepsilon}_1 \\ \sigma = E_0 \varepsilon_0 \end{cases} \quad (2.34)$$

After the calculation based on the equation (1), the elastic modulus of human skull is equation (2.35),

$$E = \left(\frac{E_0 E_1}{E_0 + E_1} \right) + \left(\frac{E_0^2}{E_0 + E_1} \right) e^{\frac{t}{P_1}} \quad (2.35)$$

Here, σ — Direct stress acted on elastic spring or impact stress acted on viscopot;

ε — Direct strain of elastic spring;

E — Elastic modulus of tensile compression;

η — Viscosity coefficient of viscopot;

$\dot{\varepsilon}$ — strain ratio;

$$P_1 = \frac{\eta}{E_0 + E_1}$$

(2) Viscoelastic model of human duramater

The generalized Kelvin model is shown in Fig2.13 (c). Fig2.13 (a) is the creep experimental curves of human duramater. Fig2.13 (b) is the curves of creep compliance for the generalized Kelvin model. It shows that the tendency of creep curve in the experiment is coincident with that of creep compliance for the generalized Kelvin model. Creep is the change law of material deformation with time under the invariable stress, so here σ is constant. For the generalized Kelvin model, the stress-strain relationship is $\varepsilon(t) = J(t)\sigma$. Thus the tendency of theoretical creep curve is totally the same as that of experimental one for human duramater. So in this paper, the generalized Kelvin model composed of three Kelvin-unit chains and a spring is adopted to simulate the viscoelasticity of human dura mater in this paper.

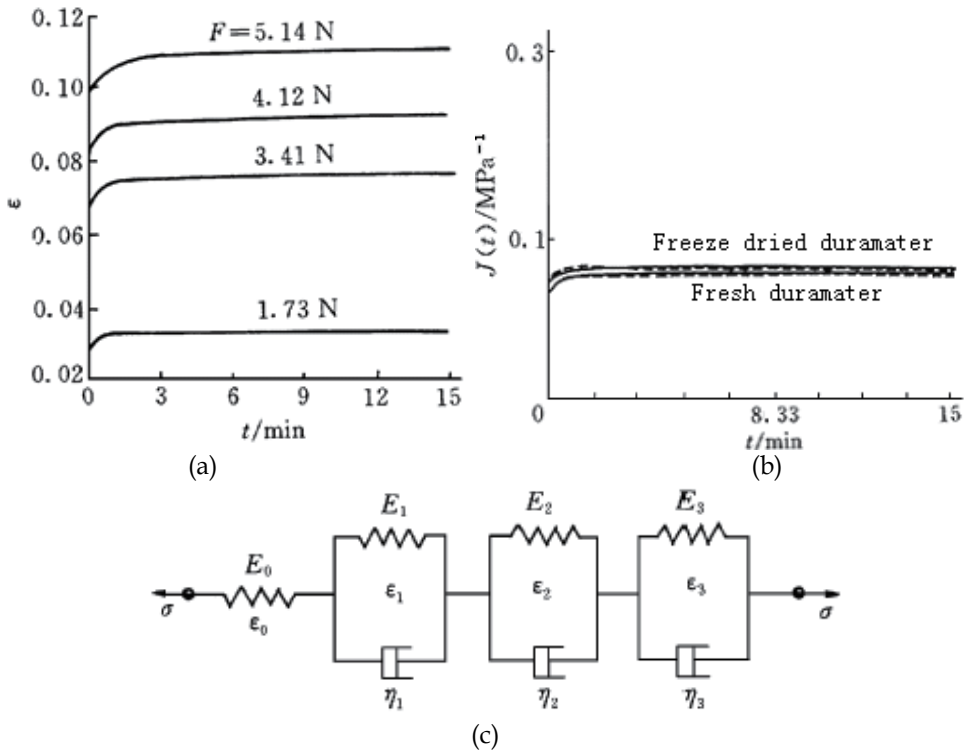


Fig. 2.17 Creep train-time curves under different loads for fresh human duramater ($L_0=23\text{mm}$, $\theta=37^\circ\text{C}$). Creep compliance curves of human duramater Kelvin model. And the Kelvin model of the duramater.

For the viscoelastic model of human dura mater composed of the three Kelvin-unit chains and a spring, the stress and strain of human dura mater are shown in equation (2.36),

$$\begin{cases} \varepsilon = \varepsilon_0 + \varepsilon_1 + \varepsilon_2 + \varepsilon_3 \\ \varepsilon_0 = \frac{\sigma}{E_0} \\ \sigma = E_1 \varepsilon_1 + \eta_1 \dot{\varepsilon}_1 = E_2 \varepsilon_2 + \eta_2 \dot{\varepsilon}_2 = E_3 \varepsilon_3 + \eta_3 \dot{\varepsilon}_3 \end{cases} \quad (2.36)$$

After the calculation based on the equation (2.36), the creep compliance of human dura mater is equation (2.37),

$$J(t) = E_0^{-1} + E_1(1 - e^{-t/\tau_1}) + E_2^{-1}(1 - e^{-t/\tau_2}) + E_3^{-1}(1 - e^{-t/\tau_3}) \quad (2.37)$$

Then the elastic modulus of human dura mater is equation (5),

$$E = [E_0^{-1} + E_1^{-1}(1 - e^{-t/\tau_1}) + E_2^{-1}(1 - e^{-t/\tau_2}) + E_3^{-1}(1 - e^{-t/\tau_3})]^{-1} \quad (2.38)$$

Here, σ , ε , E , η , $\dot{\varepsilon}$ — Ditto mark;

τ_1 , τ_2 , τ_3 — Lag time, that is $\tau_1 = \eta_1 / E_1$, $\tau_2 = \eta_2 / E_2$, $\tau_3 = \eta_3 / E_3$.

In the finite-element software ANSYS, there are three kinds of models to describe the viscoelasticity of actual materials, in which the Maxwell model is the general designation for the combined Kelvin and Maxwell models. Considering the mechanical properties of human skull and dura mater, we adopt the finite-element Maxwell model to simulate the viscoelasticity of human skull-dura mater system. The viscoelastic parameters of human skull and dura mater are respectively listed in Table 2.1 and Table 2.2.

	Elastic Modulus (GPa)		Viscosity (GPa/s)	Delay time τ (s)	
	E0	E1	η	τ_r^*	τ_d^*
Compression	5.69 ± 0.26	42.24 ± 2.09	26.9 ± 1.5	2022 ± 198	2292 ± 246
Tension	13.64 ± 0.59	51.45 ± 2.54	57.25 ± 4.27	3180 ± 300	4026 ± 372

$$^* \tau_r = \eta / (E_1 + E_2), \quad \tau_d = \eta / E_2$$

Table 1. Coefficients for the viscoelastic properties for human skull

	Elastic modulus (MPa)				Delay time τ (s)		
	E0	E1	E2	E3	τ_1	τ_2	τ_3
Duramater	16.67	125.0	150.0	93.75	40	10^4	10^6

Table 2. Creep coefficients for the viscoelastic properties for fresh human duramater

(3) The stress and strain distribution by the finite-element analysis

Fig2.14 (a) ~ (e) are the analytic graphs of stress and strain with finite-element software ANASYS when ICP variation is raised up to 2.5kPa. After considering the viscoelasticity of human skull and duramater, the stresses and strains of cranial cavity are shown in Fig2.15 as the ICP changing from 1.5kPa to 5kPa with the finite-element software ANSYS. It shows that the stress and strain distributions on the exterior surface of human skull are well-proportioned and that the stress and strain variation on the exterior surface of cranial cavity is relatively small corresponding to the ICP change. The strains of cranial cavity are

coincident with ICP variation. The deformation scope of human skull is theoretically from $0.9 \mu\epsilon$ to $3.4 \mu\epsilon$ as the ICP changing from 1.5kPa to 5.0kPa. Corresponding to ICP of 2.5kPa, 3.5kPa and 5.0kPa, the strain of skull deformation separately for mild, moderate and severe head injury is $1.5 \mu\epsilon$, $2.4 \mu\epsilon$, and $3.4 \mu\epsilon$ or so.

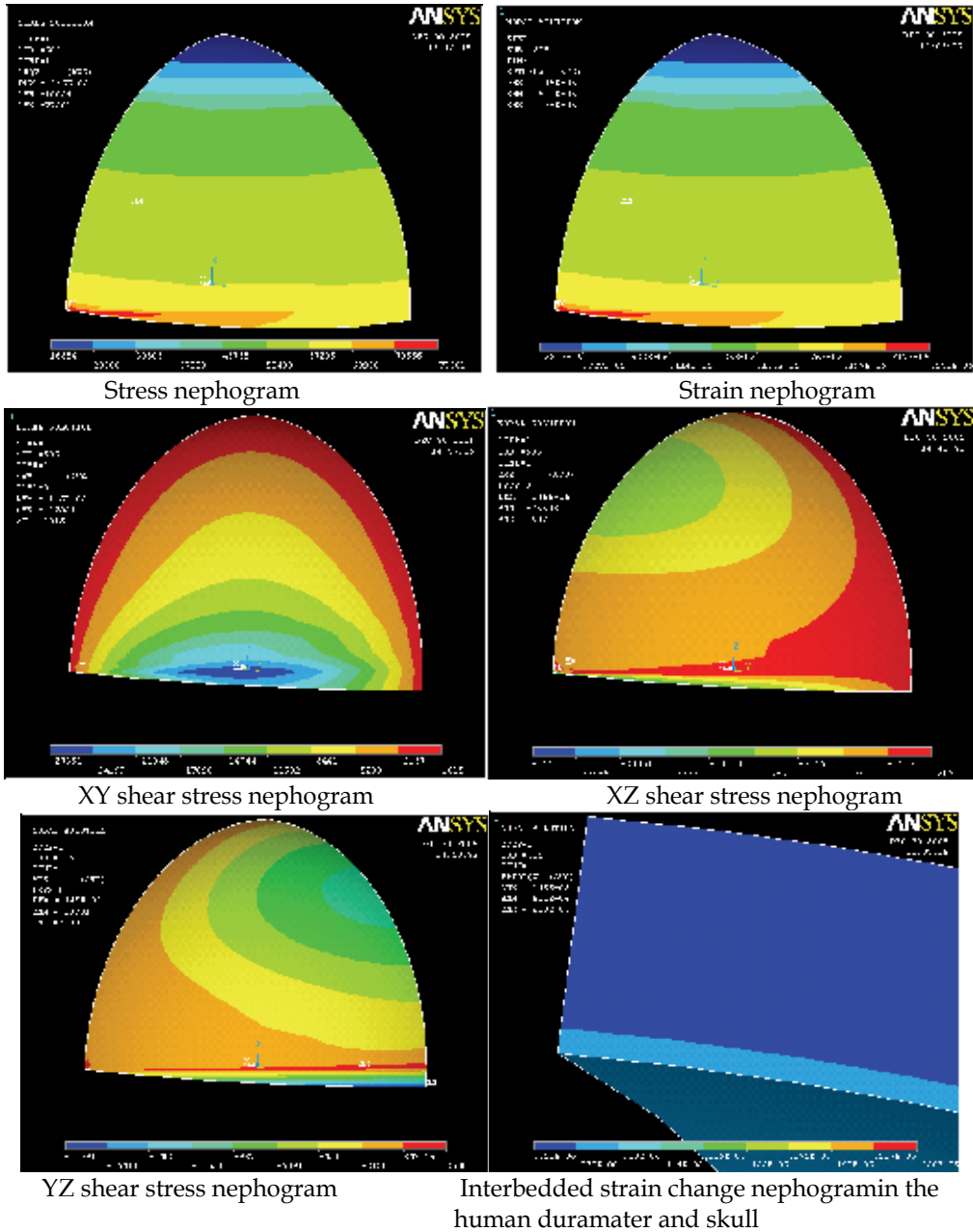


Fig. 2.18 The stress and strain distribution considering viscoelasticity of human skull and duramater

From the relationships about total, elastic and viscous strains of human skull and dura mater in Fig.4 (g), the viscous strains account for about 40% and the elastic strains are about 60% of total strains with the increasing ICP.

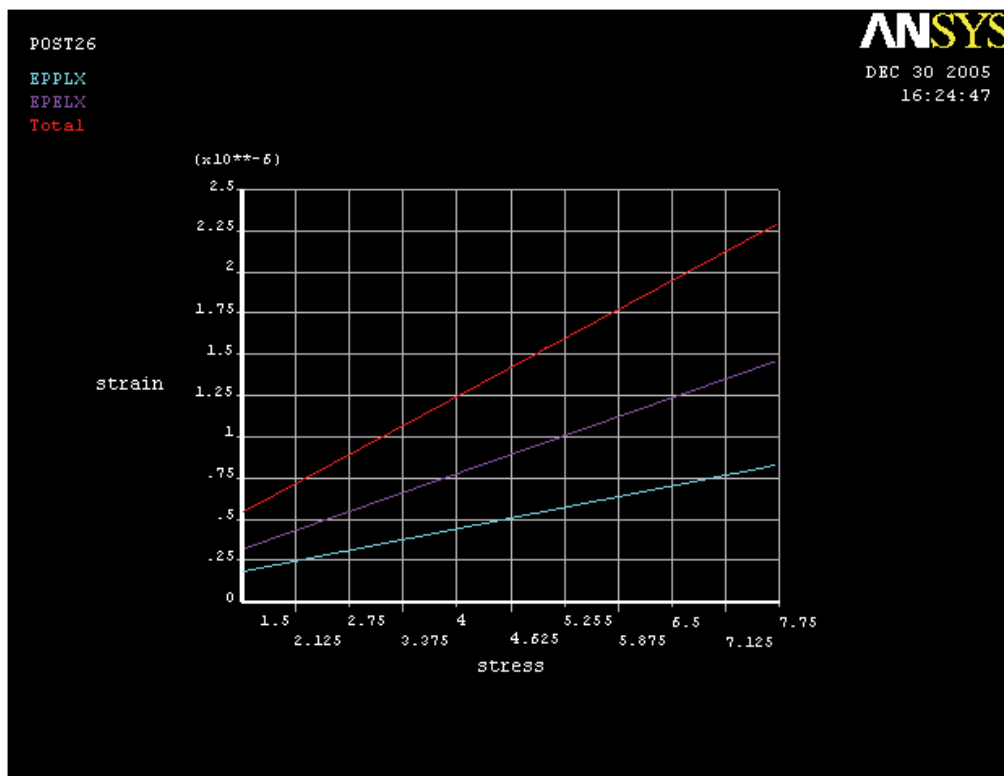


Fig. 2.19 Curves among total, elastic and viscous strain when the ICP increment is 2.5KPa. Here EPELX is elastic strain curve, EPPLX is viscous strain curve. The viscous strain is about 40% of total strain.

3. Potential Therapeutic Actions of Hypothermia

Intracranial pressure (ICP) is a main index and extremely important to the diagnoses and treatment of many diseases in neurosurgery. Acute cerebral diseases frequently lead to elevated ICP that is the early signal of illness complication in the skull and the common reason of death in the advanced stage, and intracranial pressure has been higher than 2.0kPa (15mmHg) [29]. An increase in ICP is a severe medical problem. The diseases of central nervous system, such as severe head injury, cerebrovascular accident, brain tumor, etc, can cause various central high fevers, often up to 39°C or above [30], which have serious adverse effects on the disease prognosis. Because the increased ICP affects the thermoregulation center of hypothalamus, a sustained high fever will be appeared in clinic, often up to 39°C or above [31]. Mild hypothermia therapy can significantly improve the recovery of the central nervous system, has a positive effect on the treatment of severe head

injury, can reduce intracranial hypertension and mortality, and improve the prognosis [32]. Thus, the mild hypothermia is the main temperature environment for the treatment of brain injury, intracranial hypertension, and so on. Mild hypothermia treatment of severe traumatic brain injury in recent years is another important means [33].

Yue et al [34,35] proved the human skull and duramater could be deformed with the changing ICP, and the deformation scope of human skull is from $1.30\mu\epsilon$ to $4.80\mu\epsilon$ as the ICP changing from 1.5 kPa to 5.0 kPa. The deformation tendency in this paper needs to be studied with the increasing ICP under the mild hypothermia environment.

The finite-element method extensively solves the biomechanical problem in the medical fields. Compared to other bio-mechanical modelling, the finite-element method can more accurately express the human body geometry and architecture. Therefore, the ANSYS finite-element software is in this paper used to reconstruct the three-dimensional cranial cavity of human being with the mild hypothermia treatment.

3.1 Methods

3.1.1 The geometric model

Reconstruction of the finite-element model of cranial cavity are mostly through the multi-CT scanning technology at home and abroad [36, 37], which is simple and high precision. However, the object being scanned is only a single individual. It is difficult to scan multiple images of individual unity for the universal data. Based on the average measured data of human skull of 104 Chinese people from 18 to 76 years of age, including 67 male and 37 female [38], the three-dimensional model of cranial cavity are directly drawn in the ANSYS program.

This paper studied importantly the deformation of cranial cavity, including brain tissue, cerebrospinal fluid and brain blood flow with the ICP changes. So the model of cranial cavity was properly simplified: only the cavity in which the brain lies, that is, a closed cavity is made up of the parietal bone, occipital, frontal and temporal bone, and a layer of duramater. For the approximate symmetry of person's head, the 1/2 cranial cavity model is built in this paper (Fig3.1).

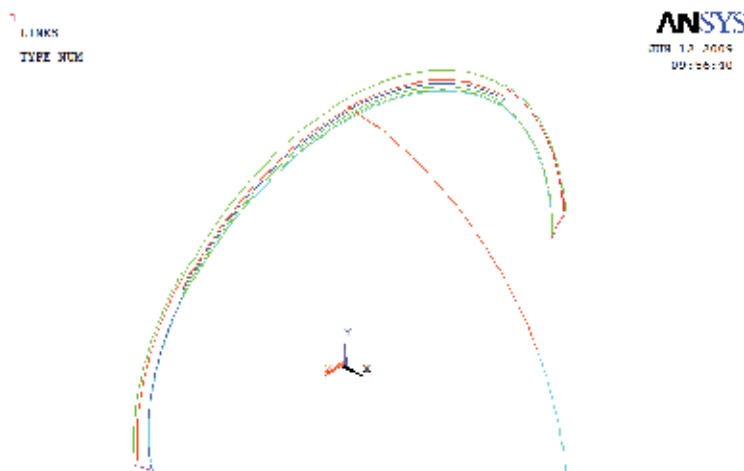


Fig. 3.1 Skeleton of the cranial cavity trendline

In Fig3.1, the thickness of parietal, frontal, occipital, temporal bone, as well as duramater are respectively 5.3560mm, 6.5558mm, 7.5286mm, 2mm and 0.4mm [39, 40]. From outside to inside each layer in turn, the thickness of external compact bone, diploe and internal compact bone are about 3mm, 1.8mm, 1mm. Due to the smoothness, the parietal bone is regarded as the main measured position. By the Extrude command, the Fig.1 can be formed the volume. After bonding the border among the frontal, parietal, occipital and temporal bone and the duramater with the glue command, the model of cranial cavity is shown in Fig3.2.

3.1.2 The meshing and load

Although the thickness is only 0.4mm [41], the viscoelasticity of duramater is strong [42], so the effect is great on the cranial deformation. The three-dimensional finite element model of cranial cavity is dealt with a composite structure made up of the skull and duramater. The elastic modulus and Poisson's ratio of parietal bone, frontal bone and other parts is shown in Table 3.1. The hexahedral grid was adopted to mesh the entire cranial cavity. The adjacent parts were dealt with the Glue command, and the grid refinement with the Meshing-Modify Mesh command were used as the irregular mesh to the edge, sharp or irregular position. Thus the 1/2 finite-element model of the cranial cavity, including the parietal, pre-frontal, occipital, temporal bone and duramater, and the cell type is block unit. The Fig3.3 is the meshing diagram of 1/2 cranial cavity, in which there are 9,700 hexahedral element and 27,256 nodes.

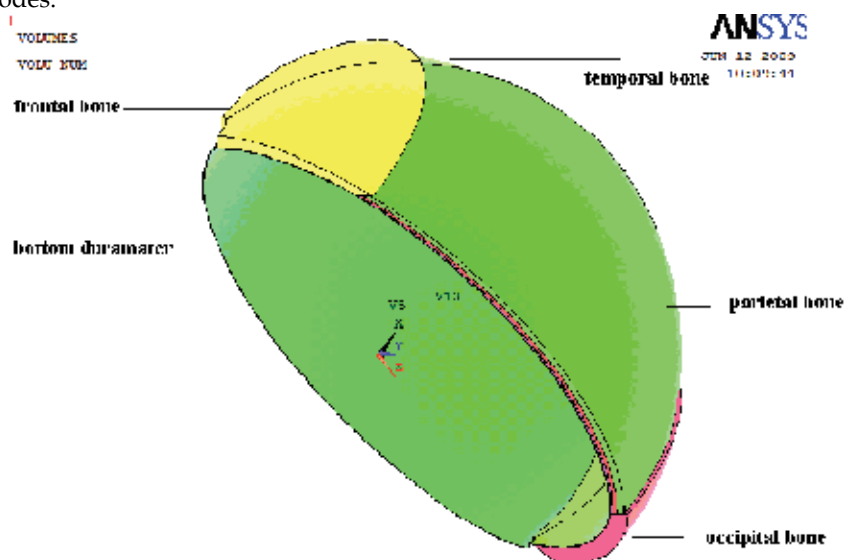


Fig. 3.2 3D model of 1/2 cranial cavity

Parietal bone	Elastic modulus	Poisson's ratio
Compact bone	$1.5 \times 10^{10} \text{Pa}$	0.21
Cancellous bone	$4.5 \times 10^9 \text{Pa}$	0.01
Duramater	$1.3 \times 10^8 \text{Pa}$	0.23

Table 3. [43, 44] The material characteristics of human skull and duramater

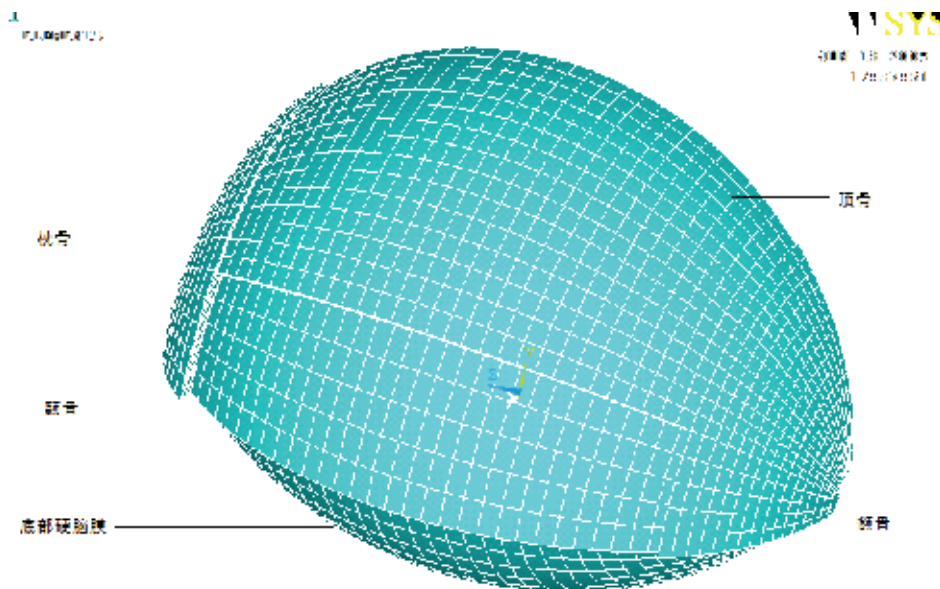


Fig. 3.3 Finite-element meshes of 1/2 cranial cavity

According to the literature [45], the scope of increased intracranial pressure is refined from 2.0kPa to 6.0kPa. Namely, the outward loads among 2.0kPa ~ 6.0kPa are imposed on the inside surfaces of the cranial cavity model to simulate the ICP changes. The temperature of high fever caused by the Intracranial hypertension is 39.5°C [30], which is the initial temperature inside the cranial cavity. The medical definition range of mild hypothermia is from 28°C to 35°C [46]. The over-temperature protection will make the protective effect isn't obvious. If the treatment temperature is too low, the serious complications will come into being. So the optimal temperature of treatment are consistent with the scope among 32°C ~ 35°C at home and abroad [47]. The average 33.5°C is the set point in this paper. In Clinic, the ice bag or blanket wrapped around the patients' body is used to cool the temperature during the mild hypothermia therapy [49]. Thus the temperature load of 33.5°C is exerted to the outer surface of human skull to simulate the mild hypothermia therapy.

3.2 Results

While patients have high fever or are in the mild hypothermia treatment, Fig3.4 – Fig3.7 are separately the strain graphs corresponding to ICP of 2.5kPa and 5.0kPa. And the Table 3.2 is the relevant data of finite-element simulation. Based on the above graphs and charts, it shows that the strains of human skull are in agreement with ICP variation whether under the high fever or mild hypothermia environments. But the former is slightly lower than the latter. At the same time, the strains of junction between the occipital and frontal parietal bone is the maximum, the connecting with the temporal bone is smaller, and the strains lie in the middle level at the central part of parietal bone.

Fig3.4 and Fig3.5 are the respectively strain diagrams of cranial cavity under the mild hypothermia environment and normal temperature conditions while the ICP is 3.0kPa. It

shows that the strains decreased about $0.001\mu\epsilon$ under the mild hypothermia environment than those under the normal temperature conditions during the same circumstance of ICP changes.

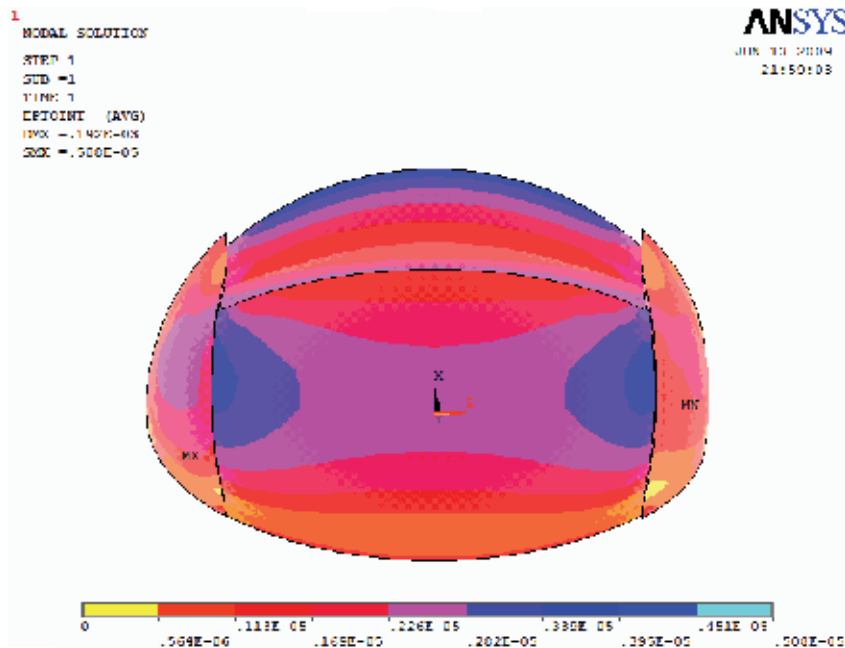


Fig. 3.4 Strain graph without hypothermia treatment when the ICP is 3.0kPa

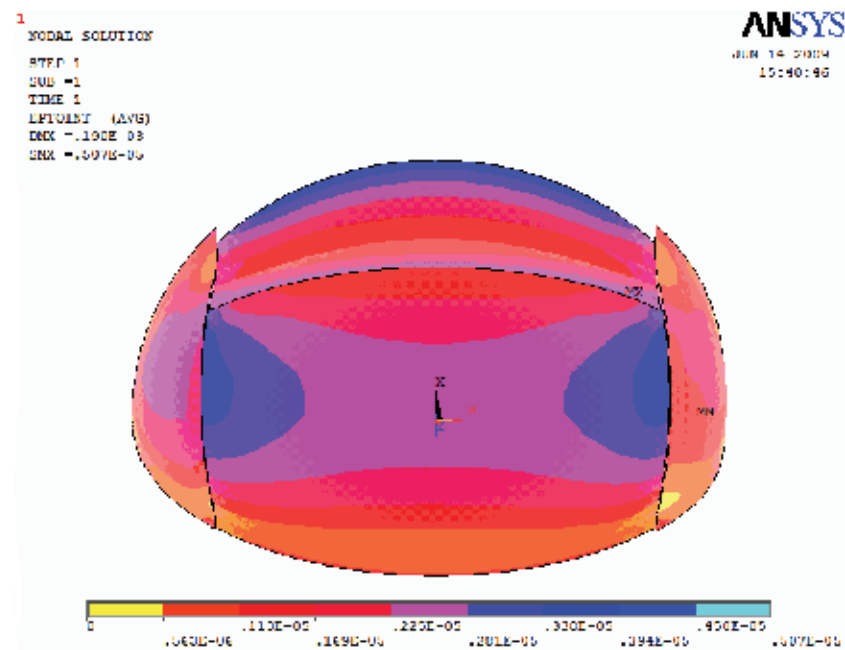


Fig. 3.5 Strain graph with hypothermia treatment when the ICP is 3.0kPa

Fig3.6 and Fig3.7 are the respectively strain diagrams of cranial cavity under the mild hypothermia environment and normal temperature conditions while the ICP is 5.0kPa. It shows that the deformation strains of cranial cavity were significantly increased with the rising ICP. The strains increased while ICP is 3.0kPa about 39.9% than those while ICP is 5.0kPa under the normal temperature conditions. Moreover, the strains decrease more obviously under the mild hypothermia environment, and the maximum is $0.03\mu\epsilon$ appeared in the junction among the parietal, frontal and occipital bone.

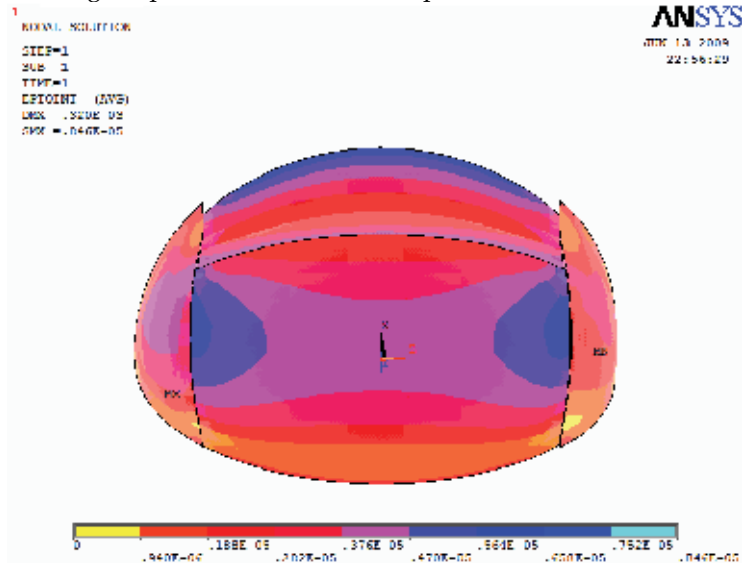


Fig. 3.6 Strain graph without hypothermia treatment when the ICP is 5.0kPa

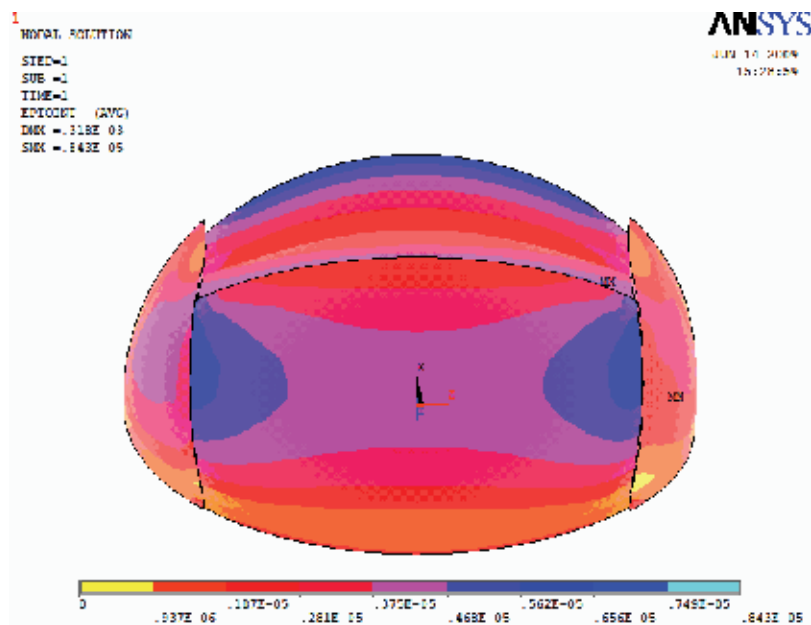


Fig. 3.7 Strain graph with hypothermia treatment when the ICP is 5.0kPa

ICP variation	Strains of 39.5°C	Strains of 33.5°C
kPa	$\mu\epsilon$	$\mu\epsilon$
2.0	1.510	1.500
2.5	1.880	1.870
3.0	2.260	2.250
3.5	2.630	2.620
4.0	3.010	2.990
4.5	3.390	3.370
5.0	3.760	3.740
6.0	4.520	4.490

Table 4. Strain data of cranial cavity simulated by ANSYS with the ICP variation

4. Results and Discussion

4.1 Results

This paper carries respectively on the stress and strain analysis on both conditions of ignoring and considering the viscoelasticity of human skull and duramater by finite-element software MSC_PPATRAN/NASTRAN and ANSYS as ICP changing from 1.5kPa to 5kPa. The three-dimensional finite element model of cranial cavity and the viscoelastic models of human skull and duramater are constructed in this paper. At the same time, the ANSYS finite-element software is in this paper used to reconstruct the three-dimensional cranial cavity of human being with the mild hypothermia treatment. The conclusion is as follows:

(1) The human skull and duramater are deformed as ICP changing, which is corresponding with mechanical deformation mechanism.

(2) When analyzing the strain of human skull and duramater as ICP changing by the finite-element software ANSYS, the strain of considering the viscoelasticity is about 14% less than that of ignoring the viscoelasticity of human skull and duramater. Because the viscoelasticity analysis by finite-element software ANSYS is relatively complex and the operation needs the huge memory and floppy disk space of computer, it is totally feasible to ignore the viscoelasticity while calculating the FEA strain of human skull and duramater as ICP changing.

(3) The viscosity plays an important role in the total deformation strain of human skull and duramater as ICP changing. In the strains analysis of human skull and duramater with the changing ICP by the finite-element software ANSYS, the viscous strain accounts for about 40% of total strain, and the elastic strain is about 60% of total strain.

(4) Because the strains of human skull are proportional to ICP variation and the caniocerebra characteristic symptoms completely correspond to different deformation strains of human skull, ICP can be completely obtained by measuring the deformation strains of human skull.

That is to say, the minitraumatic method of ICP by strain electrometric technique is feasible. Furthermore, ICP variation is respectively about 2.5kPa when the strain value of human skull is about $1.4\mu\epsilon$, about 3.5kPa when the strain value of human skull is about $2.1\mu\epsilon$, and about 5kPa when the strain value of human skull is about $3.9\mu\epsilon$.

(5) The strains decreased under the mild hypothermia environment about 0.56% than those under the normal temperature conditions during the same circumstance of ICP changes.

(6) The deformation scope of human skull is theoretically from $1.50\mu\epsilon$ to $4.52\mu\epsilon$ as the ICP changing from 2.0kPa to 6.0kPa under the normal situation, and from $1.50\mu\epsilon$ to $4.49\mu\epsilon$ under the mild hypothermia environment. Accordingly, the strains of skull deformation for mild, moderate and severe head injuries are separately $1.87\mu\epsilon$, $2.62\mu\epsilon$ and $3.74\mu\epsilon$ or so corresponding to ICP of 2.5kPa, 3.5kPa and 5.0kPa.

4.2 Discussion

From the eighties, the scholars abroad paid more attention to the effect on the mild and moderate hypothermia for the brain protection after the ischemic and traumatic brain injury, temperature from 29°C to 36°C [32]. And now the mild hypothermia has gradually been applied to the treatments of cerebral injury and ischemia. After the introduction of mild hypothermia in clinic, the impact must be considered on the protective measures to the other protection or measurement mechanisms. The deformation of cranial cavity had been calculated with the changing ICP by the finite element method [48]. Considering the increased ICP can cause central heat, taking into account, thereby requiring cooling measures, the hypothermia treatment should be carried out, which has been another important means of treating the severe craniocerebral trauma in recent years. Since the strains decreased under the mild hypothermia environment about 0.56% than those under the normal temperature conditions during the same circumstance of ICP changes, the mild hypothermia has obvious impact on the deformation of cranial cavity. That is to say, the effect must be considered on the mild hypothermia to measurement data in clinic or medical experiments while measuring the deformation of cranial cavity under the mild hypothermia environment simultaneously.

In this paper, the finite-element simulation was carried out to analyze the deformation of cranial cavity. Many complex relationships and influencing factors lie in the actual deformation of cranial cavity with the changing ICP. Therefore, in order to obtain the accurate deformation tendency of cranial cavity, the precise simulation to the finite-element model and further experimental studies in vivo and clinic need to be carried on.

5. References

- [1] Kaitai Li, Aixiang Huang, Qinghuai Huang. The finite element method and its application. 1st Eds, Xian'an: Xian'an Jiaotong University Publisher, 1992: 1-3
- [2] Guoqing Liu, Qingdong Yang. The applicable course of ANSYS in the engineering. China Railway Publisher, 2003
- [3] Ross, C.T.F. The finite element method of structural mechanics. 1st Eds. People Jiaotong Publisher, 1991. 1: 23-50

- [4] Siyuan Cheng. The methodology of finite element method. Chongqing University Journal (Social science), 2001, 7(4): 61-63
- [5] Yijin Wang, Yongfeng Li, Kerong Dai, etc. Stress analysis on the femoral stem in the replacement of artificial joint. Chinese Journal of Bone and Joint Injury 1993; 8(1): 40
- [6] Meichao Zhang, Weidong Zhao, Lin Yua, Jiantie Li, Lei Tang, Shizhen Zhong. Three-dimensional reconstruction of the knee joint of digitized virtual Chinese male No.1 by finite element simulation. Journal of First Military Medical University, 2003, 23(6): 527-529
- [7] Chunbao Zhang, Xuanxiang Ma, Shaofeng Zhang, etc. Finite element analysis of the stress induced by an endodontic endosseous implant placed through a central incisor with periodontal compromises. Journal of practical stomatology, 2003, 19: 103
- [8] Jianchao Gui, Qiang Zhou, Xiangjie Gu, Haiqi Shen, Xin Ma, Jinsong Chen, Yijin Wang. Three dimensional finite element modeling study of the effect of femoral quality on hip arthroplasty. The journal of bone joint injury, 2000, 15(3): 212-240
- [9] Youxin Shu, Shuliang Xu, Riqi Chen, Xianxiang Liu. The study of the relation between the wrong using of pillow and the cause of cervical spondylosis with the spacial analytical method of finite element in dynamic models. Journal of Traditional Medical Traumatology & Orthopedics, 1999, 7(6): 9-12
- [10] Liyang Dai, Yinkan Xu, Wenming Zhang, Kaiyuan Tu, Peilai Cheng. A three-dimensional finite analysis of lumbar intervertebral disc. Journal of Biomedical Engineering, 1991, 8(3): 237-441
- [11] Meichao Zhang, Shizhen Zhong. The effects of transcription factor decoy on gene therapy. Progress of anatomical sciences, 2003, 9(1): 53-56
- [12] R. M. Jones. Composite Material. 1st Eds. Shanghai Science and Technology Publisher, 1981. 6: 41-73
- [13] Zhimin Zhang. Structural Mechanics of Composite Material. 1st Eds. Beijing: Publish of BUAA, 1993. 9: 85-86
- [14] Willinger, R., Kang, H.S., Diaw, B.M. Development and validation of a human head mechanical model. Comptes Rendus de l'Academie des Sciences Series IIB Mechanics, 1999, 327: 125-131.
- [15] Pithioux M, Lasaygues P, Chabrand P. An alternative ultrasonic method for measuring the elastic properties of cortical bone. Journal of Biomechanics 2002; 35: 961-968.
- [16] Hakim S, Watkin KL, Elahi MM, Lessard L. A new predictive ultrasound modality of cranial bone thickness. IEEE Ultrason Sympos 1997; 2: 1153-1156.
- [17] Hatanaka M. Epidural electrical stimulation of the motor cortex in patients with facial neuralgia. Clinical Neurology and Neurosurgery 1997; 99: 155.
- [18] Kabel J., Rietbergen van B., Dalstra M., Odgaard A., Huiskes R. The role of an elective isotropic tissue modulus in the elastic properties of cancellous bone. Journal of Biomechanics 1999; 32: 673-680.
- [19] J. Kabel, B. van Rietbergen, M. Dalstra, A. Odgaard, R. Huiskes, Journal of Biomechanics, 32 (1999) 673-680.
- [20] Noort van R., Black M.M., Martin T.R. A study of the uniaxial mechanical properties of human dura mater preserved in glycerol. Biomaterials 1981; 2: 41-45.
- [21] Kuchiwaki H., Inao S., Ishii N., Ogura Y., Sakuma N. Changes in dural thickness reflect changes in intracranial pressure in dogs. Neuroscience Letters 1995; 198: 68-70.

- [22] Pithioux M., Lasaygues P., Chabrand P. An alternative ultrasonic method for measuring the elastic properties of cortical bone. *Journal of Biomechanics* 2002; 35: 961-968.
- [23] Wood J.L. Dynamic response of human cranial bone. *Journal of Biomechanics* 1971; 4: 1-12.
- [24] Hilton G. Cerebral oxygenation in the traumatically brain-injured patient: are ICP and CPP enough? *The Journal of Neuroscience Nursing*, 2000; 32: 278-282.
- [25] Wu G.R., Zhang, Y.R., Wang, Y.Q., You, G.X. Changes of Intracranial Pressure during Head Impact in Monkeys and Protection of Head Impact Injuries (Chinese). *Acta Aeronautica et Astronautica Sinica (Sup)* 1999; 20: 54-56.
- [26] Willinger R., Kang H.S., Diaw B.M. Development and validation of a human head mechanical model. *Comptes Rendus de l'Academie des Sciences Series IIB Mechanics* 1999; 327: 125-131.
- [27] Cattaneo P.M., Kofod T., Dalstra M., Melsen B. Using the finite element method to model the biomechanics of the asymmetric mandible before, during and after skeletal correction by distraction osteogenesis. *Computer Methods in Biomechanics & Biomedical Engineering* 2005; 8(3): 157-165.
- [28] Amit G., Nurit G., Qiliang Z., Ramesh R., Margulies S.S. Age-dependent changes in material properties of the brain and braincase of the rat. *Journal of Neurotrauma* 2003; 20(11): 1163-1177.
- [29] Andrus C. Dynamic observation and nursing of ICP. *Foreign Medical Sciences (Nursing Foreign Medical Science)* 1992; 11(6): 247-249.
- [30] Min W. Observation and nursing for fever caused by Acute Cerebrovascular Disease. *Contemporary Medicine* 2008; 143: 100-101.
- [31] Shiozaki T., Sugimoto H., Taneda M., et al. Selection of Severely Head Injured Patients for Mild Hypothermia Therapy. *Journal of Neurosurgery* 1998; 89: 206–211.
- [32] Lingjuan C. Current situation and development trend at home and abroad of Hypothermia Therapy Nursing. *Chinese Medicine of Factory and Mine* 2005; 18(3): 268-269.
- [33] Yue X.F., Wang L., Zhou F. Experimental Study on the strains of skull in rats with the changing Intracranial Pressure. *Tianjin Medicine Journal*, 2007, 35(2): 140-141.
- [34] Yue X.F., Wang L., Zhou F. Strain Analysis on the Superficial Surface of Skull as Intracranial Pressure Changing. *Journal of University of Science and Technology Beijing*, 2006, 28(12) : 1143-1151.
- [35] Qiang X., Jianuo Z. Impact biomechanics researches and finite element simulation for human head and neck. *Journal of Clinical Rehabilitative Tissue Engineering Research*, 2008, 12(48) : 9557-9560.
- [36] Zong Z., Lee H., Lu C. A three-dimensional human head finite element model and power in human head subject to impact. *Journal of Biomechanics*, 2006, 39(2): 284-292.
- [37] Ren L., Hao L., Shuyuan L., et al. A study of the volume of cranial cavity calculating from the dimension of cranial outer surface in X-ray films—its stepwise regressive equation and evaluation. *Acta Anthropologica Sinica*, 1999, 18(1): 17-21.
- [38] Yunhong L. Yanbo G. Yunjian W., et al. Experiment study on shock resistance of skull bone. *Medical Journal of the Chinese People's Armed Police Forces*, 1998, 9(7): 408-409.

- [39] Haiyan L., Shijie R., Xiang P., et al. The thickness measurement of alive human skull based on CT image. *Journal of Biomedical Engineering*, 2007, 24(5): 964-968.
- [40] Zhou L.F., Song D.L., Ding Z.R. Biomechanical study of human dura and substitutes. *Chinese Medical Journal*, 2002, 115(11): 1657-1659.
- [41] Willinger R., Kang H.S., Diaw B.M. Development and validation of a human head mechanical model. *Comptes Rendus de l'Academie des Sciences Series IIB Mechanics*, 1999, 327:125-131.
- [42] Odgaard A., Linde F. The Underestimation of young's modulus in compressive testing of cancellous bone specimens. *Journal of Biomechanics*, 1991, 24, 691-698.
- [43] Wenjun S., Jialin H. *Handbook of critical illness care*. People's Military Medical Press, 1994:75-77.
- [44] Dashi Z. The development and present status of mild hypothermia in cerebral protection[J]. *Moden Journal Neurosurg*, 2002, 2 (3): 133-135.
- [45] Bin W. Hong L. Investigation on the hypothermia treatment time course to the severe traumatic brain injury. *Chinese Journal of Clinical Neurosurgery*, 2008, 13(10): 628-629.
- [46] Joji Inamasu, Kiyoshi Ichikizaki. Mild hypothermia in neurologic emergency: an update. *Hypothermia in a Neurologic Emergency*, 2002, 40(2): 220-230.
- [47] Cuihua Z. Hui R. Study status of cooling methods to induce mild hypothermia. *Chinese Journal of Nursing*, 2006, 41(2): 170-172.
- [48] Yue X.F., Wang L., Zhou F. Finite element analysis on strains of viscoelastic human skull and duramater. *Journal of Basic Science and Engineering*, 2008, 16 (5): 686-694.

Application of finite element analysis in dentistry

Ming-Lun Hsu and Chih-Ling Chang
*Department of Dentistry, National Yang-Ming University
Taipei, Taiwan*

1. Introduction

Since Brånemark introduced the concept of osseointegration and the possibility of anchoring dental prostheses by intraosseous implantation in 1969, the clinical use of implants for oral and maxillofacial rehabilitation has rapidly expanded over the past 20 years. Biomechanical factors play a substantial role in implant success or failure. The application of occlusal forces induces stresses and strains within the implant-prosthesis complex and affects the bone remodeling process around implants. To achieve optimized biomechanical conditions for implant-supported prostheses, conscientious consideration of the biomechanical factors that influence prosthesis success is essential.

Many different methods have been used to study the stress/strains in bone and dental implants. Photoelasticity provides good qualitative information pertaining to the overall location of stresses but only limited quantitative information. Strain-gauge measurements provide accurate data regarding strains only at the specific location of the gauge. Finite element analysis (FEA) is capable of providing detailed quantitative data at any location within mathematical model. Thus FEA has become a valuable analytical tool in the assessment of implant systems in dentistry.

2. Assumptions in the use of FEA in the implant-bone biomechanical system

The power of the FEA resides principally in its versatility and can be applied to various physical problems. The structure analyzed can have arbitrary shape, loads, and supporting conditions, furthermore, the mesh can mix elements of different types, shapes, and physical properties. This great versatility is contained within a single computer program and the selection of program type, geometry, boundary conditions, element selection are controlled by user-prepared input data. The principal difficulty in simulating the mechanical behavior of dental implants lies in the modeling of human maxilla and mandible and its response to applied load. Certain assumptions are needed to make the modeling and solving process possible and these involve many factors which will potentially influence the accuracy of the FEA results: (1) detailed geometry of the implant and surrounding bone to be modeled, (2) boundary conditions, (3) material properties, (4) loading conditions, (5) interface between bone and implant, (6) convergence test, (7) validation.

3. Geometry

The attractive feature of finite element is the close physical resemblance between the actual structure and its finite element model. Excessive simplifications in geometry will inevitably result in considerable inaccuracy. The model is not simply an abstraction; therefore, experience and good engineering judgment are needed to define a good model. Whether to perform a two-dimensional (2-D) or three-dimensional (3-D) finite element model for the study is a significant query in FEA. It is usually suggested that, when comparing the qualitative results of one case with respect to another, a 2-D model is efficient and just as accurate as a 3-D model; although the time needed to create finite element models is decreasing with advanced computer technology, there is still a justified time and cost savings when using a 2-D model over 3-D, when appropriated. However, 2-D models cannot simulate the 3-D complexity within structures, and as a result are of little clinical values. The group of 3-D regional FE models is by far the largest category of mandible related researches. This is because modeling only the selected segment of mandible is much easier than modeling the complete mandible. In many of these regional models, reproduced boundary conditions are often oversimplified, and yield too much significance to their predictive, quantitative outcome.

When a model is supposed to be 2-D, the z axis (third dimension) must be specified to have either a plane-strain or a plane-stress condition. Plane strain assumes the model to be infinitely thick, so no strain occurs but some stress will progress in the z direction. Plane stress supposes the model to be thin enough, so no stress occurs but it has some strain in the z direction. In 3-D analysis, the stress and strain condition can be evaluated in all three axes (x, y, and z). The first step in FEA modeling is to represent the geometry of interest in the computer. In some 2-D FEA studies, the bone was modeled as a simplified rectangular configuration with the implant (Fig.1). The mandible was treated as an arch with rectangular section or a simplified segment as cancellous core surrounded by a 1.3-mm cortical layer with the overall dimensions of this block were 23.4 mm in height, 25.6 mm in mesiodistal length, and 9.0 mm in buccolingual width in 3-D FEA models (Fig.2). A dried specimen was scanned and imported into image analysis software (Image Tool 1.21; UTHSC, San Antonio, Tex, U.S.A.) to create the digital image of a sagittal cut of the palatine process of the 2-D maxilla. The outline of the image was manually plotted and each point converted into x and y coordinates. The coordinates were finally imported into the ANSYS software as keypoints of the definitive image. The same procedure was used to create the implant image (Fig.3). Computerized tomographic images of a human edentulous maxillary first molar area exhibiting buccal bone irregularities were acquired. The maxilla was approximately 11 mm in width bucco-lingually and 13 mm in height infero-superiorly. The cross-sectional image was then extruded to create a three-dimensional section of maxilla 6.5 mm in length in the mesio-distal direction. Due to symmetry with respect to the bucco-lingual plane in the geometry and loading, only half of the FE model needed to be considered (Fig.4). With the development of digital imaging techniques recently, more efficient methods are available included the application of specialized software for the direct transformation of 2- or 3-D information in image data from computed tomography (CT) or magnetic resonance imaging (MRI) into FEA meshes. Solid models of a mandibular segment, crown, and dental implants were constructed using the computer-aided design (CAD) system (Pro-Engineering, PTC, New York, NY, U.S.A.) to create 3-D FE models from the data basis originally stemmed from CT images. The need for accurate FE models of the complete mandible (Fig.5) in realistic

simulation is becoming more acknowledged to evaluate an optimal biomechanical distribution of stresses in mandibular implant-supported fixed restorations both at the level of the prosthetic superstructure and at the level of the implant infrastructure.

4. Material Properties

Material properties greatly influence the stress and strain distribution in a structure. These properties can be modeled in FEA as isotropic, transversely isotropic, orthotropic, and anisotropic. The properties are the same in all directions, therefore, only two independent material constants of Young's modulus and Poisson's ratio exist in an isotropic material. In most reported studies, an assumption was made that the materials were homogenous and linearly isotropic. How to determine the complex cancellous pattern was very tough, so the cancellous bone network ignored in early FEA studies. Therefore, it was assumed that cancellous bone has a solid design inside the inner cortical bone shell. There are several methods to determine the physical properties of bone, such as tensile, compressive, bending, and torsion testing, pure shear tests, micro- and nano-indentation tests, acoustic tests, and back-calculation using FE models (Table1). The values 13.7 GPa and 1.37 GPa have been frequently used for the Young's modulus of cortical and cancellous bone, respectively. The original source for those values is a compressive test study on human vertebrae. However, compressive tests are subject to the confounding factors of proper specimen alignment and compliance of the loading fixture, which are not factors in ultrasonic pulse technique. Consequently, in the current study, cortical and cancellous bone were given a Young's modulus of 20.7 GPa and 14.8 GPa, respectively, according to the ultrasound study by Rho et al. Poisson's ratio were assumed to be 0.3 for both cortical and cancellous bone. Several studies incorporated simplified transversely isotropy (Table2) instead of orthotropy into their FE models demonstrated the significance of using anisotropy (transversely isotropy) on bone-implant interface stresses and peri-implant principal strains. It was concluded that anisotropy increased what were already high levels of stress and strain in the isotropic case by 20-30% in the cortical crest. In cancellous bone, anisotropy increased what were relatively low levels of interface stress in the isotropic case by three- to four folds. To incorporate more realistic anisotropic materials for bone tissues in maxilla or mandible, the FE model may employ fully orthotropy for compact bone and transversely isotropy for cancellous bone (Table 3), since they are currently available material property measurements of human mandible. Because of material properties for human maxillary bone were not available, this may influence the accuracy and applicability of the study results. However, by assigning fully orthotropic material to compact bone, the high quality anisotropic FE model of the segmental maxilla may bring us one important step closer toward realizing realistic maxilla related simulation. An orthotropic material has three planes of mirror symmetry and nine independent constants as compared to one axis of symmetry and five independent constants for transverse isotropy. Orthotropy is not in itself a problem for the finite element method. However, the cross-sectional shape of the mandible does not easily lend itself to the use of orthotropic material properties, for which the symmetry axes would presumably change from point to point, following the irregular elliptical shape of the mandibular cross section. A transversely isotropic material behaves identically in all planes perpendicular to the axis of symmetry. The unique symmetry axis for compact bone was along the mesio-distal direction with the bucco-lingual plane being a plane of elastic isotropy. The unique

symmetry axis for cancellous bone of the edentulous mandible was in the infero-superior direction with the anatomic transverse plane being a plane of elastic isotropy.

5. Boundary Conditions

Zero displacement constraints must be placed on some boundaries of the model to ensure an equilibrium solution. The constraints should be placed on nodes that are far away from the region of interest to prevent the stress or strain fields associated with reaction forces from overlapping with the bone-implant interface. In the maxillary FEA models, the nodes along the external lines of the cortical bone of oral and nasopharyngeal cavities were fixed in all directions (Fig.3). Most FEA studies modeling the mandible set the boundary condition was constrained in all directions at the nodes on mesial and distal borders.

Since only half of the model was meshed, symmetry boundary conditions were prescribed at the nodes on the symmetry plane. Models were constrained in all directions at the nodes on the mesial bone surface. Because of symmetry conditions, these constraints were also reproduced on the distal bone surface (Fig.6).

An individual geometry of the complete range of mandible was created, meanwhile the functions of the mastication muscles, ligaments and functional movement of temporomandibular joints simulated. The boundary conditions included constraining all three degrees of freedom at each of the nodes located at the joint surface of the condyles and the attachment regions of the masticatory muscles (masseter, temporalis, medial pterygoid, and lateral pterygoid) (Fig.7). Expanding the domain of the model can reduce the effect of inaccurate modeling of the boundary conditions. This, however, is at the expense of computing and modeling time. Modeling a 3-D mandibular model at distances greater than 4.2 mm mesially or distally from the implant did not result in any significant further yield in FEA accuracy.

6. Loading Conditions

Mastication involves a repeated pattern of cyclic impacts that causes loading to the implant components and distributes the force to the bone interface. When applying FE analysis to dental implants, it is important to consider not only axial loads and horizontal forces (moment-causing loads) but also a combined load (oblique occlusal force) because the latter represents more realistic masticatory pattern and will generate considerable localized stresses in compact bone. Bite force studies indicated considerable variation from one area of the mouth to another and from one individual to the next. In the premolar region, reported values of maximal bite force range from 181-608 N. Average forces of more than 800 N for male young adults and 600 N for female young adults have been recorded in the molar region. Small forces of 290 and 240 N, respectively, have been measured in the incisal region. The variation may be related to many factors, such as muscle size, bone shape, sex, age, degree of edentulism, and parafunction. In the maxillary anterior region, the occlusal force was assumed to be 178 N could not impair osseointegration or induce bone resorption may be appropriate (Fig.8). A 200-N vertical and a 40-N horizontal load were applied to the occlusal surface of the crown (Fig.9). These loads represent average means recorded on patients with endosseous implants. It should be noted that a great spectra of vertical loads/forces have been reported for patients with endosseous implants (means range :

91-284 N), and the loads appear to be related to the location of the implant, as well as to food consistency. In the previous studies, the locations for the force application were specifically described as cusp tip, distal fossa, and mesial fossa. When occlusal forces exerted from the masticatory muscles, the buccal functional cusps of the mandibular teeth will be forced to contact with central, distal, and mesial fossa. Hence, bite force applied to the occlusal surface of the crown may be more reasonable than the abutment of the implant.

7. Bone-implant interface

Analyzing force transfer at the bone-implant interface is an essential step in the overall analysis of loading, which determines the success or failure of an implant. It has long been recognized that both implant and bone should be stressed within a certain range for physiologic homeostasis. Overload can lead to bone resorption or fatigue failure of the implant, whereas underloading of the bone may cause disuse atrophy and subsequent bone loss. Most FEA models, the bone-implant interface was assumed to be perfect, simulating 100% osseointegration. This does not occur so exactly in clinical situations. Up until recently, linear static models have been employed extensively in finite element studies of dental implants. However, the validity of a linear static analysis is questionable for more realistic situations such as immediate loading.

Currently FEA programs provide several types of contact algorithms for simulation of contacts. Three different contact types defined in ANSYS—"bonded", "no separation", and "frictionless"—are used to describe the integration quality at the implant-compact bone interface. The "bonded" type simulates perfect osseointegration in which the implant and the surrounding compact bone are fully integrated so that neither sliding nor separation in the implant-bone interface is possible. The "no separation" type indicates an imperfect osseointegration in which separation at the contact interface is not allowed but frictionless sliding between the implant and compact bone may take place. The poorest osseointegration is modeled by a standard unilateral "frictionless" contact, which implies that a gap between the implant and the peri-implant compact bone may exist under an occlusal force. To obtain initial stability for the situation of immediate loading after implantation, it was modeled using nonlinear frictional contact elements, which allowed minor displacements between implant and bone. Under these conditions, the contact zone transfers pressure and tangential forces (i.e., friction), but no tension. The friction coefficient was set to 0.3. The friction between contact surfaces can also be modeled with contact algorithms. Ding's study was modeled using nonlinear frictional contact elements, which allow minor displacements between implant and bone to keep the implant stable and provide an excellent simulation of the implant-bone interface with immediate load.

8. Convergence Test

The p-element method in ANSYS was used for the convergence tests, and by this method the polynomial level (p-level) of the element shape functions was manipulated. This differs from the more traditional h-method in which the mesh must be refined to obtain a suitable convergence in displacement or stress results (Fig.10). It is difficult to obtain a suitable mesh of a 3-D object with irregular shaped volumes and refining such a mesh in a consistent manner to ensure convergence is a cumbersome process. By contrast, once a suitable mesh is

constructed in the p-method, it is kept unchanged while the polynomial level is increased from two to as high as eight until convergence is obtained. When an iterative solution method was used with a starting p-value of two and a tolerance of 1% for convergence checking, the analysis was considered to have converged if the global strain energy changed by less than 1%. Changing of the global strain energy was required to be less than 5% at a p-level of four at convergence could be also considered to have converged.

9. Validation

To validate the FE model, Sekine and coworkers measured the labiolingual mobility of 41 isolated osseointegrated implants in 8 human mandibles clinically using a displacement-measuring lever with electric strain gauges. The measuring point was 6 mm from the margin of bone shown on standardized x-rays of each implant. The load was increased linearly up to 20 N and observed implant displacement was 17 to 58 μm . The results of the FEA model could be compared with a real clinical situation, a similar load applied to the test implant in the study. This means that result of the FEA was similar to the clinical situation, thus the FE model was valid. The resulting level of implant displacement of Hsu's study was 17 μm for a high-density model and 19 μm for a low density bone model which revealed the calculated load-displacement values were close to values reported for osseointegrated implants in vivo. Therefore, an in vivo experiment could be conducted to verify the FEA results.

10. Statistical analysis

Statistical analysis has seldom been used in FEA. However, Hsu et al used a pair-wise t-test in his study to analyze results obtained from FE model. In this manuscript biomechanical performance of endodontically treated teeth restored with three post materials in three different length of post were evaluated with a 3-D FE model. The choice of the applicable stress representation criterion was based on an evaluation of the failure predictive potential of the analysis performed. The von Mises energetic criterion was then chosen as a better representative of a multiaxial stress state. These evaluations were carried out in three regions and 25 equally spaced points were sampled for plotting various pattern graphics as well as conducting statistical tests. A pair-wise t-test was applied to evaluate the difference among different groups. Statistical analysis was utilized properly to enrich the result and make the FEA meaningful.

11. Conclusion

With rapid improvements and developments of computer technology, the FEA has become a powerful technique in dental implant biomechanics because of its versatility in calculating stress distributions within complex structures. By understanding the basic theory, method, application, and limitations of FEA in implant dentistry, the clinician will be better equipped to interpret results of FEA studies and extrapolate these results to clinical situations. Thus, it is a helpful tool to evaluate the influence of model parameter variations once a basic model is correctly defined. Further research should focus in analyzing stress distributions under dynamic loading conditions of mastication, which would better mimic the actual clinical situation.

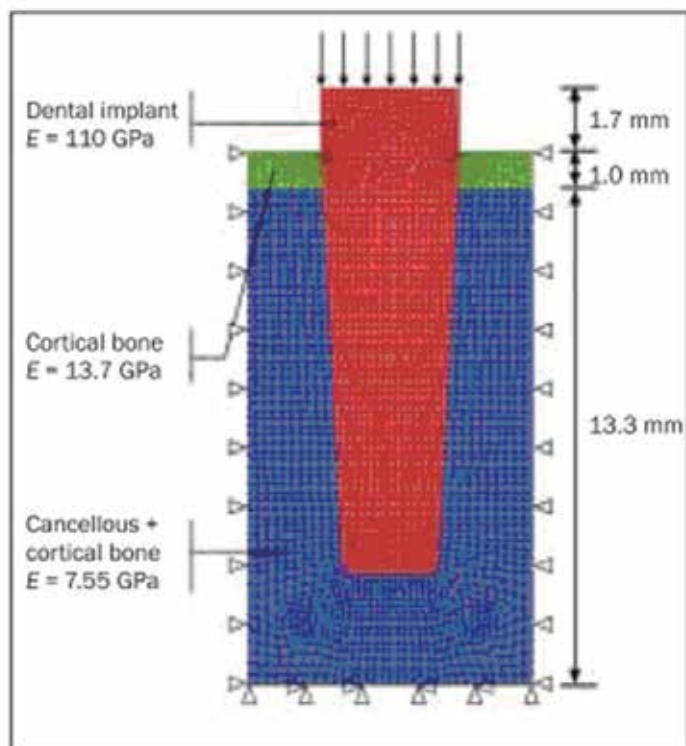


Fig. 1. The bone was modeled as a simplified rectangular configuration with the implant in 2-D FEA model (Courtesy from Shi L. et al. Int J Oral Maxillofac Implants 2007).

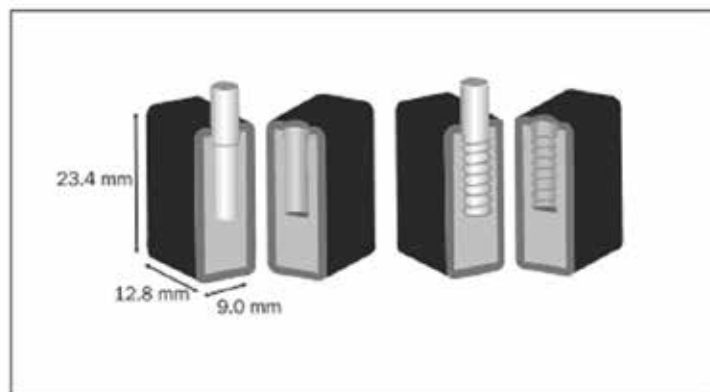


Fig. 2. The mandible was treated as a simplified segment as cancellous core surrounded by a 1.3-mm cortical layer with the overall dimensions of this block were 23.4 mm in height, 25.6 mm in mesiodistal length, and 9.0 mm in buccolingual width in 3-D FEA models (Courtesy from Tada S. et al. Int J Oral Maxillofac Implants 2003).

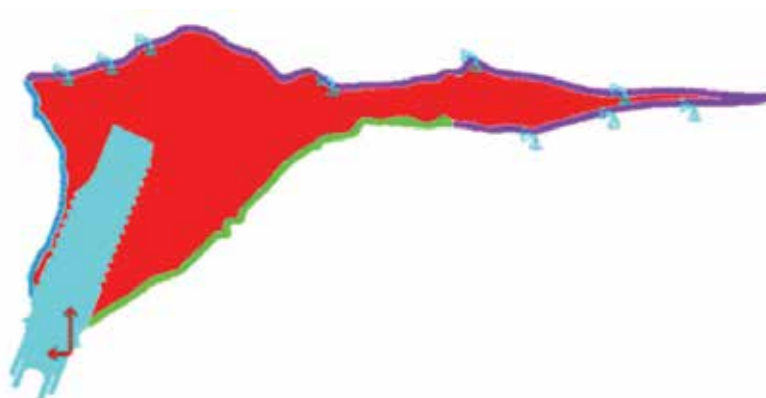


Fig. 3. The outline of the digital image was manually plotted and each point converted into x and y coordinates. The coordinates were finally imported into the ANSYS software as keypoints of the definitive image of the 2-D maxilla with implant (Courtesy from Saab XE et al. J Prosthet Dent 2007).

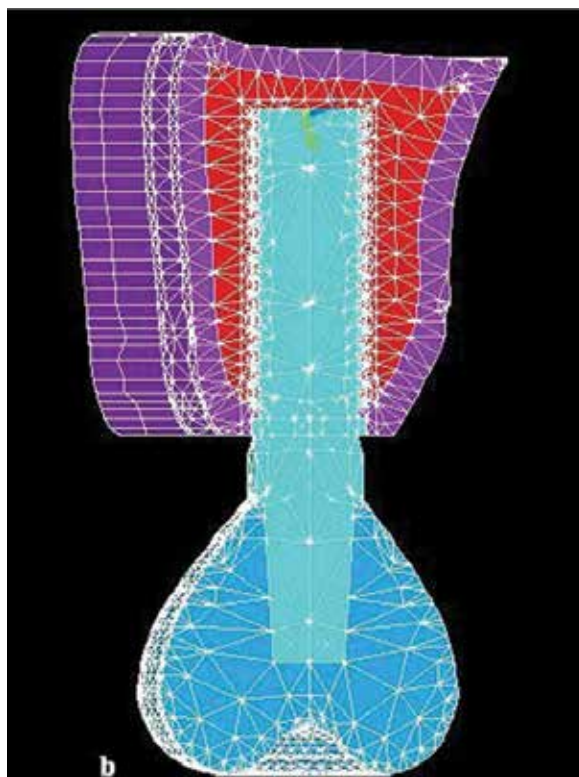


Fig. 4. Cross-sectional view on the symmetry plane of the meshed models with the implant embedded in the maxillary right first molar area and a gold alloy crown with 2-mm occlusal thickness was applied over the titanium abutment.



Fig. 5. A complete range of mandible reconstruction from CT and implants embedded in the posterior zone (Courtesy from Liao SH et al. Comput Med Imaging Graph 2008).

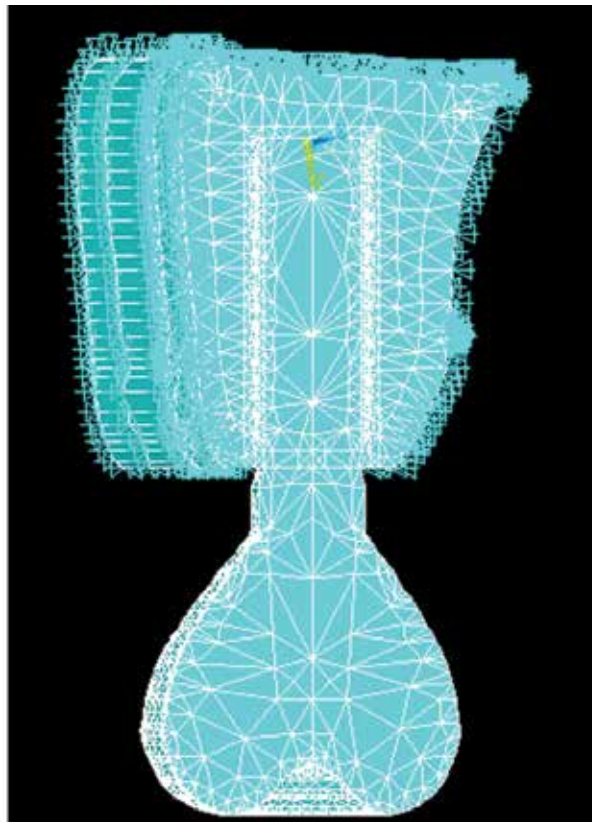


Fig. 6. Symmetry boundary conditions were prescribed at the nodes on the symmetry plane and the models were constrained in all directions at the nodes on the mesial and distal bone surface.

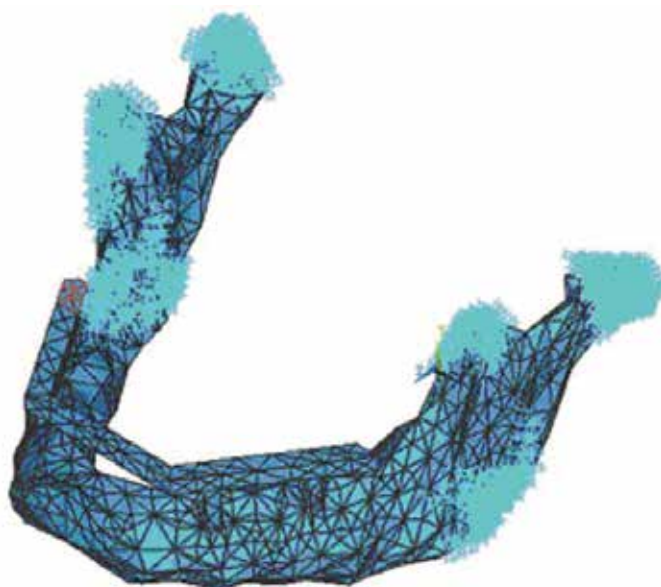


Fig. 7. All three degrees of freedom at each of the nodes located at the joint surface of the condyles and the attachment regions of the masticatory muscles (masseter, temporalis, medial pterygoid, and lateral pterygoid) were constrained (Courtesy from Nagasao T. et al. J Craniomaxillofac Surg 2002).

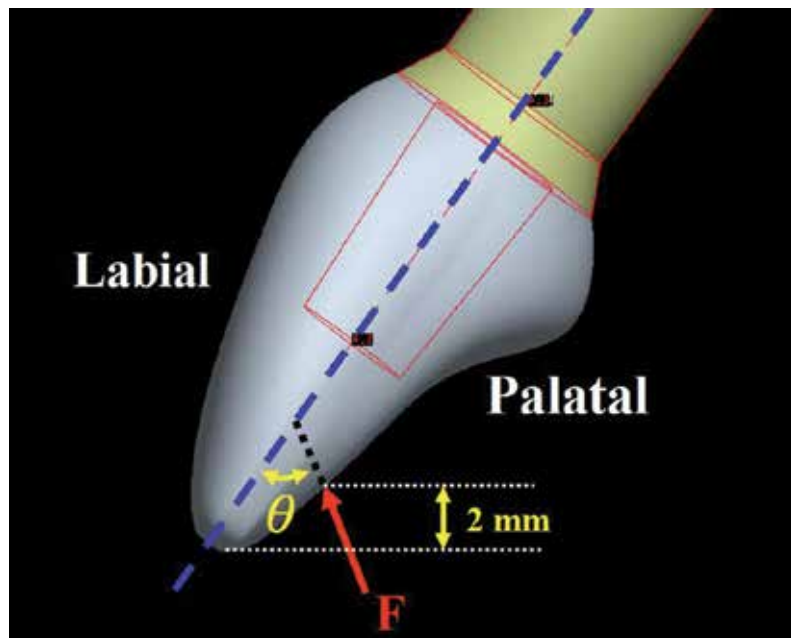


Fig. 8. In the maxillary anterior region, an occlusal load (F) of 178 N was applied on a node at the crown.

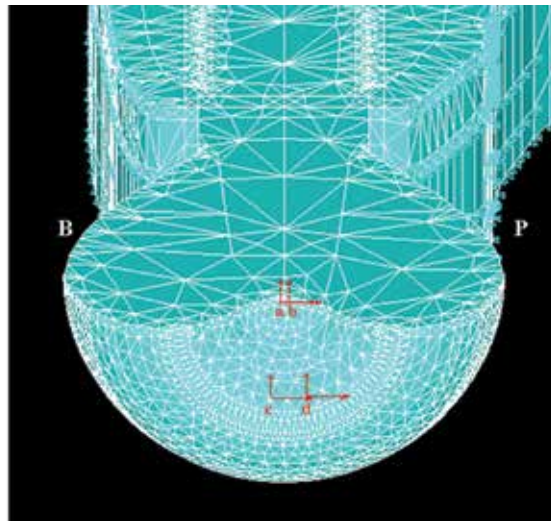


Fig. 9. Because a symmetric half model was used, loading was simulated by applying an oblique load (vertical load of 100 N and horizontal load of 20 N) from buccal to palatal at four different locations on the central (a, b) and distal fossa (c, d) of the crown.

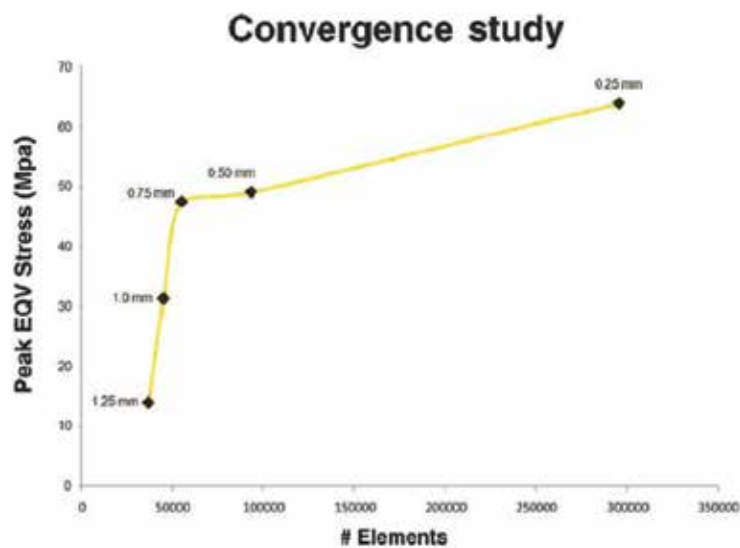


Fig. 10. Influence of element size (1.25, 1.0, 0.75, 0.50, and 0.25 mm) on bone mesh density and peak equivalent (EQV) stress in bone model (Courtesy from Pessoa RS et al. Clin Implant Dent Relat Res 2009).

Study	Compact bone E (Gpa)	Poisson's ratio (v)	Cancellous bone E (Gpa)	Poisson's ratio (v)
Geng et al ³⁷	13.4	0.3	1.37	0.31
Borchers and Reichart ³⁸	13.7	0.3	1.37	0.3
Meijer et al ³⁹	13.7	0.3	1.37	0.3
Menicucci et al ⁴⁰	13.7	0.3	1.37	0.3
Teixeira et al ⁴¹	13.7	0.3	1.37	0.3
Benzing et al ⁴²	15	0.25	2	0.495
Stegarioiu et al ⁴³	15	0.3	1.5	0.3
Ciftci and Canay ⁴⁴	14	0.3	1	0.3
Siegele and Soltesz ⁴⁵	20	0.3	2	0.3
Canay et al ⁴⁶	19.73	0.3		
Geng et al ⁴⁷	13.4	0.3	1.37	0.31
	10	0.3	1.37	0.31
	7.5	0.3	1.37	0.31
	5	0.3	1.37	0.31
	1.37	0.3	1.37	0.31

Table 1. Young's modulus (E) and Poisson's ratio (v) of compact and cancellous bone used in previous FEA studies.

Material	Young's modulus E (Mpa)		Poisson's ratio (v)		Shear modulus G (Mpa)	
compact bone	Ex	12,600	vxy	0.300		
			vyz	0.253	Gxy	4,850
			vxz	0.253		
	Ey	12,600	vyx	0.300	Gyz	5,700
			vzy	0.390		
			vzx	0.390	Gxz	5,700
cancellous bone	Ex	1,148	vxy	0.055		
			vyz	0.010	Gxy	68
			vxz	0.322		
	Ey	210	vyx	0.010	Gyz	68
			vzy	0.055		
			vzx	0.322	Gxz	434

Table 2. Material properties used in the transversely isotropic model (Courtesy from Huang HL et al. Clin Oral Implants Res 2005).

	E_y	E_x	E_z	G_{yx}	G_{yz}	G_{xz}	ν_{yx}	ν_{yz}	ν_{xz}
Com.	12.5	17.9	26.6	4.5	5.3	7.1	0.18	0.31	0.28
Can.	0.21	1.148	1.148	0.068	0.068	0.434	0.055	0.055	0.322

Table 3. Anisotropy elastic coefficients for compact (Com.) and cancellous (Can.) bone.

E_i represents Young's modulus (GPa); G_{ij} represents shear modulus (GPa); ν_{ij} represents Poisson's ratio.

* The y-direction is infero-superior, the x-direction is medial-lateral, and the z-direction is anterior-posterior (Courtesy from Chang CL et al. Int J Oral Maxillofac Implants 2010).

12. References

- Assuncao WG, Gomes EA, Barao VA, de Sousa EA. Stress analysis in simulation models with or without implant threads representation. *Int J Oral Maxillofac Implants* 2009;24:1040-1044.
- Akca K, Iplikcioglu H. Evaluation of the effect of the residual bone angulation on implant-supported fixed prosthesis in mandibular posterior edentulism. Part II: 3-D finite element stress analysis. *Implant Dent* 2001;10:238-245.
- Baggi L, Cappelloni I, Di Girolamo M, Maceri F, Vairo G. The influence of implant diameter and length on stress distribution of osseointegrated implants related to crestal bone geometry: a three-dimensional finite element analysis. *J Prosthet Dent* 2008;100:422-431.
- Bell GH, Dunbar O, Beck JS, Gibb A. Variations in strength of vertebrae with age and their relation to osteoporosis. *Calcif Tissue Res* 1967;1:75-86.
- Benzing UR, Gall H, Weber H. Biomechanical aspects of two different implant-prosthetic concepts for edentulous maxillae. *Int J Oral Maxillofac Implants* 1995;10:188-198.
- Bidez MW, Misch CE. Force transfer in implant dentistry: basic concepts and principles. *J Oral Implantol* 1992;18:264-274.
- Borchers L, Reichart P. Three-dimensional stress distribution around a dental implant at different stages of interface development. *J Dent Res* 1983;62:155-159.
- Branemark PI, Adell R, Breine U, Hansson BO, Lindstrom J, Ohlsson A. Intra-osseous anchorage of dental prostheses. I. Experimental studies. *Scand J Plast Reconstr Surg* 1969;3:81-100.
- Branemark PI, Zarb GA, Albrektsson T (eds). *Tissue-integrated prostheses: Osseointegration in clinical dentistry*. Chicago: Quintessence, 1985:129.
- Brosh T, Pilo R, Sudai D. The influence of abutment angulation on strains and stresses along the implant/bone interface: comparison between two experimental techniques. *J Prosthet Dent* 1998;79:328-334.
- Caglar A, Aydin C, Ozen J, Yilmaz C, Korkmaz T. Effects of mesiodistal inclination of implants on stress distribution in implant-supported fixed prostheses. *Int J Oral Maxillofac Implants* 2006;21:36-44.
- Canay S, Hersek N, Akpınar I, Asik Z. Comparison of stress distribution around vertical and angled implants with finite-element analysis. *Quintessence Int* 1996;27:591-598.

- Chang CL, Chen CS, Hsu ML. Biomechanical effect of platform switching in implant dentistry: a three-dimensional finite element analysis. *Int J Oral Maxillofac Implants* 2010;25:295-304.
- Chun HJ, Shin HS, Han CH, Lee SH. Influence of implant abutment type on stress distribution in bone under various loading conditions using finite element analysis. *Int J Oral Maxillofac Implants* 2006;21:195-202.
- Ciftci Y, Canay S. Stress distribution on the metal framework of the implant-supported fixed prosthesis using different veneering materials. *Int J Prosthodont* 2001;14:406-411.
- Clelland NL, Lee JK, Bimbenet OC, Brantley WA. A three-dimensional finite element stress analysis of angled abutments for an implant placed in the anterior maxilla. *J Prosthodont* 1995;4:95-100.
- Cowin SC. *Bone mechanics handbook*. 2nd ed. Boca Raton (FL): CRC Press;2001. p. 1.1-1.23.
- Cruz M, Wassall T, Toledo EM, Barra LP, Lemonge AC. Three-dimensional finite element stress analysis of a coneiform-geometry implant. *Int J Oral Maxillofac Implants* 2003;18:675-684.
- Dalkiz M, Zor M, Aykul H, Toparli M, Aksoy S. The three-dimensional finite element analysis of fixed bridge restoration supported by the combination of teeth and osseointegrated implants. *Implant Dent* 2002;11:293-300.
- Ding X, Liao SH, Zhu XH, Zhang XH, Zhang L. Effect of diameter and length on stress distribution of the alveolar crest around immediate loading implants. *Clin Implant Dent Relat Res* 2009;11:279-287.
- Ding X, Zhu XH, Liao SH, Zhang XH, Chen H. Implant-bone interface stress distribution in immediately loaded implants of different diameters: a three-dimensional finite element analysis. *J Prosthodont* 2009;18:393-402.
- Eckert SE, Choi YG, Sanchez AR, Koka S. Comparison of dental implant systems: quality of clinical evidence and prediction of 5-year survival. *Int J Oral Maxillofac Implants* 2005;20:406-415.
- Ferrigno N, Laureti M, Fanali S, Grippaudo G. A long-term follow-up study of non-submerged ITI implants in the treatment of totally edentulous jaws. Part I: Ten-year life table analysis of a prospective multicenter study with 1286 implants. *Clin Oral Implants Res* 2002;13:260-273.
- Geng JP, Tan KB, Liu GR. Application of finite element analysis in implant dentistry: a review of the literature. *J Prosthet Dent* 2001;85:585-598.
- Geng JP, Ma QS, Xu W, Tan KB, Liu GR. Finite element analysis of four thread-form configurations in a stepped screw implant. *J Oral Rehabil* 2004;31:233-239.
- Geng JP, Xu DW, Tan KB, Liu GR. Finite element analysis of an osseointegrated stepped screw dental implant. *J Oral Implantol* 2004;30:223-233.
- Hagberg C. Assessment of bite force: a review. *J Craniomandib Disord* 1987;1:162-169.
- Heckmann SM, Karl M, Wichmann MG, Winter W, Graef F, Taylor TD. Loading of bone surrounding implants through three-unit fixed partial denture fixation: a finite-element analysis based on in vitro and in vivo strain measurements. *Clin Oral Implants Res* 2006;17:345-350.
- Holmgren EP, Seckinger RJ, Kilgren LM, Mante F. Evaluating parameters of osseointegrated dental implants using finite element analysis--a two-dimensional comparative study examining the effects of implant diameter, implant shape, and load direction. *J Oral Implantol* 1998;24:80-88.

- Hsu JT, Fuh LJ, Lin DJ, Shen YW, Huang HL. Bone strain and interfacial sliding analyses of platform switching and implant diameter on an immediately loaded implant: experimental and three-dimensional finite element analyses. *J Periodontol* 2009;80:1125-1132.
- Hsu ML, Chen FC, Kao HC, Cheng CK. Influence of off-axis loading of an anterior maxillary implant: a 3-dimensional finite element analysis. *Int J Oral Maxillofac Implants* 2007;22:301-309.
- Hsu ML, Chen CS, Chen BJ, Huang HH, Chang CL. Effects of post materials and length on the stress distribution of endodontically treated maxillary central incisors: a 3D finite element analysis. *J Oral Rehabil* 2009;36:821-830.
- Huang HL, Huang JS, Ko CC, Hsu JT, Chang CH, Chen MY. Effects of splinted prosthesis supported a wide implant or two implants: a three-dimensional finite element analysis. *Clin Oral Implants Res* 2005;16:466-472.
- Huang HL, Chang CH, Hsu JT, Fallgatter AM, Ko CC. Comparison of implant body designs and threaded designs of dental implants: a 3-dimensional finite element analysis. *Int J Oral Maxillofac Implants* 2007;22:551-562.
- Juodzbaly G, Kubilius R, Eidukynas V, Raustia AM. Stress distribution in bone: single-unit implant prostheses veneered with porcelain or a new composite material. *Implant Dent* 2005;14:166-175.
- Kao HC, Gung YW, Chung TF, Hsu ML. The influence of abutment angulation on micromotion level for immediately loaded dental implants: a 3-D finite element analysis. *Int J Oral Maxillofac Implants* 2008;23:623-630.
- Kitamura E, Stegaroiu R, Nomura S, Miyakawa O. Biomechanical aspects of marginal bone resorption around osseointegrated implants: considerations based on a three-dimensional finite element analysis. *Clin Oral Implants Res* 2004;15:401-412.
- Kong L, Hu K, Li D, Song Y, Yang J, Wu Z, et al. Evaluation of the cylinder implant thread height and width: a 3-dimensional finite element analysis. *Int J Oral Maxillofac Implants* 2008;23:65-74.
- Kong L, Gu Z, Li T, Wu J, Hu K, Liu Y, et al. Biomechanical optimization of implant diameter and length for immediate loading: a nonlinear finite element analysis. *Int J Prosthodont* 2009;22:607-615.
- Liao SH, Tong RF, Dong JX. Influence of anisotropy on peri-implant stress and strain in complete mandible model from CT. *Comput Med Imaging Graph* 2008;32:53-60.
- Li T, Kong L, Wang Y, Hu K, Song L, Liu B, et al. Selection of optimal dental implant diameter and length in type IV bone: a three-dimensional finite element analysis. *Int J Oral Maxillofac Surg* 2009;38:1077-1083.
- Lin CL, Chang SH, Wang JC, Chang WJ. Mechanical interactions of an implant/tooth-supported system under different periodontal supports and number of splinted teeth with rigid and non-rigid connections. *J Dent* 2006;34:682-691.
- Lin CL, Wang JC, Ramp LC, Liu PR. Biomechanical response of implant systems placed in the maxillary posterior region under various conditions of angulation, bone density, and loading. *Int J Oral Maxillofac Implants* 2008;23:57-64.
- Maeda Y, Miura J, Taki I, Sogo M. Biomechanical analysis on platform switching: is there any biomechanical rationale? *Clin Oral Implants Res* 2007;18:581-584.
- Meijer GJ, Starmans FJ, de Putter C, van Blitterswijk CA. The influence of a flexible coating on the bone stress around dental implants. *J Oral Rehabil* 1995;22:105-111.

- Menicucci G, Lorenzetti M, Pera P, Preti G. Mandibular implant-retained overdenture: finite element analysis of two anchorage systems. *Int J Oral Maxillofac Implants* 1998;13:369-376.
- Mellal A, Wiskott HW, Botsis J, Scherrer SS, Belser UC. Stimulating effect of implant loading on surrounding bone. Comparison of three numerical models and validation by in vivo data. *Clin Oral Implants Res* 2004;15:239-248.
- Morneburg TR, Proschel PA. Measurement of masticatory forces and implant loads: a methodologic clinical study. *Int J Prosthodont* 2002;15:20-27.
- Morneburg TR, Proschel PA. In vivo forces on implants influenced by occlusal scheme and food consistency. *Int J Prosthodont* 2003;16:481-486.
- Nagasao T, Kobayashi M, Tsuchiya Y, Kaneko T, Nakajima T. Finite element analysis of the stresses around endosseous implants in various reconstructed mandibular models. *J Craniomaxillofac Surg* 2002;30:170-177.
- Nagasao T, Kobayashi M, Tsuchiya Y, Kaneko T, Nakajima T. Finite element analysis of the stresses around fixtures in various reconstructed mandibular models--part II (effect of horizontal load). *J Craniomaxillofac Surg* 2003;31:168-175.
- Natali AN, Pavan PG, Ruggero AL. Analysis of bone-implant interaction phenomena by using a numerical approach. *Clin Oral Implants Res* 2006;17:67-74.
- Nomura T, Powers MP, Katz JL, Saito C. Finite element analysis of a transmandibular implant. *J Biomed Mater Res B Appl Biomater* 2007;80:370-376.
- O'Mahony AM, Williams JL, Katz JO, Spencer P. Anisotropic elastic properties of cancellous bone from a human edentulous mandible. *Clin Oral Implants Res* 2000;11:415-421.
- O'Mahony AM, Williams JL, Spencer P. Anisotropic elasticity of cortical and cancellous bone in the posterior mandible increases peri-implant stress and strain under oblique loading. *Clin Oral Implants Res* 2001;12:648-657.
- Perry J, Lenchewski E. Clinical performance and 5-year retrospective evaluation of Frialit-2 implants. *Int J Oral Maxillofac Implants* 2004;19:887-891.
- Pessoa RS, Muraru L, Junior EM, Vaz LG, Sloten JV, Duyck J, et al. Influence of Implant Connection Type on the Biomechanical Environment of Immediately Placed Implants - CT-Based Nonlinear, Three-Dimensional Finite Element Analysis. *Clin Implant Dent Relat Res* 2009.
- Petrie CS, Williams JL. Comparative evaluation of implant designs: influence of diameter, length, and taper on strains in the alveolar crest. A three-dimensional finite-element analysis. *Clin Oral Implants Res* 2005;16:486-494.
- Petrie CS, Williams JL. Probabilistic analysis of peri-implant strain predictions as influenced by uncertainties in bone properties and occlusal forces. *Clin Oral Implants Res* 2007;18:611-619.
- Pierrisnard L, Hure G, Barquins M, Chappard D. Two dental implants designed for immediate loading: a finite element analysis. *Int J Oral Maxillofac Implants* 2002;17:353-362.
- Rho JY, Ashman RB, Turner CH. Young's modulus of trabecular and cortical bone material: ultrasonic and microtensile measurements. *J Biomech* 1993;26:111-119.
- Rieger MR, Mayberry M, Brose MO. Finite element analysis of six endosseous implants. *J Prosthet Dent* 1990;63:671-676.
- Rosenberg ES, Torosian JP, Slots J. Microbial differences in 2 clinically distinct types of failures of osseointegrated implants. *Clin Oral Implants Res* 1991;2:135-144.

- Saab XE, Griggs JA, Powers JM, Engelmeier RL. Effect of abutment angulation on the strain on the bone around an implant in the anterior maxilla: a finite element study. *J Prosthet Dent* 2007;97:85-92.
- Sahin S, Cehreli MC, Yalcin E. The influence of functional forces on the biomechanics of implant-supported prostheses--a review. *J Dent* 2002;30:271-282.
- Satoh T, Maeda Y, Komiyama Y. Biomechanical rationale for intentionally inclined implants in the posterior mandible using 3D finite element analysis. *Int J Oral Maxillofac Implants* 2005;20:533-539.
- Schrotenboer J, Tsao YP, Kinariwala V, Wang HL. Effect of microthreads and platform switching on crestal bone stress levels: a finite element analysis. *J Periodontol* 2008;79:2166-2172.
- Schwartz-Dabney CL, Dechow PC. Edentulation alters material properties of cortical bone in the human mandible. *J Dent Res* 2002;81:613-617.
- Schwartz-Dabney CL, Dechow PC. Variations in cortical material properties throughout the human dentate mandible. *Am J Phys Anthropol* 2003;120:252-277.
- Sekine H, Komiyama Y, Hotta H, Yoshida Y. Mobility characteristics and tactile sensitivity of osseointegrated fixture-supporting systems. In: van Steenberghe D (ed). *Tissue integration in Oral Maxillofacial Reconstruction*. Amsterdam: Elsevier, 1986:306-332.
- Sertgoz A. Finite element analysis study of the effect of superstructure material on stress distribution in an implant-supported fixed prosthesis. *Int J Prosthodont* 1997;10:19-27.
- Sevimay M, Turhan F, Kilcarslan MA, Eskitascioglu G. Three-dimensional finite element analysis of the effect of different bone quality on stress distribution in an implant-supported crown. *J Prosthet Dent* 2005;93:227-234.
- Shi L, Li H, Fok AS, Ucer C, Devlin H, Horner K. Shape optimization of dental implants. *Int J Oral Maxillofac Implants* 2007;22:911-920.
- Siegele D, Soltesz U. Numerical investigations of the influence of implant shape on stress distribution in the jaw bone. *Int J Oral Maxillofac Implants* 1989;4:333-340.
- Stegarioiu R, Sato T, Kusakari H, Miyakawa O. Influence of restoration type on stress distribution in bone around implants: a three-dimensional finite element analysis. *Int J Oral Maxillofac Implants* 1998;13:82-90.
- Tada S, Stegaroiu R, Kitamura E, Miyakawa O, Kusakari H. Influence of implant design and bone quality on stress/strain distribution in bone around implants: a 3-dimensional finite element analysis. *Int J Oral Maxillofac Implants* 2003;18:357-368.
- Teixeira ER, Sato Y, Akagawa Y, Shindoi N. A comparative evaluation of mandibular finite element models with different lengths and elements for implant biomechanics. *J Oral Rehabil* 1998;25:299-303.
- Tie Y, Wang DM, Ji T, Wang CT, Zhang CP. Three-dimensional finite-element analysis investigating the biomechanical effects of human mandibular reconstruction with autogenous bone grafts. *J Craniomaxillofac Surg* 2006;34:290-298.
- Tonetti MS. Determination of the success and failure of root-form osseointegrated dental implants. *Adv Dent Res* 1999;13:173-180.
- Van Oosterwyck H, Duyck J, Vander Sloten J, Van der Perre G, De Cooman M, Lievens S, et al. The influence of bone mechanical properties and implant fixation upon bone loading around oral implants. *Clin Oral Implants Res* 1998;9:407-418.

- Van Oosterwyck H, Duyck J, Vander Sloten J, Van Der Perre G, Naert I. Peri-implant bone tissue strains in cases of dehiscence: a finite element study. *Clin Oral Implants Res* 2002;13:327-333.
- Wakabayashi N, Ona M, Suzuki T, Igarashi Y. Nonlinear finite element analyses: advances and challenges in dental applications. *J Dent* 2008;36:463-471.
- Waltimo A, Kemppainen P, Kononen M. Maximal contraction force and endurance of human jaw-closing muscles in isometric clenching. *Scand J Dent Res* 1993;101:416-421.
- Waltimo A, Kononen M. Bite force on single as opposed to all maxillary front teeth. *Scand J Dent Res* 1994;102:372-375.
- Yang J, Xiang HJ. A three-dimensional finite element study on the biomechanical behavior of an FGBM dental implant in surrounding bone. *J Biomech* 2007;40:2377-2385.
- Zampelis A, Rangert B, Heijl L. Tilting of splinted implants for improved prosthodontic support: a two-dimensional finite element analysis. *J Prosthet Dent* 2007;97:S35-43.
- Zarone F, Apicella A, Nicolais L, Aversa R, Sorrentino R. Mandibular flexure and stress build-up in mandibular full-arch fixed prostheses supported by osseointegrated implants. *Clin Oral Implants Res* 2003;14:103-114.

Finite element analysis for dental prosthetic design

Akikazu Shinya, DDS, PhD

Department of Crown and Bridge, School of Life Dentistry at Tokyo, The Nippon Dental University, Japan, Honorary Assistant Professor, Dental Materials Science Unit, Faculty of Dentistry, The University of Hong Kong and Research scientist, Department of Biomaterials Science, Institute of Dentistry and BioCity Turku Biomaterials Research Program, University of Turku

Daiichiro Yokoyama, DDS, PhD

Department of Crown and Bridge, School of Life Dentistry at Tokyo, The Nippon Dental University, 1-9-20 Fujimi, Chiyoda-ku, Tokyo 102-8158, Japan.

1. Background

There is a limited amount of dental research which is focused on the aforementioned aspects. In a finite element (FE) analysis, a large structure is divided into a number of small simple shaped elements, for which individual deformation (strain and stress) can be more easily calculated than for the whole undivided large structure. Using the traditional biophysical knowledge database in a rational validation process¹⁾, the use of FE analysis in the dental field has been significantly refined during the last decade²⁻⁹⁾. However, fixed prosthesis cannot be assimilated to a simplified geometric representation due to both their anatomical shape and layered structure. Sophisticated techniques have therefore been developed to refine geometry acquisition, such as the recreation and digitization of planar outlines. In general, a 3D FE modeling is a complicated process which is often the most time-consuming for the scientists. In addition, this process is prone to errors and simplifications which may induce faulty predictions.

Hybrid composite are clinically applicable to inlays, onlays, crowns and by the recent experiences of fiber reinforced composites (FRC) composite resins have created a new era in metal-free, esthetic dentistry¹⁰⁻¹³⁾. (Figure 1 and 2) The most commonly used FRC contains continuous unidirectional glass fibers, typically E (electrical) glass embedded in dimethacrylate or semi-IPN (interpenetrating polymer network) matrix of dimethacrylate and polymethylmethacrylate. Flexural modulus of metal alloys is generally higher than that of FRC. However, there are differences in flexural modulus between FRC as well. Highly cross-linked polymer matrix made of dimetharylates provides higher flexural modulus for FRC than obtained with semi-IPN matrix with similar fibers. It has been suggested that there are two factors which explain the better clinical function of FRC FPD compared to their metal alloy framework counterparts. Firstly, there are studies showing that composite resins adhere initially better to the FRC framework with semi-IPN matrix than to the metal alloys.



Fig. 1. FRC-FPDs in the region of the maxillary left lateral incisor.

The clinical applications for FRC have been investigated from different points of view in terms of the quantity of fiber for reinforcement, matrix coupling, biocompatibility, fatigue property, the type of fiber and bending strength with various testing methods ¹⁴⁻²⁷).

Designs for fiber framework for posterior FPD have been extensively investigated using FE method and FE analysis, clarifying that the reinforcement effects differ and depend on the position of the fiber framework ^{8,28}). The structure of framework of FRC FPD consists of (1) main framework which supports the FPD against vertical loads, (2) Bonding wings or full-coverage crowns on abutments which increase resistance against dislodgement, and (3) of additional fibers placed to support pontic against delamination.

In anterior region, curving the main framework labially could provide alternative for using additional fibers. Anterior and posterior teeth play different roles in masticatory movement. While chewing, the maxillary and mandibular anterior teeth come into contact in the edge-to-edge occlusal position, after a sliding movement on the lingual surface, and then progress into the intercuspatal position. Prosthetic appliances placed in the anterior region are subjected to the influence of the diverse occlusal contacts of maxillary and mandibular teeth, as well as to the occlusal force of the masticatory movement. Due to the different crown contours and loading conditions, the mechanical behavior of FRC-FPD in the anterior region is quite different from that of the posterior region.

The present study investigated the design of FRC framework in order to obtain its maximum reinforcement effect for anterior and posterior FRC FPD.



Fig. 2. Frontal view of the same patient.

2. Mechanical approach for designing of an anterior FRC FPD

Introduction

The optimal design of an FRC framework was obtained the maximum reinforcement for anterior FPD. FRC frameworks were designed using three different variations in the pontic. The results were compared with the control, a hybrid composite FPD without any fiber reinforcement, using the 3-D FE method.

Materials and Methods

Definition of structures: Preparation of the C-FPD FE Model

To create an FE model, a three-unit FPD replica of the maxillary right central incisor to maxillary right canine replacing the maxillary right lateral incisor was fabricated. The replica (D51-SC41, Nisshin) was made in accordance with the Textbook of Dental Anatomy²⁹⁾ and the following anatomic measurements were used: mesiodistal distance of the FPD (23.5 mm) and crown length (11.0 mm). Preparation designs are shown in Figure 3. A 1mm circumferential reduction shoulder preparation and a 2-mm incisal reduction were prepared for an assumed jacket crown. Because cementum, periodontal regiment, and alveolar bone were thought to have only a slight influence on the magnitude and distribution of stresses, they were ignored and a complete FPD-abutment interface was established.^{30–33)} A replica was measured at 0.25-mm intervals using a 3-D dental computer-aided design unit (Dental Cadim, ADVANCE) and the contact scanning method. The preprocessor of an FE analysis program (ANSYS 10.0, ANSYS) generated point clouds describing the replica surface, which were used as input data for the FE model (Figure 4).

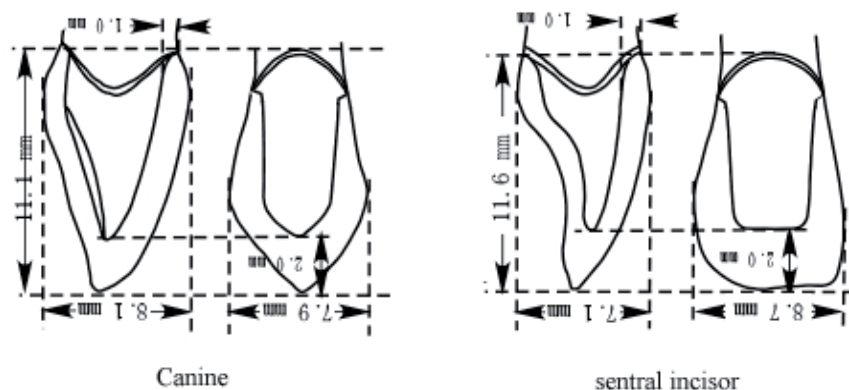


Fig. 3. The size of the abutment equipment (*left*) Maxillary right central incisor; (*right*) maxillary right canine.

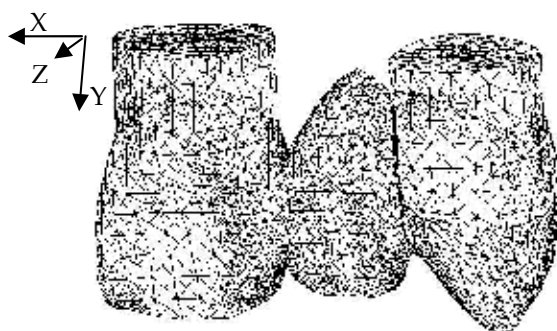


Fig. 4. External form of the FPD FE model (lingual view).

Preparation of the FRC-FPD FE Models

The FRC-FPD FE models were composed of the C-FPD with an FRC framework (3.0 mm wide by 1.0 mm high). The mesiodistal distance of the FRC framework was constructed to extend from the mesial side of the central incisor distally to the canine, covering both of the abutment teeth. The distance between the labial side of mesiodistal center of the pontic and the fiber framework was defined as A. Figure 5 shows the radius of curvature (r) varied with the value of A and both connectors. Accordingly, three different FRC-FPD FE models were constructed depending on the radius of curvature obtained: FRC1 ($A = 0.5$ mm, $r = 3.5$ mm, labial side, curved line), FRC2 ($A = 1.5$ mm, $r = 5.0$ mm, at the center, curved line), and FRC3 ($A = 3.0$ mm, $r = 0.0$ mm, lingual side, straight line). Figure 6 demonstrates three different curvatures of the FRC framework in the pontic of the FE model (ie, labially, at the center, and lingually).

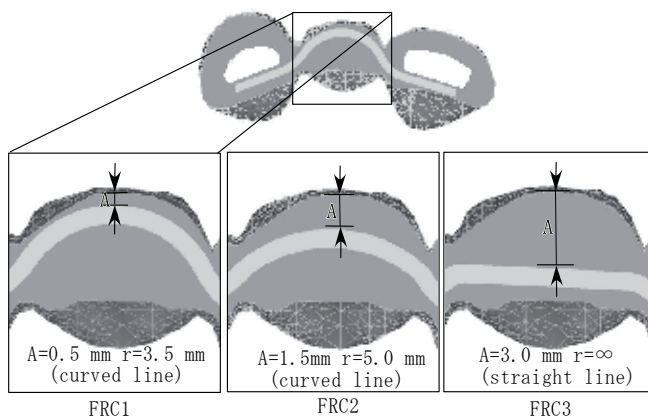


Fig. 5. Design of the FRC framework (horizontal cross section).

FRC1, $a = 0.5$ mm, $r = 3.5$ mm;

FRC2, $a = 1.5$ mm, $r = 5.0$ mm;

FRC3, $a = 3.0$ mm, $r = 0.0$ mm.

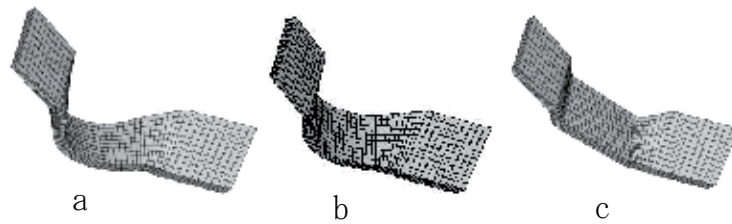


Fig. 6. Three different glass fiber framework

- a: FRC1 labial side, curved line
 b: FRC2 at the center, curved line
 c: FRC3 lingual side, straight line

Material properties, load case and data processing

The properties of the materials used for this FE analysis are listed in Table 1. Most of these values were determined according to previous literature surveys.^{34–38}) A hybrid composite with isotropic material properties (Estenia, Kuraray Medical) was used as a veneering material. The FRC framework was constructed from unidirectional glass fiber (everStick, StickTech) with anisotropic material properties. In this FE analysis, orientation of the fiber was set as the x axis. As such, material properties of the fiber framework in the x axis were set to have higher values (46 GPa), whereas in the y and z axes they were to have lower values (7GPa), thus representing the anisotropic properties(Figure 7).

	Young's modulus (MPa)	Poisson's ratio	Shear modulus (MPa)
Hybrid composite resin	2.20×10^4	0.27	
Dentin	1.80×10^4	0.31	
Pulp	2.10	0.45	
Glass fiber			
Longitudinal	4.60×10^4	0.39	1.65×10^4
Transverse	0.70×10^4	0.29	0.27×10^4

Table 1. Material properties ^{34–38})

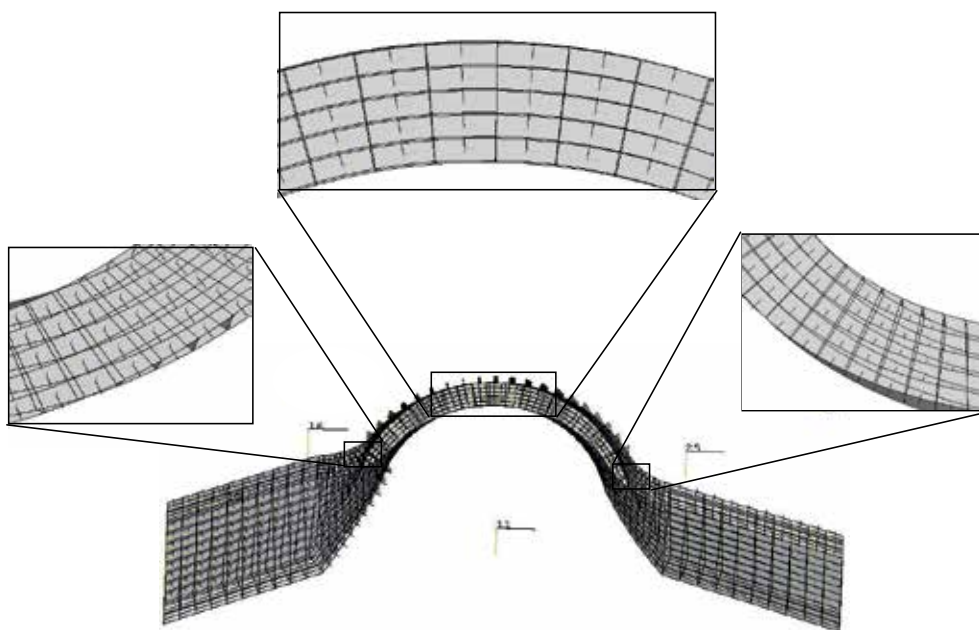


Fig. 7. A finite element model of glassfiber frame

Hexagonal elements with 20 nodes were selected for the anisotropic fiber framework, whereas tetrahedral elements with 10 nodes represented the isotropic materials. Four different FE models were constructed in this study: C-FPD (49,450 elements; 74,028 nodes), FRC1 (50,197 elements; 79,367 nodes), FRC2 (49,884 elements; 78,678 nodes), and FRC3 (48,886 elements; 77,340 nodes). As an anisotropic material, the FRC framework offers an exceptionally high elastic modulus along the orientation of the fibers (Table 1). To exhibit the intrinsic material properties of the glass fiber, a new local coordinate system, in addition to the rectangular coordinate system, with a different point of origin needed to be established. Orientation of the fibers coincided with the major axial direction of the coordinates.

Figure 8 shows the boundary and loading conditions of the FRC framework. In the present study, three different loading conditions were used to simulate edge-to-edge occlusion (1), centric occlusion (2), and a deep overbite (3), as observed during occlusion in the anterior region. A lateral load of 154 N was derived from the maximum occlusal force on healthy permanent teeth and was applied to the three different loading points of the pontic at an angle of 135 degrees from the lingual side (ie, 0 mm, 3 mm, and 6 mm from the incisal edge equidistant mesiodistally, each representing loading conditions 1, 2, and 3, respectively). The final element on the x, y, and z axes of the abutment base was assumed fixed, thereby defining the boundary conditions. FE analysis was presumed to be linear static. FE model construction and FE analysis were performed on a PC workstation (Precision Workstation 530, Dell) using ANSYS 10.0.

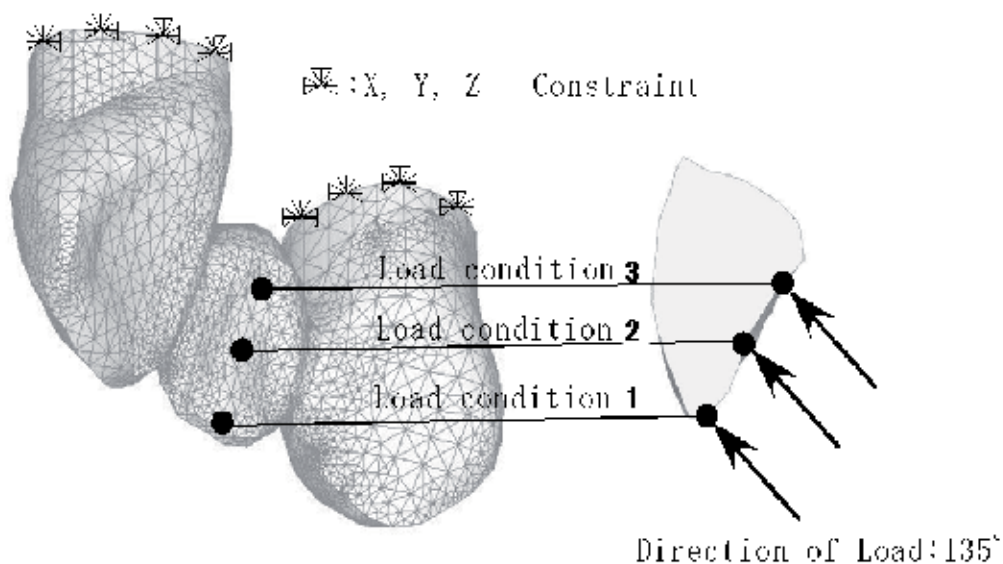


Fig. 8. Boundary conditions.

- Condition 1: 0 mm from the incisal edge simulating edge-to-edge occlusion;
- Condition 2: 3 mm from the incisal edge simulating centric occlusion;
- Condition 3: 6 mm from the incisal edge simulating a deep overbite.

Results

Maximum Principal Stress and Displacement of the C-FPD

Figure 9 shows the labial view of maximum principal stress and Figure 10 shows displacement values of the C-FPD model under the three different loading conditions.

Under condition 1, maximum principal stresses exceeding 100 MPa were distributed to the incisal and gingival embrasures of the mesiodistal connectors, peaking at a value of 240 MPa in the incisal embrasure of the mesial connector. Displacement exceeding 0.026 mm was observed over the entire surface of pontic. Displacement, however, was gradually decreased from the incisal edge to the cervical side of both abutment teeth. Maximum principal stress distribution under condition 1 revealed that the load applied to the cutting edge had caused a bending deformation toward the labial side, which further induced the twisting force toward the labial side rotating around the lower embrasure of the mesiodistal connectors. High stress concentration around the upper embrasure of the connector was mainly generated by two factors: the large displacement of the pontic and the combined effects of both the bending and twisting forces³⁹⁻⁴¹).

Under condition 2, maximum principal stresses exceeding 50 to 70 MPa were distributed across the gingival embrasure of the mesial connector, peaking at a value of 56.0 MPa. Displacement exceeding 0.020 to 0.026 mm was observed at the incisal edge of the pontic. Displacement, however, was gradually decreased from the incisal edge to the cervical side of both abutment teeth. Under condition 2, a load applied at the center of the lingual surface (3 mm from the incisal edge) induced less twisting force against the FPD, resulting in small displacement and lower stress values.

Under condition 3, maximum principal stresses exceeding 100 MPa were distributed to the gingival embrasure of the mesiodistal connectors, peaking at a value of 189 MPa in the gingival embrasure of the mesial connector. Displacement exceeding 0.026 mm was observed on the pontic base. Displacement, however, was gradually decreased toward the incisal edge. Under condition 3, a large displacement was observed around the pontic base, which was mainly induced by the turning behavior of the pontic when it rotated around the upper embrasure of the mesiodistal connectors. High stress concentrations around the lower embrasure of the connector might be induced by the twisting behavior with the connector as a fixed point. Concerning displacement of both abutment teeth, displacement was found to occur only on the side adjacent to the edentulous space.

In the C-FPD, a localized high stress concentration was observed in the connector area under all loading conditions, which might be attributable to the isotropic properties of the hybrid composite resin and the intrinsic morphology of the anterior FPD. To meet esthetic and biomechanical requirements, an anterior FPD inevitably possesses irregular and high stress concentrations around the area of the connector.

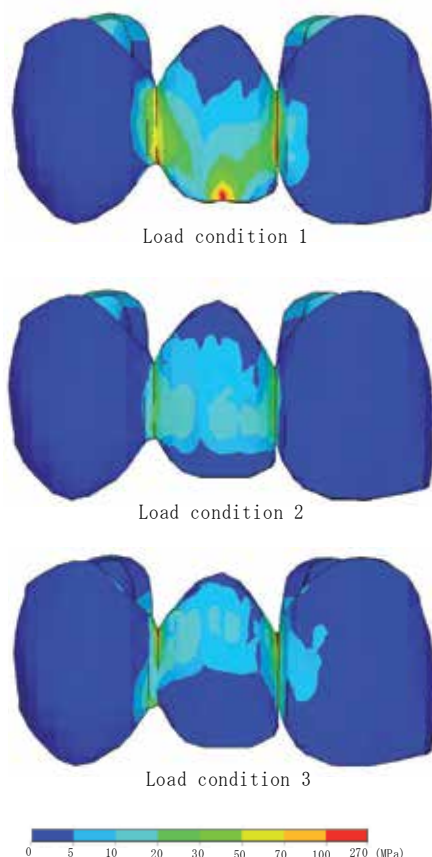


Fig. 9. Maximum principal stress distributions of FPDs. (Labial view)

The highest stress distribution was admitted in load condition 1, and the lowest stress distribution was admitted in load condition 3

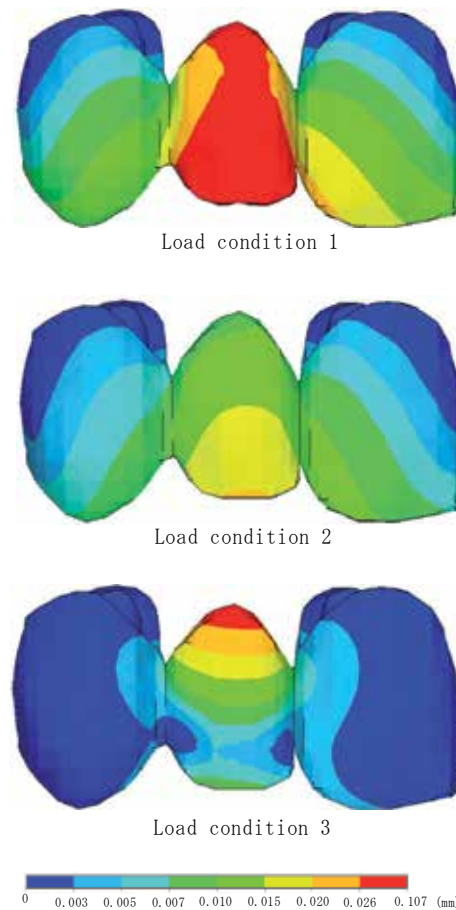


Fig. 10. Displacement distributions of FPDs(Labial view)

The biggest displacement distribution was admitted in load condition 1, and the smallest displacement distribution was admitted in load condition 3.

Vector Indication of Principal Stress on the C-FPD

Figure 11 shows a vector indication of principal stress and an outline of principal stress direction of the C-FPD. In general, the maximum principal stress was oriented from the lingual side to the connector at the abutment tooth under all loading conditions. At the connector, the maximum principal stress was oriented from the marginal ridge of the abutment tooth (missing tooth side) to the labial side of the pontic, showing a curvature along the external form of the labial side of the pontic. At the lingual side of the pontic, the maximum principal stress was directed parallel to the external form of the pontic. Maximum principal stress direction patterns from the abutment tooth to the connector, due to the shear force generated from the lingual marginal ridge to the labial side of the pontic, showed a curvature along the lingual side of the abutment tooth and the labial side of the pontic. These FE test results revealed that critical factors such as FPD morphology, loading points, and loading direction greatly affect the stress direction pattern⁴²⁻⁴⁴.

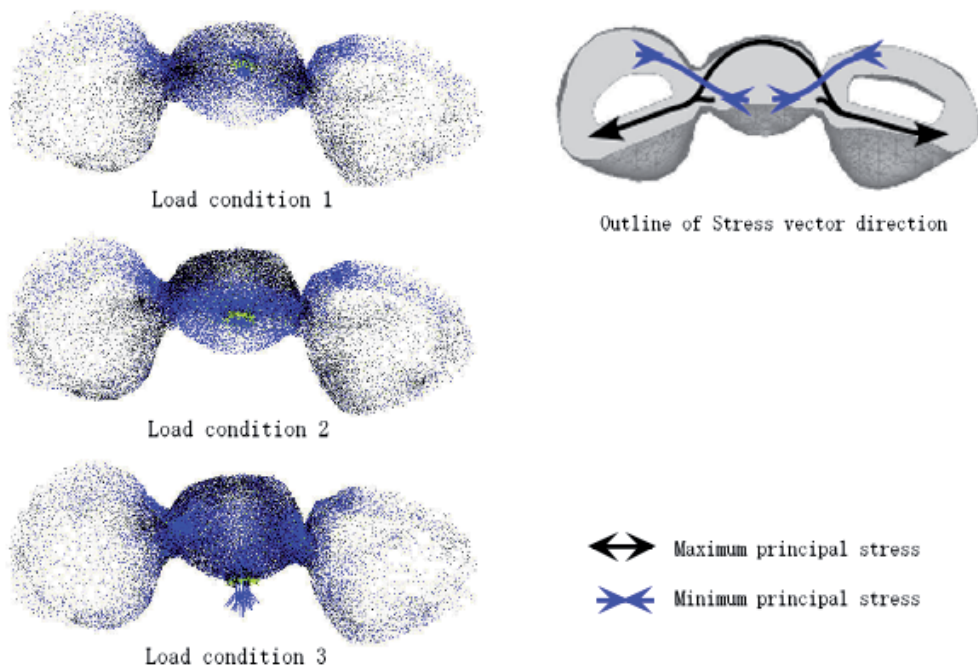


Fig. 11. Vector indication of principal stress. Horizontal cross section of the C-FPD.

Reduction Ratio of Maximum Principal Stress on the FRC-FPD Models

Stress reduction ratio was calculated based on the difference between the maximum principal stress values of the mesial connector of the C-FPD, which detected the highest values under each loading condition (condition 1: incisal embrasure, condition 2: gingival embrasure, condition 3: gingival embrasure), and the corresponding values of the FRC-FPD models. Table 2 represents the stress reduction ratio of each model with their varying loading conditions. Stress reduction ratios of the FRC-FPD models in descending order were FRC1, FRC2, and FRC3. The highest stress reduction ratio of 36% was obtained in the combination with loading condition 2 and the FRC1 framework.

	Load condition 1		Load condition 2		Load condition 3	
	Stress	reduction	Stress	reduction	Stress	reduction
FPDs	239	/	56	/	188	/
GFR1	182	24	36	36	132	30
GFR2	189	21	37	34	135	28
GFR3	197	18	44	21	159	15

Table 2. Stress reduction (%) and Maximum principal stress (MPa) at Mesial conector

Comparison between C-FPD and FRC1 Models

Figure 12 represents the labiolingual vertical crosssectional view of the mesial connector (recorded highest maximum principal stress value) of the C-FPD and FRC1 framework

models (observed highest stress reduction ratio). Under condition 1, the mesial connector was viewed from the incisal edge, while under conditions 2 and 3, the mesial connector was viewed from the cervical edge. Under condition 1, a high-stress area of more than 100 MPa, which had been observed in the C-FPD, was concentrated at the lowest part of the mesial connector in the FRC1. Furthermore, a 50 to 70 MPa stress area was also reduced compared to the C-FPD. Under condition 2, a 50- to 70-MPa stress area, which had been observed in the C-FPD, was not seen in the FRC1. Moreover, a 30 to 50 MPa stress area was only observed at the lowest part of the mesial connector. Under condition 3, a localized high stress area of more than 100 MPa was only observed at the pontic side of the connector in the FRC1 and the 70 to 100 MPa stress area was decreased when compared with the C-FPD.

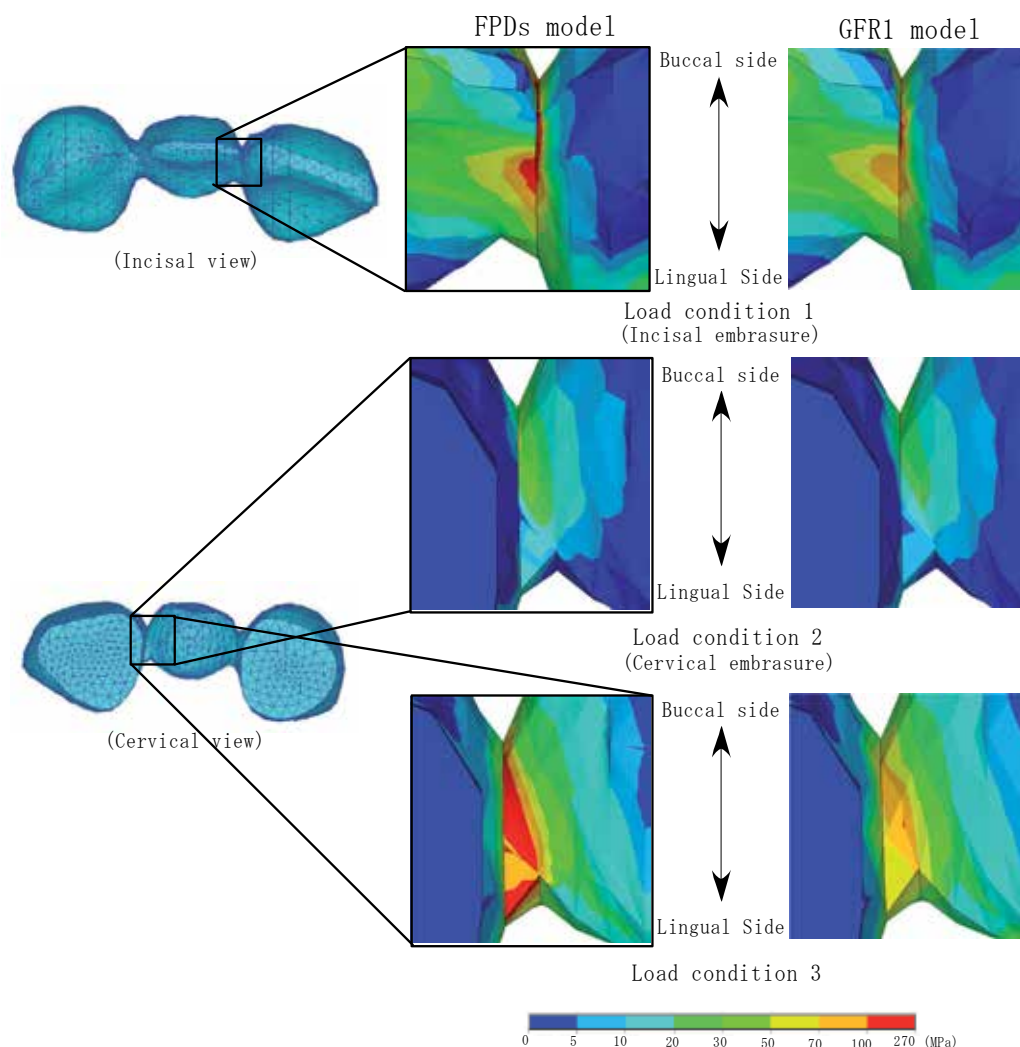


Fig. 12. Maximum principal stress at the mesial connector. A = C-FPD; B = FRC1.

Figure 13 represents the horizontal cross-sectional view of the stress distribution pattern generated by the three different loading conditions in the middle of the connectors of both the C-FPD and FRC1 frameworks. The C-FPD induced a high stress concentration in the outer surface of the connector, whereas the FRC1 induced the same stress in the fiber framework. The maximum principal stress generated in the resin matrix from the connector to the pontic in the C-FPD was relieved and transferred to the fiber framework in the FRC1. On the whole, the maximum principal stress direction pattern of the FRC-FPD models showed a similar tendency to that of the C-FPD. Once the stress direction pattern in the fiber framework of the FRC-FPD models was evaluated, the maximum principal stress was oriented uniformly along the fibers, representing a striking contrast to the C-FPD. Furthermore, the principal stress was oriented randomly at the mesiodistal connectors. This phenomenon was especially obvious in the case of the FRC1 framework.

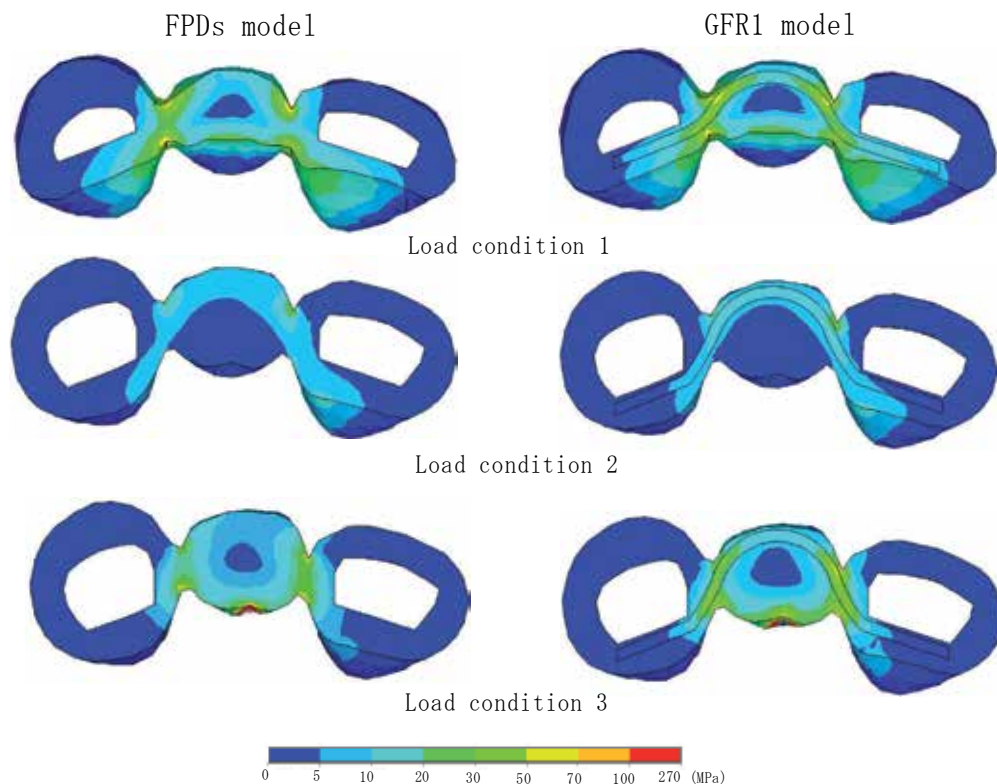


Fig. 13. Maximum principal stress at the horizontal cross section. A = C-FPD; B = FRC1.

FRC-FPD Frameworks

The stress reduction ratio of the maximum principal stress obtained from the differences between the C-FPD and FRC-FPD models (FRC1, FRC2, and FRC3) exhibited that the FRC1 model obtained the highest stress reduction ratio under all loading conditions. It was suggested that the fiber framework, which had been placed within a high stress distribution area of the labial side, effectively bore more tensile stress.

Concerning FRCC2 and FRC3, inappropriate positioning of the fiber framework placement only affected the stress distribution pattern and not the stress reduction ratio, indicating a small reduction ratio. When investigating the influence of the fiber framework on displacement, it was reduced only at the pontic of the FRC1. When evaluating the displacement distribution pattern of the abutment teeth and connectors, no significant differences were found among the C-FPD and models reinforced with a fiber framework. As for the direction of maximum principal stress on the whole, no significant differences were found among the C-FPD and models reinforced with a fiber framework. At the fiber framework, however, maximum principal stress tended to be directed along the orientation of the fiber and the direction of the fiber framework and maximum principal stress partially coincided with one another. This demonstrated a striking contrast to the C-FPD model, of which the principal stress was oriented randomly at the mesiodistal connectors.

Optimum Design of the Fiber Framework

For prostheses to function successfully for a long time in a rigorous oral environment, it is highly critical to make the most of the anisotropic nature of the fiber framework (ie, having an exceptionally high Young's and elastic modulus along the orientation of the fibers). Continuous unidirectional glass fiber is used as the reinforcement framework in an FPD. Obtaining the maximum reinforcement effects of the fiber framework and avoiding high irregular stress concentration are very important when designing an FPD. Esthetic treatment for the anterior tooth region often requires utilization of materials with varying color shades. Extensive space is needed to place different combinations of composite shades, especially in the incisal edge region of anterior teeth. In order to achieve good esthetics when working with anterior teeth, a labial space of 0.5 mm is considered to be the minimum for layered veneering composites with different color shades. FRC materials are practical in achieving the natural tooth color with both the layering as well as with the staining technique. Fiber-reinforced anterior FPDs were investigated from a biomechanical and structural point of view and it was found that the optimal fiber reinforcement effect is achieved with a curved FRC extending from the lingual side of both abutment teeth to the labial side of the pontic base.

Conclusions

1. Localized high stress concentration was observed around the connectors under all loading conditions.
2. In all FRC-FPD models, the fiber framework bore the stress generated, showing the stress-bearing capacity of the FRC framework. The highest stress reduction ratio was obtained with a curved FRC extending from the lingual side of both abutment teeth to the labial side of the pontic base.
3. Fiber reinforcement also enabled the reduction of the quantified displacement.

3. Mechanical approach for designing of a posterior FRC FPD

By means of finite element analysis, the optimal design of fiber framework placed in a posterior FRC FPD replacing the mandibular first molar was obtained (Figure 14). FRC framework was designed with three different variations each in the pontic and connector

(total nine different models). The results were compared with the control model (hybrid composite FPD without fiber reinforcement) and FRC FPD.

This term consists of

- (1) A main framework design of FRC FPD against vertical occlusal force using 2D FEM analysis (1st analysis).
- (2) The optimal thickness of a main framework for FRC FPD using 2D FEM analysis (2nd analysis).
- (3) Focused on the possible need of additional FRC support for the pontic. Two types of loading conditions, i.e. against vertical occlusal load, and loading the cusp in lateral movement of jaw, on mechanical behavior of the pontic of hybrid composite FPD (C-FPD) and FRC FPD (FRC-FPD) were studied by 3D FEM analysis (3rd analysis).



Fig. 14. FRC-FPD and FE model in the posterior region.

1st analysis: Introduction

Hybrid composite is presumed to have the isotropic material properties. The glass-fibers On the other hand, is different from that. The characteristics of fiber are, first it tensile stress is strong in the direction parallel to the fibers. Second, it is weak not only the tensile stress in the direction perpendicular to the fibers but also compression. So we assume that the fiber will have to deal with the orthogonal anisotropic material properties on the stress analysis carried out.

The aim was to study the effects of differences in the location of FRC in internal bridge structure on reinforcing effects. This was done by using stress analysis according to the FEM by designating FRCC as an anisotropic material for the purpose of clinical application of FRC hybrid composite FPD.

Materials and Methods

Definition of structures

With a view to replacing the mandibular first molar, two-dimensional finite element models of three-unit bridge from mandibular second premolar to mandibular second molar were created. This model (FRC0) was the base model with no fiber framework reinforcement. FRC0 was to be in the middle of both buccolingual and mesiodistal distances and along the long axis of a crown. Crown contour based on anatomical measurements was obtained.

Sanitary pontic design was used for missing mandibular first molar. A 1.5-mm occlusal clearance from pit and fissure and a deep chamfer (1.0 mm) were prepared for assumed jacket crown. Axial inclinations of abutment teeth were based on an anatomical study, while thickness and width of connectors were as per described in previous reports.

The medial and distal connectors and center of the pontic were selected as the FRC installation locations. The vertical distance from the apex coordinates on the bottom surface of the connectors was designated as A for the connector embedded site. For the center of the pontic, the vertical distance from the lowest point of the pontic was designated as B. These distances were then set at 0.1 mm, 0.5 mm and 1.0 mm for A, and at 0.1 mm, 1.0 mm and 2.0 mm for B (Figure 15).

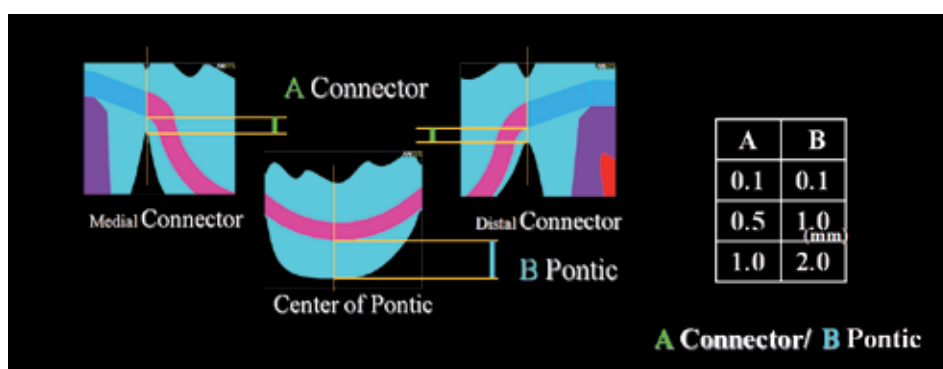


Fig. 15. The position of the fiber framework

Nine types of FRC three-unit FPD two-dimensional finite element models were then prepared by combining these distances A and B (Figure 16).. In addition, the vertical thickness of the FRC was set at 1.0 mm. Reinforcing effects were evaluated by comparing the maximum principal stress on the pontic side of the connectors with that of the base model followed by comparison of their reduction rates.

Material properties, load case and data processing

Load conditions consisted of subjecting the pontic to a bimodal concentrated load. The amount of the load was defined so that the principal stress of the results of three-dimensional model analysis coincided with the results for the base model, assuming the maximum individual tooth occlusal force of healthy permanent teeth. Restraining conditions consisted of completely restraining the bottom surface of the model. A two-dimensional finite element model that consisted of 15,046 nodes and 5,684 elements (Figure 17). Material properties used for this finite element analysis were listed in Table 1. Most of these values were determined according to a literature survey³⁴⁻³⁸). Hybrid composite, which was Estenia (Kraey medical. Tokyo, Japan.) in this study, was assumed to be an isotropic material, while glass fiber was assumed to be an anisotropic material.

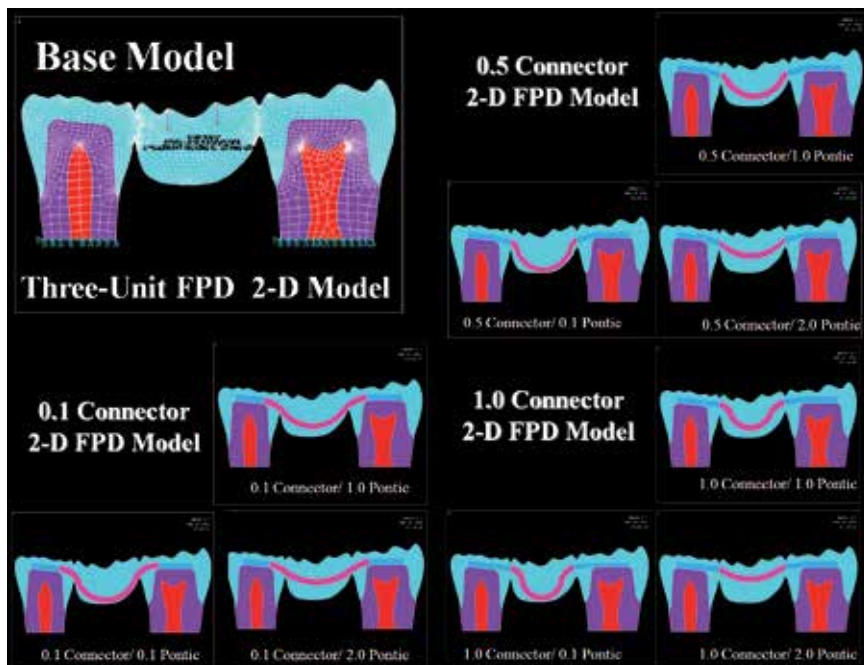


Fig. 16. Nine different 2-D FE FRC hybrid composite FPD models.

As in all finite element models, fiber orientation was set to the x-axis. As such, material properties of fiber framework in the x-axis were set to be of the highest values.

Analysis was presumed to be linear static. FE model construction and FE analysis were then performed on a PC workstation (Precision Work Station 530, Dell Computer) using a commercially available FEM software, ANSYS 5.7 (ANSYS Inc., USA).

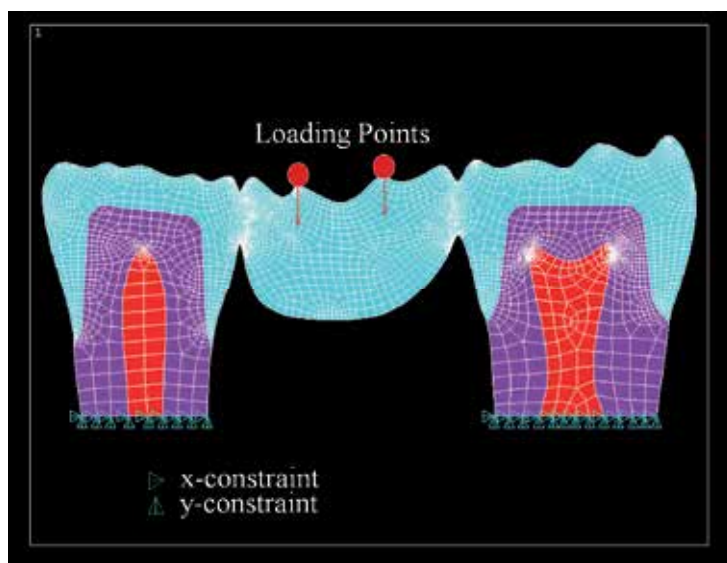


Fig. 17. Boundary conditions in 1st analysis.

Results

Stress concentration in the base model was observed on the pontic side of the medial connector, demonstrating a value of 106.82 MPa. The maximum stress reduction rate was observed in the 0.1 mm connector model and 0.1 mm pontic model, demonstrating a value of 29%, while the maximum principal stress was 75.45 MPa. A trend was observed in which the shorter the designated distance was, the greater the stress reduction rate became. A trend was also observed in which the more parallel the direction of the fibers was to the maximum principal stress, the greater the stress reduction rate. (Figure. 18-22. Table 3)

Conclusion

Contrary to my expectations, the result of the investigation is as follows the stress decrement at model (0.1 connector/1.0 and 2.0 pontic) is smaller than the others. The cause of this result is that it was not identical direction of tensile stress and fibers. So, these models could not show the stress decrement. In the preparation of FRC hybrid composite FPD, it is important that the FRC is embedded at sites at which there is occurrence of tensile stress, and that the fibers are embedded in a form so as to follow the direction of tensile stress.

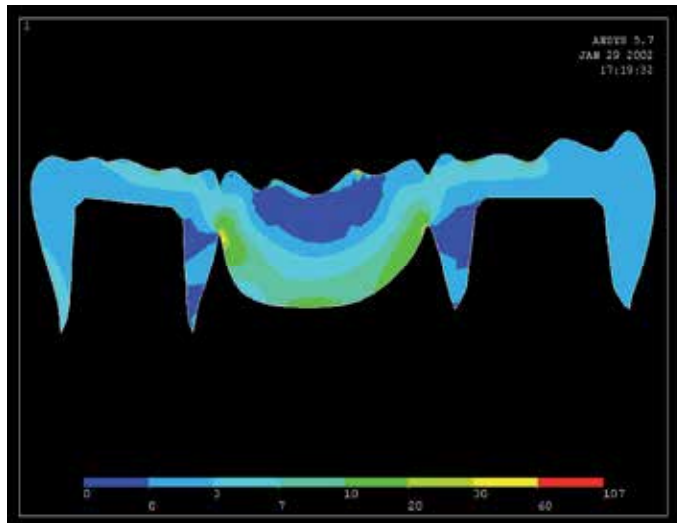


Fig. 18. Maximum principal stress on 3 Unit-CFPD 2-D Model

	0.1 Connector	0.5 Connector	1.0 Connector
0.1 Pontic	29 (75.45)	26 (78.36)	25 (80.57)
1.0 Pontic	19 (86.93)	24 (81.56)	21 (84.75)
2.0 Pontic	18 (88.12)	20 (85.13)	16 (89.69)

Table 3. Stress reduction ratio on 1st analysis. %(MPa)

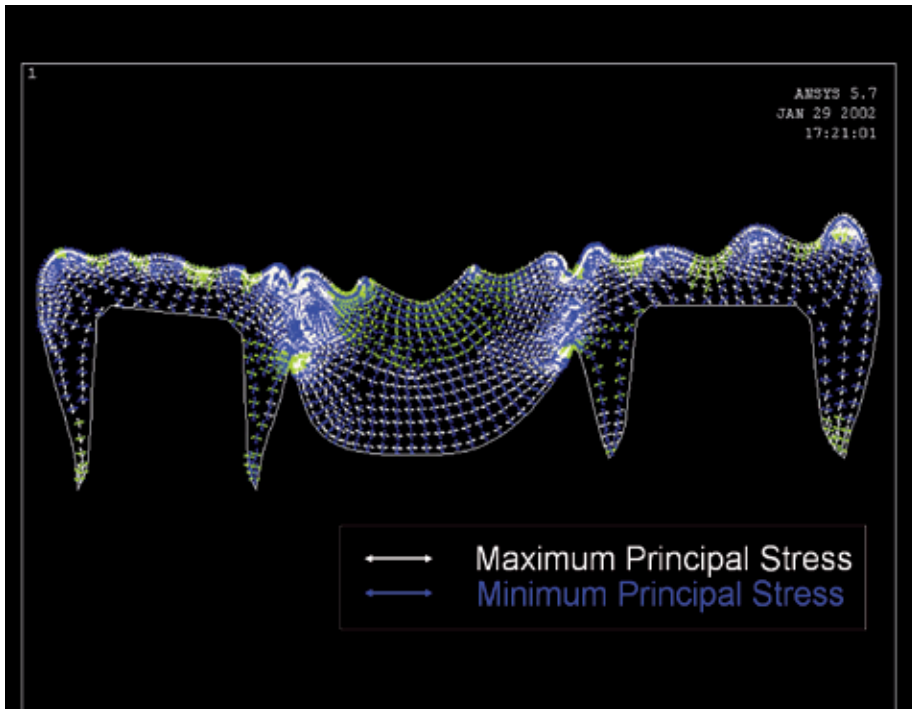


Fig. 19. Vector Indication of Stress on 3 Unit-CFPD 2-D Model

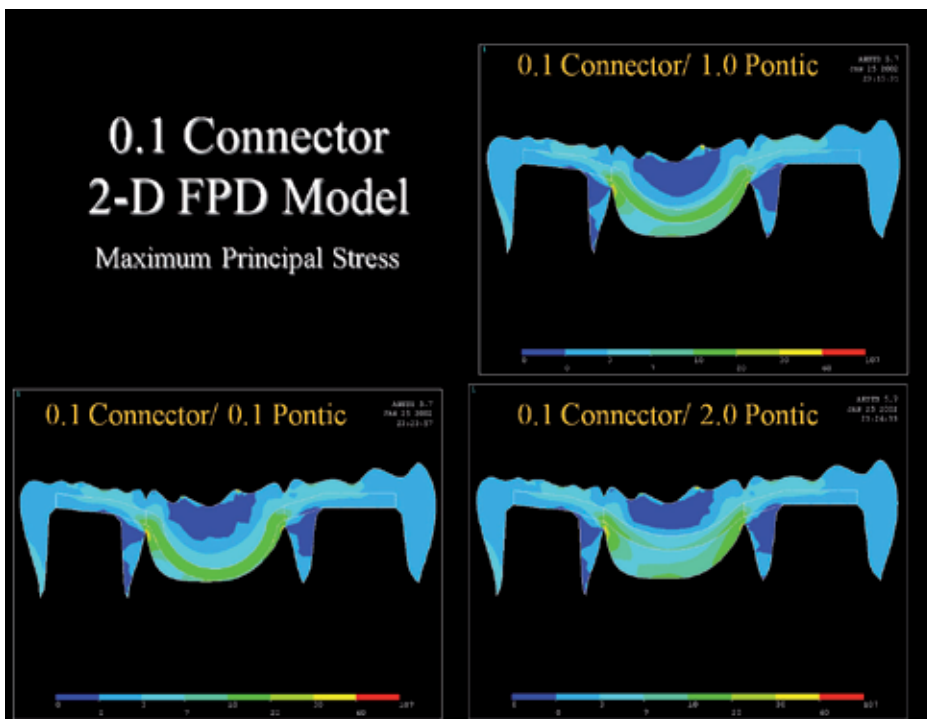


Fig. 20. Maximum principal stress on 3 Unit FRC FPD 2-D Model(Connector 0.1mm)

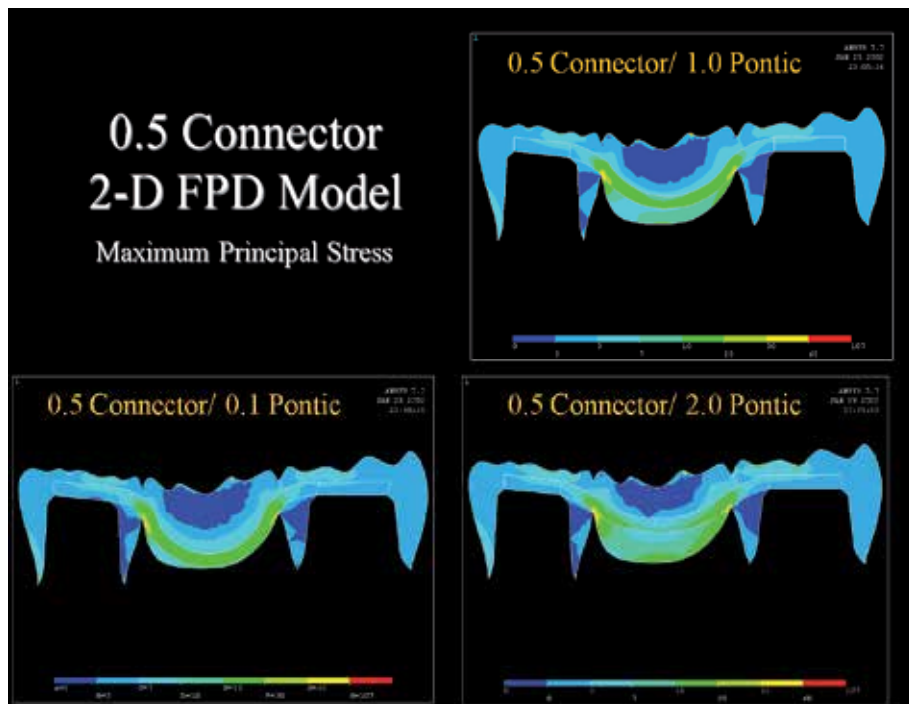


Fig. 21. Maximum principal stress on 3 Unit FRC FPD 2-D Model (Connector 0.5mm)

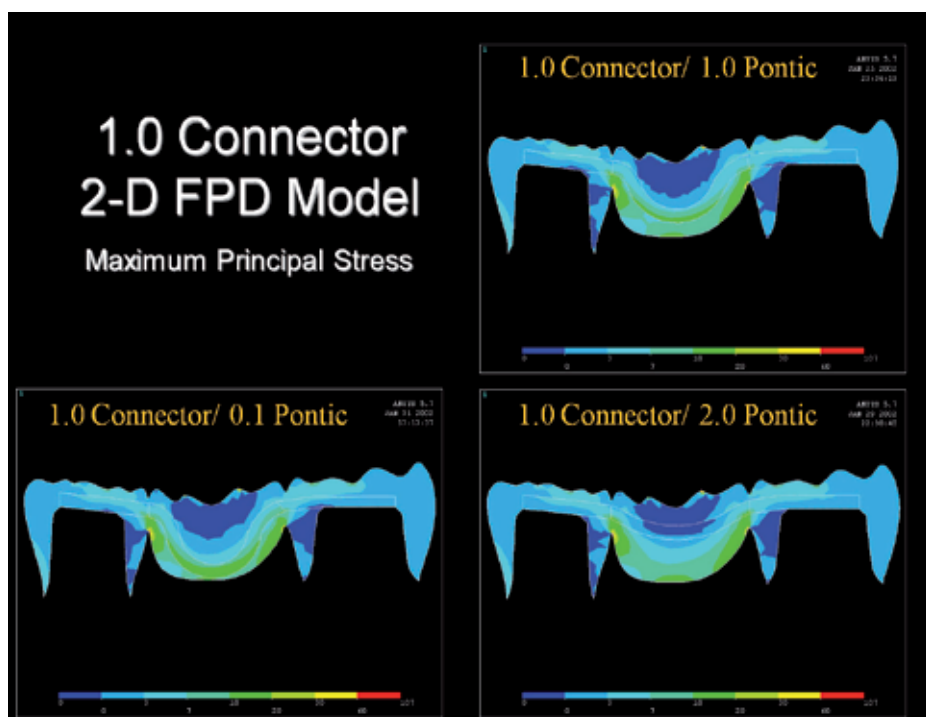


Fig. 22. Maximum principal stress on 3 Unit FRC FPD 2-D Model (Connector 1.0mm)

2nd analysis: Introduction

1st analysis demonstrated that maximum principal stress applied to mandibular first amolar was decreased from 106 MPa to 75 MPa (approximately 29% stress reduction) by maximizing fiber reinforcement effect. However, the effect of fiber quantity (i.e., vertical thickness of fiber framework) on magnitude and distribution of stress has not yet been fully investigated, which means there are no definite guidelines on this matter. Putting aside the lack of research and lack of guidelines, it is undeniably critical to clarify stress reduction mechanism (or stress-transferring mechanism), which is achieved by reinforcing with fiber framework placed within a limited occlusal clearance.

2nd analysis, preimpregnated, unidirectional, and anisotropic fiber-reinforced composite with high strength and stiffness in one direction was used to fabricate FPD. Then, from a structural and mechanical perspective, a finite element method was used to investigate the influence of the thickness of fiber framework on magnitude and distribution of stress, and thereby determine the optimal thickness of fiber framework.

Materials and Methods

Definition of structures

The base model using on 2nd analysis is same as 1st analysis. This model (FRC0) was the base model with no fiber framework reinforcement.

To optimize reinforcement effect, the shape of fiber framework was as per that described in 1st analysis. Fiber framework (figure 23) was reinforced in the axial direction (or along y-axis). With due consideration for 1.5-mm occlusal clearance, a total of eight fiber framework models were constructed with a thickness of 0, 0.2, 0.4, 0.6, 0.8, 1.0, 1.2, or 1.4 mm (FRC0, FRC0.2, FRC0.4, FRC0.6, FRC0.8, FRC1.0, FRC1.2, FRC1.4) (figure 24).

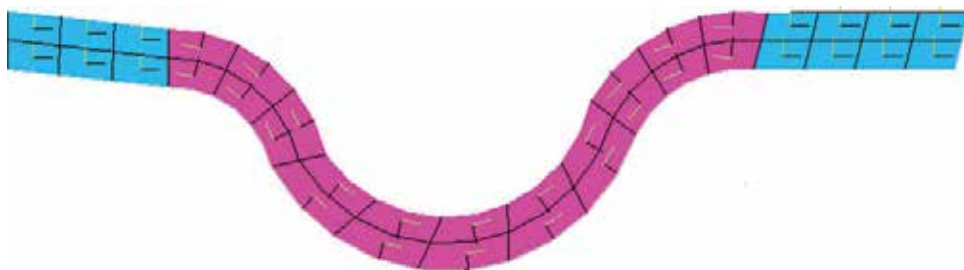


Fig. 23. A finite element model of glass fiber framework

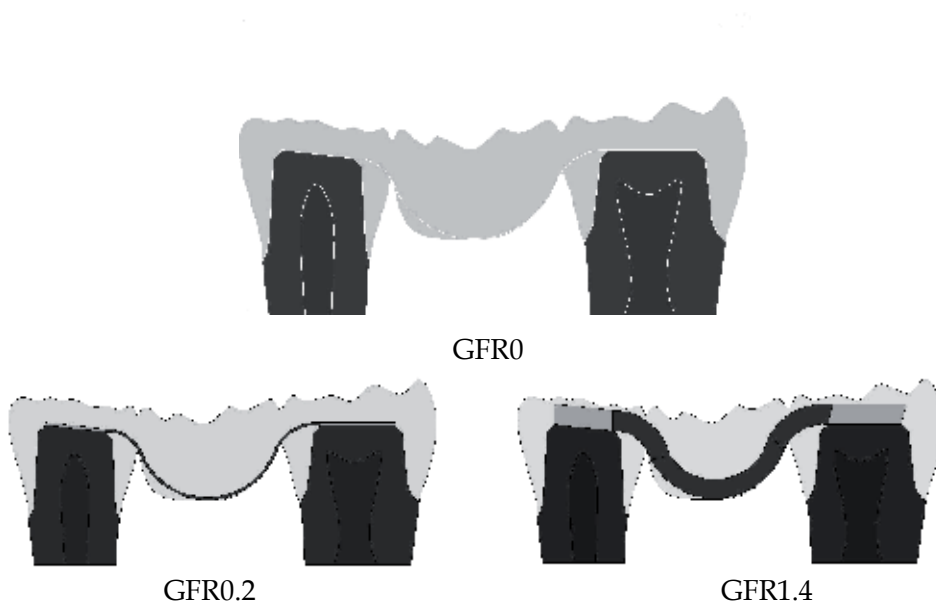


Fig. 24. Two-dimensional models of 3-unite glass fiber reinforced fixed partial dentures

Material properties, load case and data processing

Finite element models consisting of 5,684 elements with 15,046 nodes were created for plane stress analysis. Quadrilateral elements with eight nodes were chosen because they are more regular than triangular elements. The geometric conditions and materials properties are same as 1st analysis. Figure 25 shows the outline of the model and its boundary conditions.

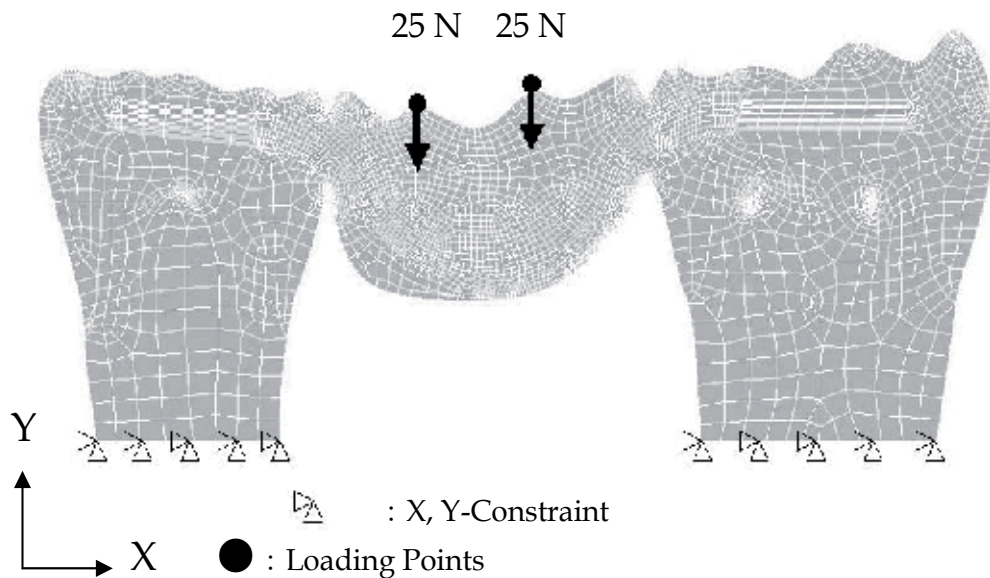


Fig. 25. Boundary conditions of finite element models (GFR1.0)

Results

Maximum principal stress distribution

Figure 26 shows the maximum principal stress distribution patterns from FRC0 to FRC1.4. With FRC0, maximum principal stresses of 3–7 MPa were observed in the mesial margin of second premolar, occlusal surface of abutment teeth, connector areas, one half of pontic base, and loading points. Maximum principal stresses of 7–10 MPa were distributed to gingival connector, pontic base, as well as pit and fissure of both abutment teeth. Maximum principal stresses exceeding 20 MPa were distributed to gingival connector, and a localized high stress of 107 MPa was observed at lower embrasure on pontic of mesial connector.

In terms of the influence of fiber reinforcement, significant differences were noted in the maximum principal stress patterns of FRC0.2–FRC1.4. With FRC0.2, maximum principal stress (103 MPa) was distributed to the fiber framework. However, maximum principal stress mostly concentrated in a limited area of fiber framework, not showing a uniform stress distribution pattern as compared with FRC0.6–FRC1.4.

Disparity in maximum principal stress was apparently observed at the interface between fiber framework and resin matrix. However, more disconcertingly, high stress concentration (84 MPa) was observed in regions pertaining to lower embrasure of the connectors, indicating that fiber framework bore little load. It should be re-emphasized that fiber reinforcement was deemed effective only if the stress were transferred from resin matrix to the fiber framework.

Concerning stress distribution pattern of pontic in all the finite element models, maximum principal stress was extensively distributed to the pontic. With increase in the fiber framework quantity (i.e., vertical reinforcement), maximum principal stress – which had been concentrated in connector area – was transferred to the fiber framework, whereby maximum principal stresses of 20–30 MPa were gradually distributed to the whole fiber framework.

Principal stress direction

Figure 27 shows the directions of maximum principal stress. With FRC0, maximum principal stress was oriented from the lower embrasure of connector to the occlusal surface of abutment teeth, and then to the marginal area parallel to the external form of bridge. At pontic, maximum principal stress was oriented from the upper embrasure of connector (or marginal ridge of pontic) to the pontic base, showing a curvature along the external form of pontic base. On the whole, maximum principal stress showed a M-shaped stress distribution pattern, i.e., connecting the proximal margins of both abutment teeth and being oriented parallel to the external form of bridge near the outer surface.

With FRC1.4, maximum principal stress was oriented uniformly along the orientation of fibers, presenting a striking contrast to FRC0, of which the principal stress was oriented randomly at the mesial connector.

Stress reduction with vertical reinforcement

Figure 28 represents the change in maximum principal stress value of high stress concentration area observed from second premolar distal margin to pontic base to second molar mesial margin. Peak value was detected in the lower embrasure of both mesial and

distal connectors, although the peak values differed from each other. Apart from the peak values, maximum principal stress value increased slightly at the center of pontic base.

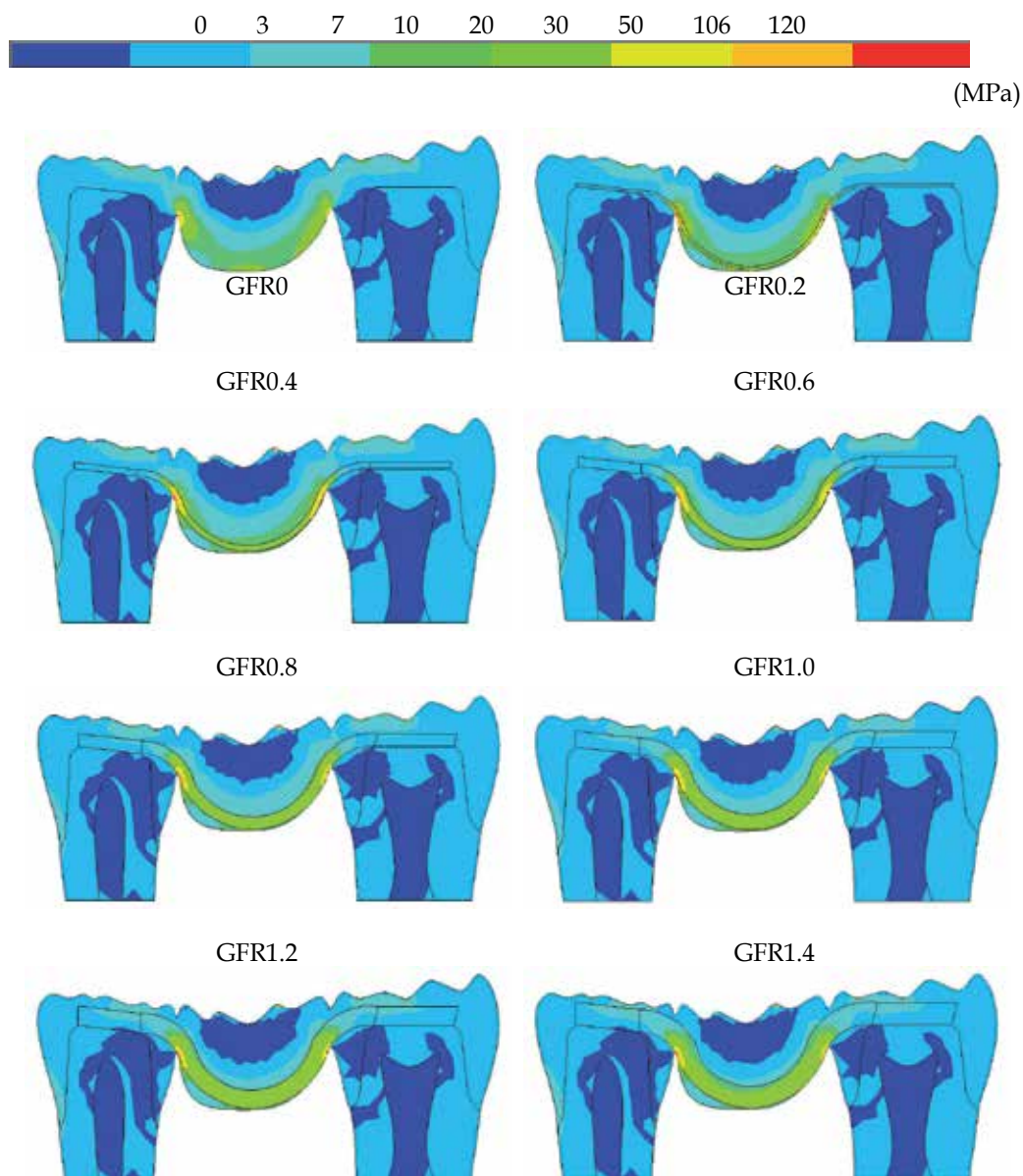
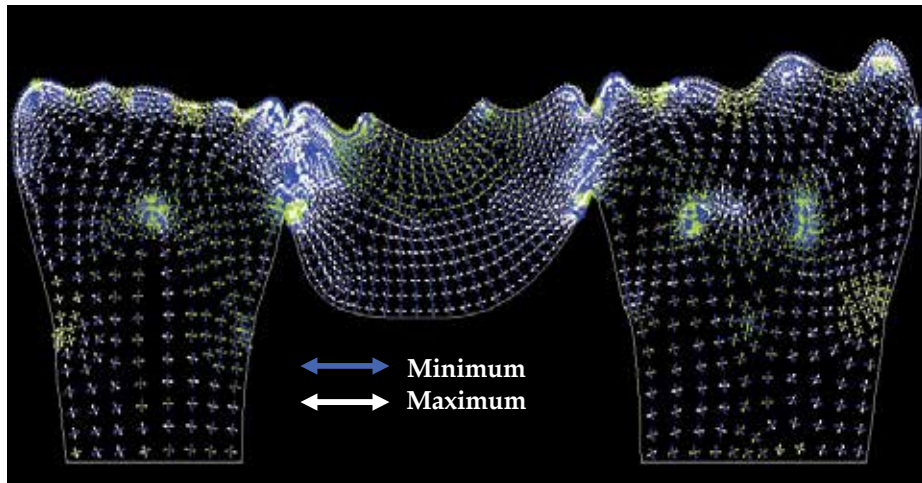


Fig. 26. Distributions of maximum principal stress

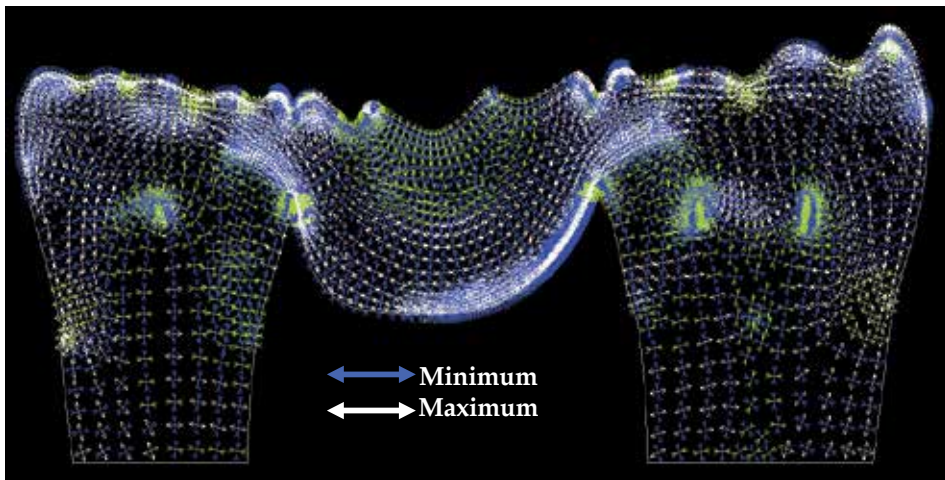
FE models reinforced with fiber framework demonstrated a sharp decline in principal stress value, compared with FE model with no fiber framework reinforcement.

Concerning FRC0.6–FRC1.4, stress reduction ratio tended to gradually decrease with increasing quantity of fiber framework. Stress reduction ratio was calculated based on the

difference between maximum principal stress values of FRC0 and FRC0.2–FRC1.4 observed at lower embrasure of mesial connector. On this note, Figure 29 shows the stress reduction ratios observed at mesial connector. Compared with FRC0 (107 MPa), FRC1.4 demonstrated the highest reduction ratio of 34% (70 MPa), while FRC0.2 demonstrated the lowest reduction ratio of 21% (84 MPa).



FPDs



GFR FPDs

Fig. 27. Depiction of stress vector

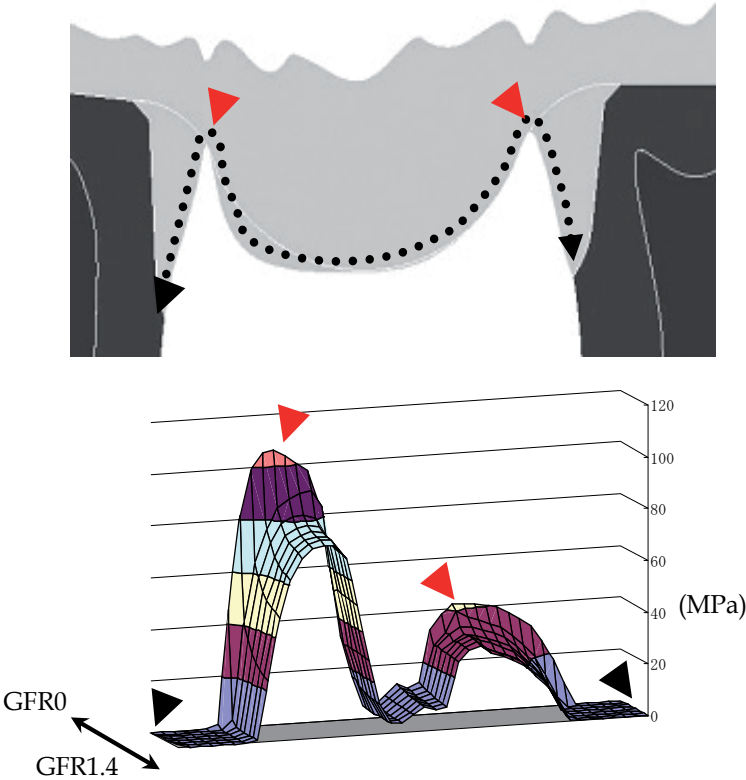


Fig. 28. Path plots of maximum principal stress on the bottom surface in all models

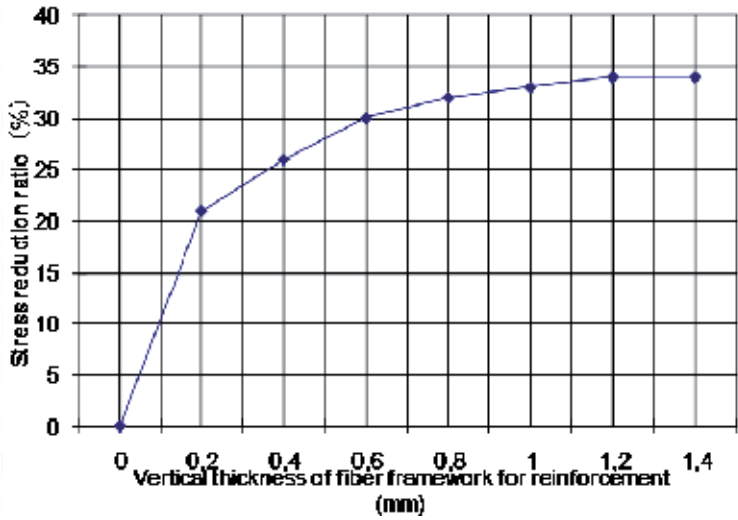


Fig. 29. Reduction ratios of maximum principal stress values observed at mesial connector

Discussion

Maximum principal stress distribution

Stress analysis with different quantities of fiber framework revealed that the value of maximum principal stress peaked at the lower embrasure of connectors. With a three-point bending test using bridge specimens, Vallittu¹⁰⁾ demonstrated that a crack which was initiated at lower embrasure of connectors propagated to the loading points. Further, it was commented that high stress concentration in connector area was closely related to bridge design — a suggestion in cordial agreement with the results of this study, which also showed high stress concentration in lower embrasure of connectors.

However, high stress concentration in connector area is mainly due to intrinsic morphology of a bridge.

As for the stress distribution pattern in pontic, maximum principal stress was efficiently transferred from resin matrix to fiber framework with increasing quantity of fiber framework. This was probably due to the excellent tensile properties of glass fiber to bear more tensile stress, thus causing less tensile stress to be generated in the resin matrix.

As for the stress distribution patterns in connectors, they showed a similar behavior as that of pontic, i.e., stress was transferred reliably from resin matrix to fiber framework with increasing quantity of fiber framework. Compared with FRC0, FRC0.2 demonstrated that stresses exceeding 78 MPa were extensively distributed along the orientation of fibers, and that maximum principal stress value at lower embrasure of connectors was reduced to 84 MPa (approximately 21% stress reduction). With increasing quantity of fiber framework (FRC0.4FRC1.4), maximum principal stress was gradually transferred from resin matrix to fiber framework, such that stresses exceeding 78 MPa were no longer observed in resin matrix from FRC0.6.

With increase in the quantity of fiber framework, the value of maximum principal stress in connectors tended to decrease. This was probably due to the exceptionally high elastic modulus of the fiber framework along the orientation of fibers. As a result, the enhanced tensile properties enabled glass fibers to bear more tensile stress generated in the connectors, thus reducing the stress in resin matrix.

In this research, FRC1.4 achieved the highest stress reduction ratio, demonstrating a uniform stress distribution pattern without any high stress concentration in the fiber framework. Indeed, Figure 28 confirmed that fiber framework had enough stress-bearing capacity to withstand occlusal force generated in bridge.

Principal stress direction

Maximum principal stress was oriented from marginal ridge of both abutment teeth to pontic base, showing a curve-like stress distribution pattern along the external form of pontic base.

On the whole, maximum principal stress distribution showed a M-shaped pattern, i.e., connecting the cervical regions of both abutment teeth. With increasing reinforcement effect of fiber framework, principal stress became more systematically oriented along the major axis of fiber framework. As a result, excellent stress-bearing capacity along the orientation of fibers led to enhanced tensile strength.

The Optimum thickness of fiber framework

In clinical scenarios with a typically limited occlusal clearance, it is highly critical that reinforcement effect be optimized with a relatively small quantity of fiber framework. In this connection, it has been reported — from a structural and mechanical perspective — that fiber framework be reinforced with 0.6–1.4 mm thickness for 1.5-mm occlusal clearance. Moreover, the higher the quantity of fiber framework, the greater was the reinforcement effect.

However, when reinforcement exceeded 0.6 mm, stress-bearing capacity did not improve significantly; on the contrary, a lower stress reduction ratio was yielded. Therefore, it did not seem necessary for fiber framework thickness to be more than 0.6 mm to withstand an average adult's masticatory force in the molar region. Moreover, from esthetic and biomechanical standpoints, at least 0.5 mm of occlusal surface must be overlaid with a veneering composite with highly filled ceramic particles — otherwise, esthetics will be compromised. Putting both stress reduction findings and esthetic/biomechanical considerations together, 0.6–1.0 mm thickness seemed to be both adequate and appropriate to achieve the desired reinforcement effect, with more than one third of occlusal clearance placed with fiber framework.

Conclusions

1. Results of finite element analysis showed that for all fiber-reinforced models, the fiber framework bore the maximum principal stress. With FRC0, maximum principal stress of 107 MPa was observed at lower embrasure of connectors. With FRC1.4, on the other hand, only a maximum principal stress of 70 MPa was observed, indicating approximately 34% stress reduction.
2. In terms of quantified displacement, an area of displacement peaking at 0.02 mm was distributed extensively from occlusal surface to pontic base in FRC0, but which was reduced to a limited area of occlusal surface in FRC1.4. By increasing the quantity of fiber framework, quantified displacement at pontic was reduced, thereby improving significantly the rigidity of bridge.
3. Significant stress reduction was found in FRC0.2–FRC0.6 models, but stress reduction seemingly remained unchanged in FRC0.6–FRC1.4.

The present test results revealed that magnitude and distribution of stress were influenced by the thickness of fiber framework, whereby 0.6 mm of vertical reinforcement was deemed both necessary and adequate from a structural and mechanical perspective.

3rd analysis: Introduction

For successful function of the FRC FPD for a long period of time, it is important to clarify biomechanical behavior of posterior FPD, in terms of magnitude of stress, stress distribution pattern and displacement under functional loading. Biomechanical information related to occlusal conditions will be useful in preparing guidelines for clinical applications of FRC FPD.

3rd analysis focused to the design of the FPD, especially on the possible need of additional FRC support for the pontic. Two types of loading conditions, *i.e.* against vertical occlusal load, and loading the cusp in lateral movement of jaw, on mechanical behavior of the pontic

and connector of hybrid composite FPD and FRC FPD were studied. Stress and displacement generated within the FPD was studied by 3-D FE method.

Material and methods

Definition of structures

Figure 30 shows a three-dimensional model of three-unit FPD from mandibular second premolar to mandibular second molar replacing the mandibular first molar. To create a finite element model, a FPD replica was fabricated. The geometry of the abutment replica has been described by Wheeler. Sanitary pontic design was chosen for missing mandibular first molar. The thickness and width of connectors of the FPD were followed described in other author's report. A replica was measured with 0.75 mm intervals using 3-D coordinate measuring machine (Micro Scribe 3D, Micro Scribe Inc., San Jose, USA), captured data was plotted on 3-D CAD (Mechanical Desktop R2.0, Auto Desk Inc., California, USA) to construct 3-D solid model. The model was inputted into FE analysis software (ANSYS 10.0 Sp, ANSYS Inc., Houston, USA).

A 1.8mm occlusal clearance from pit and fissure, 1.0mm cervical clearance and shoulder margin were prepared for assumed jacket crown. The FRC framework used in this study was 0.7mm in height and 2.1mm in width, and consisted of three continuous curvatures of mesial connector, pontic and distal connector, having an each individual curvature. The FRC framework was placed as well as from mesial side of mandibular second premolar to distal side of mandibular second molar.

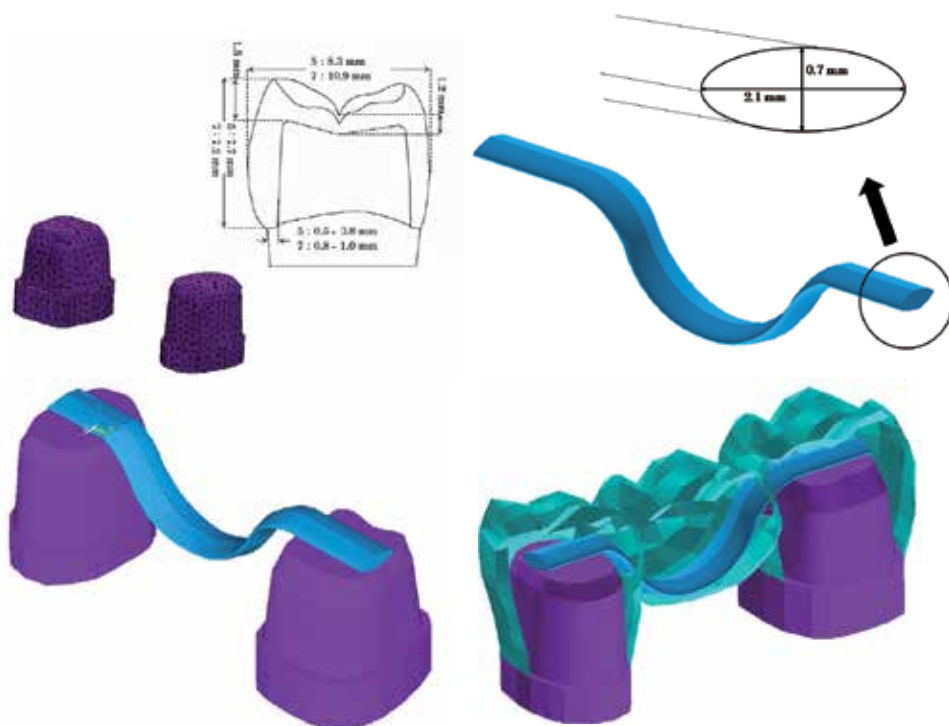


Fig. 30. Structure of the FRC-FPD used in this study

Material properties, load case and data processing

Veneering material was supposed to be hybrid composite resin (Estenia, Kuraray medical Inc., Tokyo, Japan) and it was assumed to be isotropic material. FRC framework was supposed to be unidirectional glass fiber and it was assumed to be anisotropic material. Hexagonal elements with 8 nodes were chosen for cross-sectionally anisotropic FRC framework, in the same way, hexagonal elements with 20 nodes were chosen for hybrid resin-dentin-pulp continuum, and a 3-D finite element model consisting of 70,393 elements with 155,070 nodes was created. To define FRC framework as rectangular anisotropic material, curvature of FRC framework was to be consisted of three continuous curvatures of mesial connector, pontic and distal connector, having an each individual curvature. A new central coordinate of the circle was established for each curvature. Thus, longitudinal axes of glass fibers and 3-D FRC framework were parallel. The final element in all directions of FE model base was fixed, which was defined as boundary condition.

Two loading conditions, *i.e.* against vertical occlusal force referred as vertical load, and loading the cusp in lateral movement of jaw were tested in the model (Figure 31). The latter is referred as lateral load. A total of 629N vertical load was applied to eight points on the occlusal surface of the pontic to simulate maximum occlusal force^{31,45}. To simulate loading cusp in lateral movement of jaw, a total of 250 N was applied to three points of the pontic (buccal-cusp incline) at an angle of 45 degrees from lingual direction^{46,47}. FE analysis was presumed to be linear static. Finite element model construction and FE analysis were performed on PC workstation (Precision Work Station 670, Dell Inc., Texas, USA) using FE analysis software ANSYS 10.0 Sp. In considering biomechanical properties of hybrid resin, maximum principal stress and displacement were used to demonstrate the results at certain cross-sections and on the surface of the FPD.

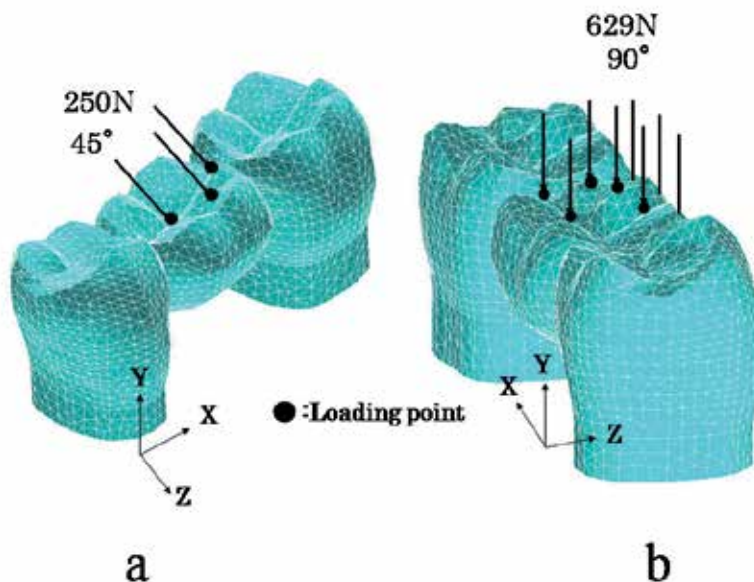


Fig. 31. Loading conditions used in this study (a: Lateral load b: Vertical load)

Results

In the study, the maximum principle stress (Pmax, MPa) and the displacement (mm) were used to assess the stress distribution of FRC-FPD. The maximum principle stress and the displacement distributions are shown in Figs. 6, 7, 8, and 9. A color scale with 8 stress values served to measure quantitatively the stress distribution in the model components. The scaling was selected not to represent the yield strength but rather to provide clear visualization of the region of stress. However in the results of maximum principle stress, the range of less than 80 MPa (yellow), from 80 to 90 MPa (orange) and more than 90 MPa (red) were focused. Because the tensile strength of hybrid composite (ESTENIA) is nearly 90 MPa.

Displacement at the pontic

Lateral load (Figure 32):

Surface and mesiodistal cross-section

C-FPD showed relatively large displacement of up to 0.04 mm at one third of the buccal side of the pontic. The largest displacement of 0.05~0.07 mm was observed from buccal groove to distobuccal cusp. Displacement of FRC-FPD was 0.03~0.04 mm being slightly less than that of C-FPD. The displacement distribution between C-FPD and FRC-FPD showed similar patterns.

Buccolingual cross-section

On lingual side, displacement of 0.02~0.03mm at C-FPD was observed at one third of the central part of the pontic. Displacement of 0.03~0.04 mm was observed from one third of central part of the pontic to buccal direction. Displacement of 0.04~0.05mm was observed from buccal one third to buccal side. The largest displacement (0.05~0.07mm) was found at the upper outermost surface of buccal side, buccal groove and its surrounding area. Compared displacement distribution of FRC-FPD with that of C-FPD, it showed a similar pattern.

Vertical load (Figure 33):

Surface and mesiodistal cross-section

The largest displacement of C-FPD (0.05~0.07mm) was seen from buccal one third of the pontic to buccal cusp of the pontic and to pontic base. Calculation of FRC-FPD showed a similar displacement distribution than that of around the connector. However, the largest displacement (0.05~0.07mm) was lower in the pontic area.

Buccolingual cross-section

C-FPD showed displacement of 0.04~0.05mm from lingual side to the middle of the pontic. The largest displacement of 0.05~0.07mm was found from the middle to buccal side of the pontic. Displacement of FRC-FPD was 0.04~0.05mm was seen through the whole pontic.

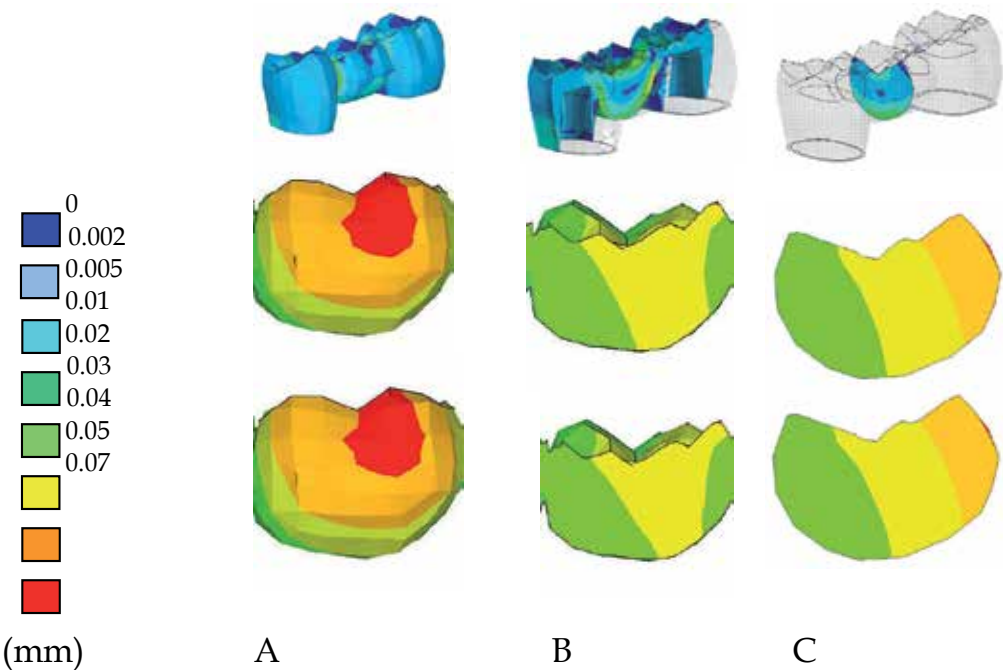


Fig. 32. Displacement of the pontic against lateral load
Upper: C-FPD. Lower: FRC-FPD.
A; Surface view of buccal side. B; Mesiodistal section. C; Buccolingual section

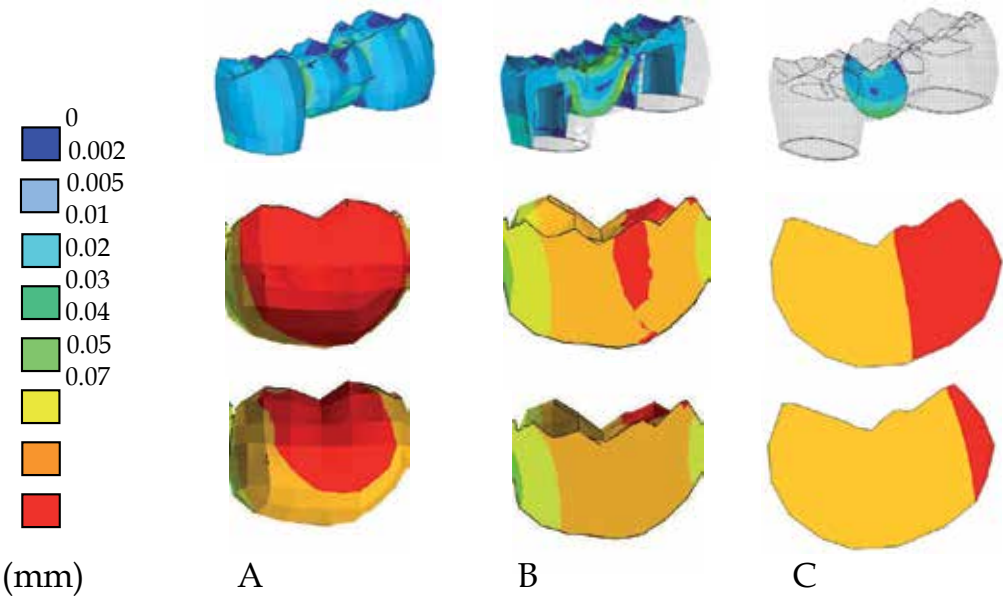


Fig. 33. Displacement of the pontic against vertical load
Upper: C-FPD. Lower: FRC-FPD
A; Surface view of buccal side. B; Mesiodistal section. C; Buccolingual section

Maximum principal stress at the pontic

Lateral load (Figure 34):

Surface and mesiodistal cross-section

In calculation of maximum principal stress of C-FPD, 70~80 MPa was distributed to areas including pontic base of distal connector and within a pontic. Maximum principal stress of the FRC-FPD, i.e. 80~90 MPa was distributed to the FRC framework of distal connector.

Buccolingual cross-section

In C-FPD, 15~30 MPa principal stresses were distributed to the area from the middle of the pontic to the buccal direction as well as to cusp and fossa on the occlusal surface.

Maximum principal stress of FRC-FPD, 15~30 MPa was distributed to the areas from one third of the pontic to the buccal direction of the pontic. Higher than 30 MPa stresses were exclusively found in FRC framework surrounded by around buccal side of pontic base. Compressive stresses were observed in the center of the pontic.

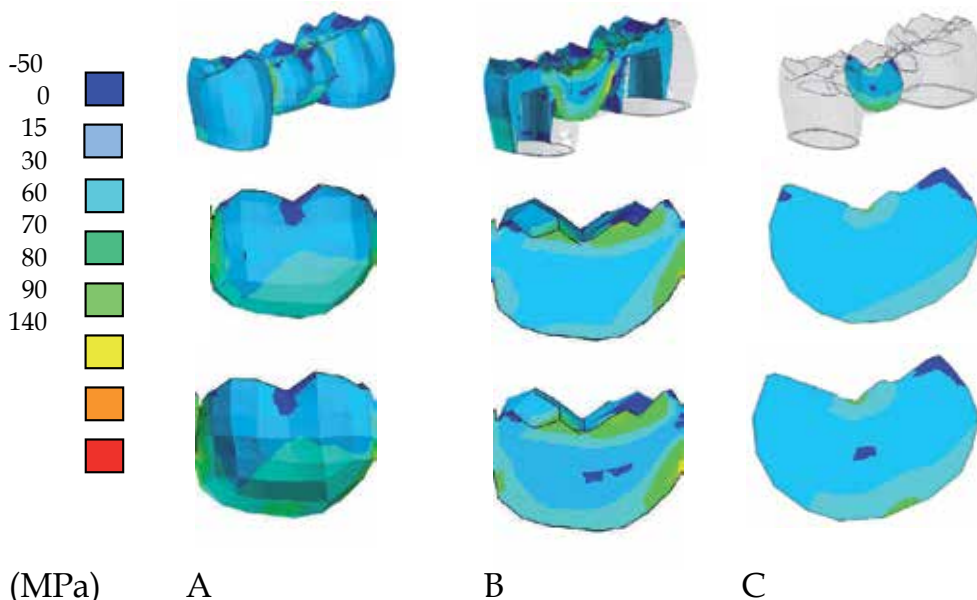


Fig. 34. Maximum principal stress of the pontic against lateral load

Upper: C-FPD. Lower: FRC-FPD

A; Surface view of buccal side. B; Mesiodistal section. C; Buccolingual section

Vertical load (Figure 35)

Surface and mesiodistal cross-section

Maximum principal stress of C-FPD, higher than 90 MPa was distributed to the area including pontic's of distal connector as well as to the inner portion of the pontic. In FRC-FPD, higher than 90 MPa principal stresses were found in the FRC framework surrounded by the pontic side of lower embrasure of distal connector, but its area was reduced compared with C-FPD. In a buccolingual section of C-FPD, 15~30 MPa principal stresses were distributed to the area from the middle to the lower one third of the pontic,

and to cusp and fossa in the occlusal surface. Maximum principal stress of FRC-FPD was 15 ~ 30 MPa in the middle to lower one third of the pontic, in the cusp and fossa, demonstrating similar stress distribution pattern of that of C-FPD. Compressive stress was observed in the middle of the pontic.

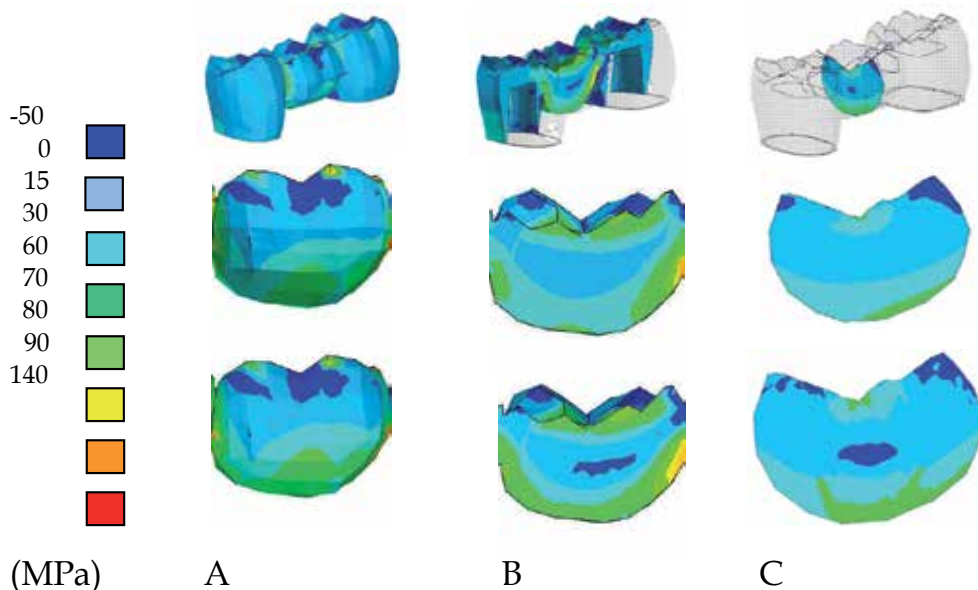


Fig. 35. Maximum principal stress of the pontic against vertical load

Upper: C-FPD. Lower: FRC-FPD

A; Surface view of buccal side. B; Mesiodistal section. C; Buccolingual section

Displacement at the connector

Lateral load:

Surface and mesiodistal cross-section Figure 36

C-FPD showed the largest displacement of 0.05-0.07 mm and was observed at the buccal groove to the distobuccal cusp around loading points. Displacement of the FRC-FPD was 0.03-0.04 mm being slightly less than that for the C-FPD. The displacement distribution between C-FPD and FRC-FPD showed similar patterns.

Buccolingual cross section of mesial connector Figure 37

Displacement for C-FPD was 0.01-0.02 mm and the displacement area was located in the middle of connector base and on lingual side. Displacement of FRC-FPD was 0.01-0.02 mm and the displacement area was found in the middle part of the connector base and the stress at the lingual side was lower compared with C-FPD.

Buccolingual cross section of distal connector Figure 37

Displacement for C-FPD was 0.03-0.04 mm and the displacement area was at the buccal side of the distal connector. Displacement distribution pattern of FRC-FPD as whole was close to that of C-FPD.

Vertical load:**Surface and mesiodistal cross-section Figure 38**

The largest displacement of C-FPD (0.05-0.07 mm) was seen from buccal one third of the pontic to buccal cusp of the pontic and to the pontic base. Calculation of the FRC-FPD showed a similar displacement distribution around the connectors.

Buccolingual cross section of mesial connector Figure 39

Displacement of C-FPD was 0.02-0.03 mm and the displacement area was located from middle to lower part of the connector, whereas other areas showed 0.03-0.04 mm displacement. A similar displacement distribution pattern was observed between the FRC-FPD and the C-FPD.

Buccolingual cross section of distal connector Figure 39

In C-FPD, the largest displacement of 0.05-0.07 mm was seen at the outermost surface of the buccal side. Displacement of FRC-FPD was 0.05-0.07 mm. Although the displacement of FRC-FPD was similar to that of C-FPD, displacement quantification showed significant reduction.

Maximum principal stress at the connector**Lateral load:****Surface and mesiodistal cross-section Figure 36**

By calculating the maximum principal stress for C-FPD, a stress of 70-80 MPa was distributed to areas including pontic the of distal connector and also inside the pontic. Maximum principal stress of the FRC-FPD, 80-90 MPa, was distributed to the FRC framework of distal connector.

Buccolingual cross section of mesial connector Figure 37

Maximum principal stresses were above 90 MPa and they were distributed to the upper line angle of buccal side of the connector. Maximum principal stress was above 90 MPa and it located in the upper buccal side of connector.

Buccolingual cross section of distal connector Figure 37

Maximum principal stress was above 90 MPa. Displacement of FRC-FPD was 0.02-0.03 mm and the displacement area was minor compared with C-FPD. Maximum principal stress was higher than 90 MPa and the stresses were distributed to the bottom of fiber framework, indicating that stress of the distal connector was transferred to fiber framework.

Vertical load:**Surface and mesiodistal cross-section Figure 38**

Maximum principal stress of C-FPD was above 90 MPa and it was distributed to the area including pontic's distal connector as well as to the inner portion of the pontic. In FRC-FPD, maximum principal stresses higher than 90 MPa were found in the FRC framework around the pontic side of lower embrasure of distal connector.

Buccolingual cross section of mesial connector Figure 39

When evaluated maximum principal stress, stresses higher than 90 MPa were distributed to the upper line angle of the buccal side of mesial connector. In addition, localized highest stress concentration of up to 120 MPa was observed in the same area (upper line angle of buccal side of mesial connector). In the maximum principal stress analysis, higher stresses than 90 MPa were observed in the outermost surface of the fiber framework of the lower connector. Comparison to the C-FPD demonstrated stresses being distributed more extensively from the lower part of the mesial connector to the buccal side of the mesial connector. Peak stress was found in the occlusal line angles and they were transferred from the hybrid composite matrix to the fiber framework.

Buccolingual cross section of distal connector Figure 39

In maximum principal stress calculation, higher than 90 MPa stresses were observed in the line angle of the lower lingual side as well as in the lingual and buccal sides of the central part of connector. When evaluated maximum principal stress, stresses higher than 90 MPa were found and they were distributed in the middle part of the distal connector in the C-FPD with a concentration to the bottom part of the fiber framework. This is suggesting stress-transferring mechanism of anisotropic fiber framework which enable distribute stress uniformly.

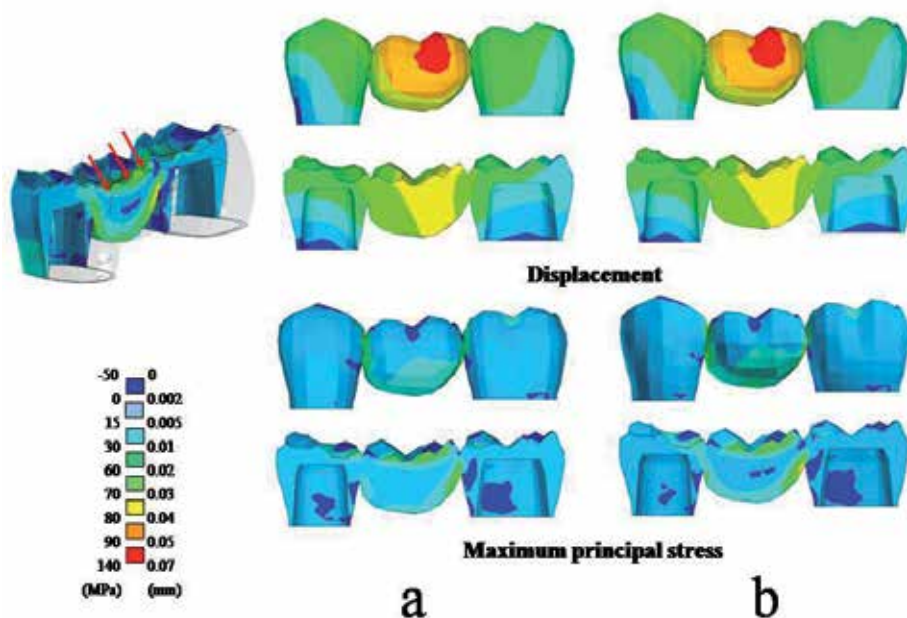


Fig. 36. Lateral load at a surface and mesiodistal cross section a: C-FPD. b: FRC-FPD

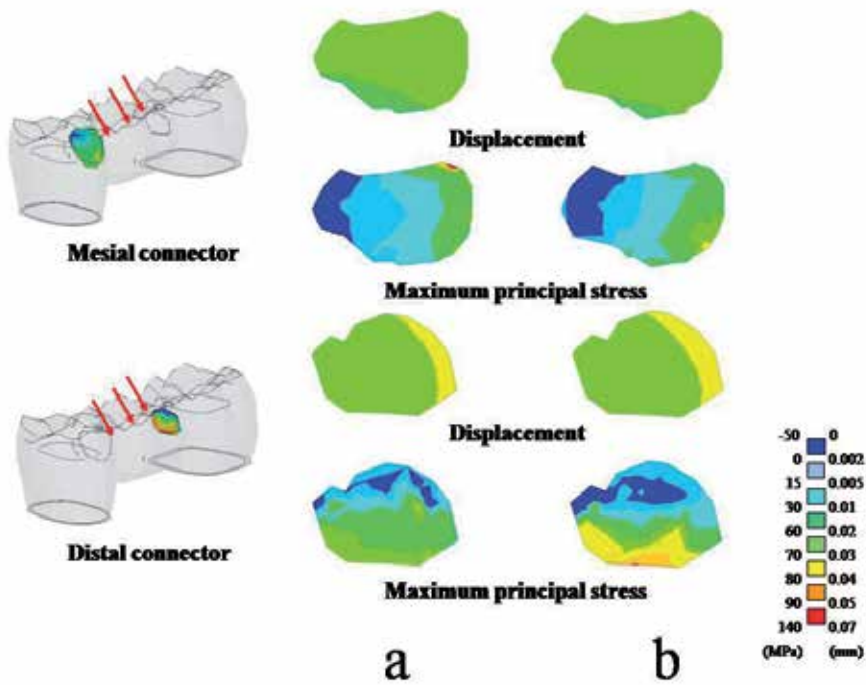


Fig. 37. Lateral load at buccolingual cross section of mesial and distal connector a: C-FPD. b: FRC-FPD

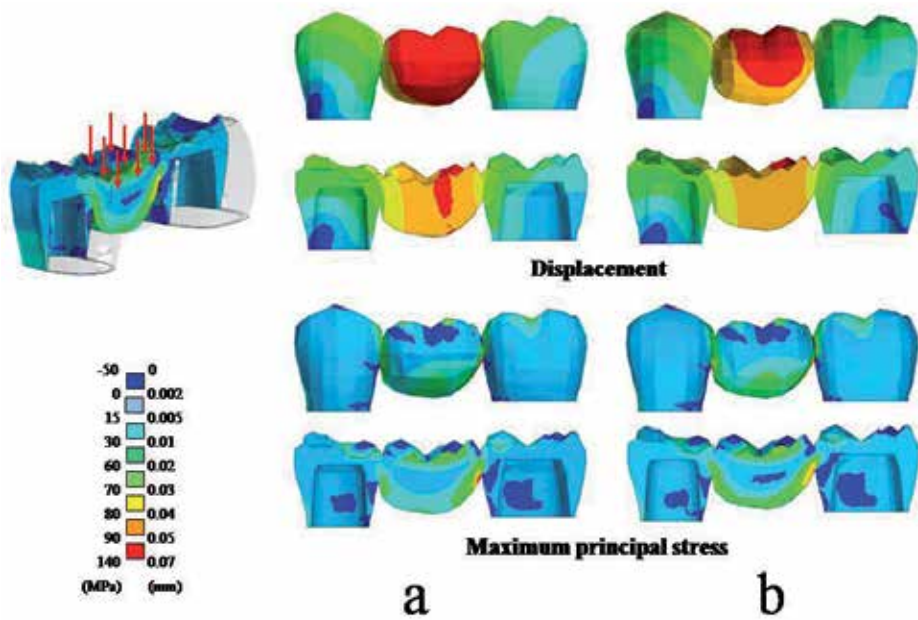


Fig. 38. Vertical load at a surface and mesiodistal cross section a: C-FPD. b: FRC-FPD

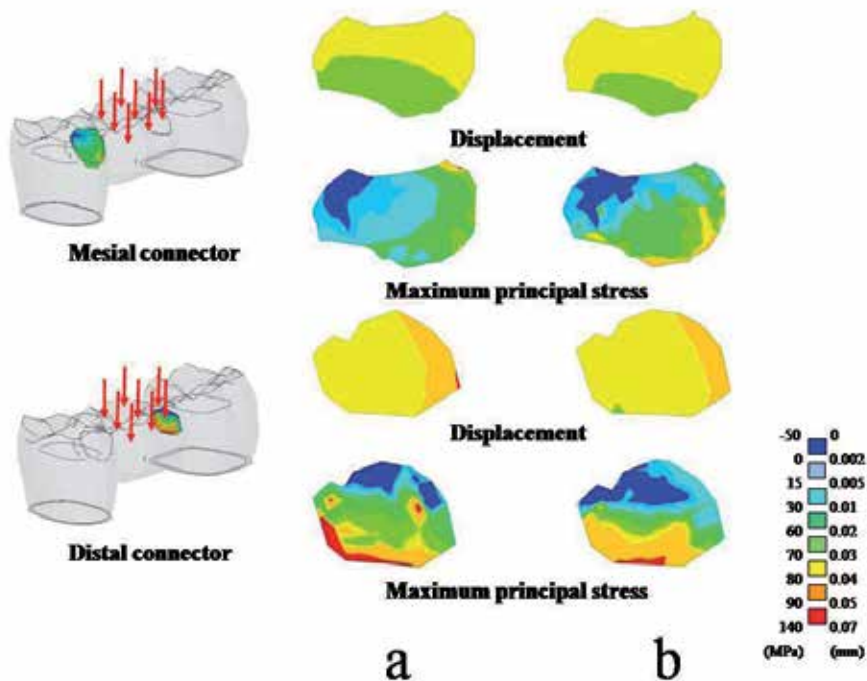


Fig. 39. Vertical load at buccolingual cross section of mesial and distal connector a: C-FPD. b: FRC-FPD

Discussion

In this modeling study, it was demonstrated that there are some differences in mechanical behavior of isotropic particulate composite resin and anisotropic FRC FPDs. Also, the stress distribution variation by changing the loading condition was seen. This is important because it has been reported that there might occur delamination of veneering composite from the FRC framework because at undesirable loading conditions (lateral load) with some framework design.

Mechanical behavior of the FPD under lateral load

When taking into consideration the displacement distribution, no significant difference was found between C-FPD and FRC-FPD, showing a similar distribution pattern in all reference cross sections. Lateral load applied to incline of buccal-cusp from lingual direction caused bending force and it was concentrated at buccal side of the pontic. This is probably due to FRC framework placed on the bottom of the pontic could not exhibit good load-bearing capability and buccal-cusp was not veneered with high enough quantity of hybrid resin (insufficient thickness), which is most likely inducing delamination of the composite veneers. When compared maximum principal stress distribution of C-FPD with that of FRC-FPD, higher stress concentrations were not found at the pontics. Stress generated in the pontic of C-FPD was transferred from the areas to the bottom of the FRC framework with higher Young's modulus material (horizontal cross-section) in the FRC-FPD. No significant

difference in the displacement distribution pattern was noted between C-FPD and FRC-FPD. This may suggest that longitudinal FRC along the direction of framework does not provide any support for the veneer over the regular veneering composite.

Mechanical behavior of a bridge under vertical load

Viewed from buccolingual cross section, the stress found in the side and the bottom of the pontic was transferred to the FRC framework, and relatively low stress of 30~60 MPa was found in the buccal side of the pontic base which was transferred to FRC framework, demonstrating good load-bearing capacity that FRC framework can withstand.

Occlusal contact and guideline for FRC framework design

In this analysis, a total of 629N vertical load was applied to eight points on the occlusal surface of the pontic to simulate maximum occlusal force, and a total of 250N lateral load was applied to three points on buccal-cusp incline of the pontic at an angle of 45 degrees from lingual direction to simulate lateral movement of jaw. This can be considered as worst-case scenario loading condition when there is no cuspid protection available and teeth are loaded laterally.

The most effective way to reduce stress concentration in mesiodistal connectors and buccal side of the pontic is to decrease magnitude or avoid occlusal contacts of buccal side of the pontic. Technically, the pontic can be reinforced by additional fibers on the main framework. It was suggested that in order to get better stress distribution to the FRC-FPD, the framework design should contain FRC perpendicular to that of main framework. To obtain optimal hygiene, function and esthetics, a wide variety of pontic design are proposed. As for occlusal morphology, buccolingual diameter of a pontic should be 2/3~3/4 smaller than that of a missing tooth, and as for occlusal contact, centric stop is defined as an important occlusal contact. Rudo⁴⁸⁾ et al suggested in his study that to resist twisting force generated in pontic, orthotropic FRC framework is more effective to anisotropic, unidirectional FRC framework with high strength and stiffness in one direction. Rudo⁴⁸⁾ et al also suggested the placement of FRC on the outermost surface of a bridge where crack might initiate. In considering plaque accumulation and esthetics, however, placing FRC on the outermost surface of a bridge can be difficult. There is evidence that exposed fibers, especially polyethylene fibers, adhere significant quantities of oral pathogens on the surface⁴⁹⁻⁵²⁾. The authors suggest that it is more appropriate to place FRC framework to the bottom of the pontic, where crack might initiate. The results of this study indicate that high stress concentration in the connectors can be reduced by weakening contacts of buccal cusp. When applied 250 N lateral loads to simulate loading the cusp in lateral movement of jaw, localized high stress concentrated in upper line angle of buccal side of mesial connector in C-FPD was disappeared by the reinforcement of FRC framework, avoiding high stress concentration. Peak stress value of 80 MPa is not likely to cause fracture or delamination of the veneer. When applying 629 N vertical load to simulate maximum occlusal force, higher than 90MPa stress were concentrated in upper line angle of the buccal side of the mesial connector and from lingual side to the bottom of the distal connector in C-FPD. The high stress concentration in the mesiodistal connectors was thought to have been caused by twisting moment imposed on the pontic. With the FRC-FPD, this localized high stress concentration area was subsequently transferred to be carried by the FRC framework, and

thus avoiding the high stress concentrations. In addition to stress-transfer mechanism, the FRC framework reinforcement also improved the rigidity of the FPD. FRC framework bore more than 90MPa of stress, but considering superb flexure strength and load-bearing capacity of the FRC framework, it corresponds to one seventh of flexure strength FRC framework exhibits. These test results confirmed that correctly designed FRC reinforced composite system possesses adequate strength and safety to withstand maximum occlusal force.

Conclusion

These results clarified the magnitude, distribution of stress and displacement generated in C-FPD and FRC-FPD of pontic caused by two different loading conditions. The results suggested that in order to get better stress distribution to the FRC-FPD, the framework design should contain additional pontic FRC perpendicular to that of main framework.

4. References

- Korioth TW, Versluis A. Modeling the mechanical behavior of the jaws and their related structures by finite element (FE) analysis. *Crit Rev Oral Biol Med* 1997; 8: 90-104.
- Ausiello P, Apicella A, Davidson CL, Rengo S. 3D-finite element analyses of cusp movements in a human upper premolar, restored with adhesive resin-based composites. *J Biomech* 2001; 34: 1269-1277.
- Ausiello P, Apicella A, Davidson CL. Effect of adhesive layer properties on stress distribution in composite restorations a 3D finite element analysis. *Dent Mater* 2002; 18: 295-303.
- Ausiello P, Rengo S, Davidson CL, Watts DC. Stress distributions in adhesively cemented ceramic and resin-composite class II inlay restorations: a 3D-FEA study. *Dent Mater* 2004; 20: 862-872.
- Cattaneo PM, Dalstra M, Frich LH. A three-dimensional finite element model from computed tomography data: a semi-automated method. *Proc Inst Mech Eng [H]* 2001; 215: 203-213.
- Cattaneo PM, Dalstra M, Melsen B. The transfer of occlusal forces through the maxillary molars: a finite element study. *Am J Orthod Dentofacial Orthop* 2003; 123: 367-373.
- Shinya K, Shinya A, Nakahara R, Nakasone Y, Shinya A. Characteristics of the Tooth in the Initial Movement: The Influence of the Restraint Site to the Periodontal Ligament and the Alveolar Bone. *The Open Dentistry Journal* 2009; 3: 61-67.
- Ootaki M, Shinya A, Gomi H, Shinya A. Optimum design for fixed partial dentures made of hybrid resin with glass fiber reinforcement on finite element analysis: Effect of vertical reinforced thickness to fiber frame. *Dent Mater J* 2007; 26: 280-289.
- Versluis A, Tantbirojn D, Pintado MR, DeLong R, Douglas WH. Residual shrinkage stress distributions in molars after composite restoration. *Dent Mater* 2004; 20: 554-564.
- Vallittu PK. The effect of glass fiber reinforcement on the fracture resistance of a provisional fixed partial denture. *J Prosthet Dent* 1998; 79: 125-130.
- Vallittu PK, Sevelius C. Resin-bonded, glass fiber-reinforced composite fixed partial dentures: A clinical study. *J Prosthet Dent* 2000; 84: 413-418.
- Bae JM, Kim KN, Hattori M, Hasegawa K, Yoshinari M, Kawada E, Oda Y. The flexural properties of fiber-reinforced composite with light-polymerized polymer-matrix. *Int J Prosthodont* 2001; 14: 33-39.

- Lastumaki TM, Kallio TT, Vallittu PK. The bond strength of light-curing composite resin to finally polymerized and aged glass fiber-reinforced composite substrate. *Biomaterials* 2002; 23: 4533-4539.
- Handa I, Shinya A, Gomi H, Sailynoja ES. Bending strength of fiber reinforced posterior hybrid resin material. *J J Dent Mater* 2003; 22:171-180.
- Sailynoja ES, Shin-ya A, Gomi H, Ishii Y. The effect of immersion temperature on the flexural strength of a pre-coated fiber reinforced composite resin. *Prosthodont Res Pract* 2003; 2:1-10.
- Nakamura T, Waki T, Kinuta S, Tanaka H. Strength and elastic modulus of fiber-reinforced composites used for fabricating FPDs. *Int J Prosthodont* 2003; 16:549-553.
- Drummond JL, Bapna MS. Static and cyclic loading of fiber-reinforced dental resin. *Dent Mater* 2003; 19:226-231.
- Tezvergil A, Lassila LVJ, Vallittu PK. Composite-composite repair bond strength. Effect of adhesion primers. *J Dent* 2003; 31: 521-525.
- Tezvergil A, Lassila LVJ, Vallittu PK. Strength of adhesive-bonded fiber-reinforced composites to enamel and dentin substrates. *J Adhes Dent* 2003; 5: 301-311.
- Vallittu PK. Survival rates of resin-bonded, glass fiber-reinforced composite fixed partial dentures with a mean follow-up of 42 months: A pilot study. *J Prosthet Dent* 2004; 91:241-246.
- Sailynoja ES, Shin-ya A, Koskinen MK, Salonen JI, Masuda T, Shin-ya A et al. Heat curing of UTMA-based hybrid resin: Effects on the degree of conversion and cytotoxicity. *Odontology* 2004; 92: 27-35.
- Shimizu K, Shinya A, Gomi H, Nakasone Y. Fatigue properties of hybrid resin with glass fiber reinforcement ; The influence of cyclic loading, glass fiber reinforcement and test environments. *J J Dent Mater* 2004; 23:294-305.
- Bae JM, Kim KN, Hattori M, Hasegawa K, Yoshinari M, Kawada E, et al. Fatigue strengths of particulate filler composites reinforced with fibers. *Dent Mater J* 2004; 23:166-174.
- Lassila LVJ, Tezvergil A, Lahdenpera M, Alander P, Shin-ya A, Shin-ya A, et al. Evaluation of some properties of two fiber-reinforced composite materials. *Acta Odontol Scand* 2005; 63:196-204.
- Nakamura T, Ohyama T, Waki T, Kinuta S, Wakabayashi K, Takano N, et al. Finite element analysis of fiber-reinforced fixed partial dentures. *Dent Mater J* 2005; 24:275-279.
- Fennis WM, Tezvergil A, Kuijs RH, Lassila LVJ, Kreulen CM, Creugers NH, et al. In vitro fracture resistance of fiber reinforced cusp-replacing composite restorations. *Dent Mater* 2005; 21: 565-572.
- Waki T, Nakamura T, Nakamura T, Kinuta S, Wakabayashi K, Yatani H. Fracture resistance of inlay-retained fixed partial dentures reinforced with fiber-reinforced composite. *Dent Mater J* 2006; 25:1-6.
- Shinya A, Matsuda T, Shinya A, Nakasone Y. Hybrid resin fixed partial dentures reinforced with glass fiber: Optimum posterior fiber frame design with finite element analysis. *J J Dent Mater* 2004; 23: 183-192.
- Fujita T. Textbook of Dental Anatomy, ed 22. Tokyo; Kanehara, 1995:75-104.
- el-Ebrashi MK, Craig RG, Peyton FA. Experimental stress analysis of dental restorations. VII: Structural design and stress analysis of fixed partial dentures. *J Prosthet Dent* 1970;23:177-186.

- Krejci I, Reich T, Lutz F, Albertoni M. An in vitro test procedure for evaluating dental restoration systems. 1. A computer-controlled mastication simulator [in German]. *Schweiz Monatsschr Zahnmed* 1990;100:953-960.
- Belvedere PC. Single-sitting, fiber-reinforced fixed bridges for the missing lateral or central incisors in adolescent patients. *Dent Clin North Am* 1998;42:665-682.
- Göhring TN, Mörmann WH, Lutz F. Clinical and scanning electron microscopic evaluation of fiber-reinforced inlay fixed partial dentures: Preliminary results after one year. *J Prosthet Dent* 1999;82:662-668.
- Shinya A, Matsuda T, Shinya A, Nakasone Y. Hybrid resin fixed partial dentures reinforced with glass fiber -Optimum posterior fiber frame design with finite element analysis-. *J J Dent Mater* 2004;23:186-192.
- Farah JW, Craig RG, Merouel KA. Finite element analysis of a mandibular model. *J Oral Rehabil* 1988;15:615-624.
- Farah JW, Craig RG. Finite element stress analysis of a restored axisymmetric first molar. *J Dent Res* 1974;53:859-866.
- Rubin C, Krishnamurthy N, Capilouto E. Stress analysis of the human tooth using a three-dimensional finite element model. *J Dent Res* 1983;62:82-86.
- Aydin AK, Tekkaya AE. Stresses induced by different loadings around weak abutments. *J Prosthet Dent* 1992;68:879-884.
- Freilich MA, Meiers JC, Duncan JP, Goldberg AJ. Fiber-reinforced composites in clinical dentistry. Chicago: Quintessence, 2000:1-21.
- Rosentritt M, Behr M, Lang R, Handel G. Experimental design of FPD made of all-ceramics and fiber-reinforced composite. *Dent Mater* 2000;16:159-165.
- Magne P, Perakis N, Belser UC, Krejci I. Stress distribution of inlay-anchored adhesive fixed partial dentures: A finite element analysis of the influence of restorative materials and abutment preparation design. *J Prosthet Dent* 2002;87:516-527.
- Yang HS, Lang LA, Felton DA. Finite element analysis on the effect of splinting in fixed partial dentures. *J Prosthet Dent* 1999;81:721-728.
- Kelly JR, Tesk JA, Sorensen JA. Failure of all-ceramic fixed partial dentures in vitro and in vivo: Analysis and modeling. *J Dent Res* 1995;74:1253-1258.
- Meiers JC, Kazemi RB, Donadio M. The influence of fiber reinforcement of composites on shear bond strengths to enamel. *J Prosthet Dent* 2003;89:388-393.
- Bakke M, Holm B, Jensen BL, Michler L, Möller E. Unilateral, isometric bite force in 8-68-year-old women and men related to occlusal factors. *Scand J Dent Res* 1990;98:149-58.
- Tortopidis D, Lyons MF, Baxendale RH, Gilmour WH. The variability of bite force measurement between sessions, in different positions within the dental arch. *J Oral Rehabil* 1998;25:681-686.
- Lundn D, Laurell L. Occlusal force pattern during chewing and biting in dentitions restored with fixed bridges of cross-arch extension. I. Bilateral end abutments. *J Oral Rehabil* 1986;13:57-71.
- Rudo DN, Karbhari VM. Physical behaviors of fiber reinforcement as applied to tooth stabilization *Dent Clin North Am* 1999;43:7-35.
- Tanner J, Carlén A, Söderling E, Vallittu PK. Adsorption of parotid saliva proteins and adhesion of *Streptococcus mutans* ATCC 21752 to dental fiber-reinforced composites. *J Biomed Mater Res* 2003;15:391-398.

- Tanner J, Robinson C, Söderling E, Vallittu PK. Early plaque formation on fiber reinforced composites in vivo. *Clin Oral Invest* 2005;9:154-160.
- Tanner J, Vallittu PK, Söderling E. Effect of water storage of E-glass fiber-reinforced composite on adhesion of *Streptococcus mutans*. *Biomaterials* 2001;22:1613-1618.
- Tanner J, Vallittu PK, Söderling E. Adherence of *Streptococcus mutans* to an E-glass fiber-reinforced composite and conventional restorative materials used in prosthetic dentistry. *J Biomed Mater Res* 2000;49:250-256.

Application of finite element analysis in root canal therapy

Tao Hu, Ran Cheng, Meiying Shao, Hui Yang,
Ru Zhang, Qianhua Gao and Liyang Guo
West China College of Stomatology, Sichuan University

1. Introduction

Finite element analysis (FEA) is an engineering method for the numerical analysis of complex structures based on their material properties. It is used to determine the distribution of stress when a structure is subjected to force. Recently, FEA has been employed in endodontic. Endodontic diseases, such as pulpitis or periradicular periodontitis, often require root canal therapy. During or after treatment, the pulpless teeth are more prone to fracture as an increasing amount of tooth structure is lost from disease or operative procedures.

An important cause of endodontic failure is vertical root fracture (VRF), defined as a longitudinal fracture confined to the root that usually begins on the internal wall and extends outward to the root surface (Walton 2002). VRF is an increasingly common cause of failure of tooth restoration. Beside the difficult diagnosis, the management often requires rather aggressive approaches, i.e., extraction or root amputation. Although a conservative approach has been suggested, the long-term outcomes are still questionable (Kawai & Masaka 2002, Trope & Rosenberg 1992).

It has been reported that approximately 0.2% of root filled teeth each year are lost directly because of VRF (Dammaschke *et al.* 2003, Fuss *et al.* 1999). Due to the high prevalence, poor diagnosis and prognosis of VRF, prevention is therefore of great importance. To avoid the problems caused by VRF, a full understanding of the etiological factors of root fractures is pivotal before preventive measures can be developed. Finite element analysis was first used more than 10 years ago to determine the factors that influencing fracture susceptibility. This chapter will focus on the contribution of finite element analysis to root canal therapy.

Root canal therapy and the use of intraradicular dowels are the two main iatrogenic factors associated with VRF. In root canal instrumentation, enlargement of the coronal third of the root canal space is considered important to support root canal length measurement, debris removal, and canal obturation. However, extensive use of rotary instruments, to be cut dentin to straight lines at curvatures during preparation of the canal space, weakens the root structure. With infected root canals especially, a balance should therefore be sought between the need to remove infected dentin and to maintain sufficient root thickness to withstand the forces of mastication. Special attention to securing sufficient remaining dentin should be given to the teeth and roots most susceptible to fracture, i.e., the maxillary and mandibular premolars and the mesial roots of the mandibular molars.

2. FEA of root canal instrumentation

2.1 Analysis of different conventional preparation techniques

Over decades, clinicians have promoted several preparation techniques for canal instrumentation. Here, we introduced three common techniques.

2.1.1 Instrumentation using the crown-down technique

In the crown-down technique, the clinician essentially works from the crown of the tooth to shape the canal towards the apex. For example, one may start from the coronal portion with gates gliddens that do the coronal flaring or initiate the coronal widening with a #35 K file. This file will only pierce a few millimeters into the coronal aspect of the canal, which is then widened carefully by the circumferential filing. Following this coronally initiated instrumentation, smaller files will penetrate ever deeper into the canal until the smallest files reach the root apex, and achieve a funnel shape.

2.1.2 Instrumentation using the step-back technique

In the step-back technique, the dentist instruments the apical portion first and then progressed toward coronal. Following the bulk removal of the pulp in the extirpation phase, small hand files are adopted nearest the apex. The dentist progresses to larger sizes of files and works back up the canal. A little bit more dentin is cut and removed by each file. Usually these advance about one millimeter with each new instrument. Subsequently the coronal portion of the root canal is conically shaped using the step-back technique through four additional instruments sizes. Finally, the original root canal obtains a funnel shape.

2.1.3 Instrumentation using the reverse-flaring technique (Yang *et al.* 2004)

In the reverse-flaring technique, emphasis is placed on flaring the coronal portion of the preparation before the completion of the apical portion. The dentist employs the hand instruments (files) destined for later apical instrumentation in a reverse order. Because the canal entrance is an anatomically determined narrowing, it is referred to as a "coronal constriction". Therefore, the early elimination of this constriction simplifies the subsequent instrumentation steps. "Reverse flaring" is performed in this coronal widening. After the coronal widening, the K file is adopted again to widen the apical portion. During the procedure of apical instrumentation of the root canal, the small files must be applied over a sufficient time. In curved canals, more effective preparation of the apical area will be made if the file has fewer obstructions in the coronal part.

2.1.4 Comparison of different conventional preparation techniques

As discussed above, different preparation techniques induce various canal morphologies. To assess the stress and fracture possibilities, Finite Element Analysis (FEA) has been used to estimate these preparation techniques.

Cheng *et al.* (2007, 2009) studied the stress distribution on endodontically-treated teeth with curved canals under various loads and determined the differences among three preparation techniques. They pointed out that the three techniques (crown-down, step-back and reverse-flaring techniques) displayed a similar stress distribution at the lower part of FEA model when occlusal loads and condensation loads were applied. In the case of vertical

condensation, the maximum stress was close to the reported tensile strength of dentin. The warm vertical compaction technique was likely to result in root fractures when excessive compaction forces (50 N) were loaded. The FEA model also revealed a tendency toward stress concentration below the compacting level.

3. Root canal morphology after root preparation

There is reasonable evidence for the putative causal linkage between endodontic treatment and tooth fracture. Predisposing factors of root fracture of endodontically-treated teeth include altered physical properties of dentin, i.e., moisture loss in pulpless teeth, previous cracks in the dentin, loss of alveolar bone support, etc. However, based on both clinical and experimental studies, the main cause is considered to be changes in gross canal morphology, including loss of dentin thickness, altered canal curvature, and altered canal cross-section shape. It is likely that these factors interact cumulatively to influence distribution of stresses and tooth vulnerability to loading, ultimately increasing the possibility of catastrophic failure.

To discuss the relative contribution of geometrical parameters after root canal preparation to tooth fracture, FEA models have been constructed to analyze the stress distribution of teeth quantitatively after root canal preparation. By far, the most repeatedly discussed morphological parameters affecting stress distribution on modified FEA models are dentin thickness, radius of canal curvature, canal cross-sectional shapes, canal irregularities, and canal taper.

3.1 Dentin thickness

As reported in many clinical and experimental studies, the dentin thickness is in inverse proportion to the fracture susceptibility. Using FEA, Ricks-Williamson *et al.* (2009) found that the magnitude of generated radicular stresses was directly correlated with the simulated canal diameters. Wilcox *et al.* (1997) found that root surface craze lines formed on roots where greater percentages of the canal wall were removed.

Hong *et al.* (2003) constructed FE models of the mandibular first molar with diameter of root canal modified to 1/4, 1/3 and 1/2 of that of the root. They revealed that enlargement of the root canal resulted in an increased concentration of stress on the root canal surface at the orifice and coronal 1/3. In the lower part of the root canal, the stress distribution disparity smoothened out with the decrease of stress on the root canal surface.

Cheng *et al.* (2006) studied the stress distribution with FEA models of normal wall thickness (12 mm) and roots with 75%, 50% and 25%, respectively, of normal wall thickness. They found that the enlargement of root canal diameter increased the stress of the root canal wall up to 37% under lateral loading. Sathorn *et al.* built eight FE models indicating that the more dentin removed, the greater the fracture susceptibility. These results also concur with those reported by Wilcox *et al.* (1997).

In many clinical and experimental observations, the typical and most frequently observed vertical root fracture pattern is in a predominantly buccolingual direction, which occurs outwardly from the wall of root canal and through the thickest part of dentin.

Lertchirakarn *et al.* (2003a,2003b) also showed the non-uniform tensile-stress distribution is in a predominantly buccolingual direction, even though the dentin in this direction is

typically thicker than in the mesiodistal direction. They constructed a series of FEA models with progressive reduction in proximal dentin thickness, which shows that decreasing proximal dentin thickness increased the tendency for buccolingual stress concentration and hence the predisposition to buccolingual fracture, but not the direction of maximum tensile stress.

This pattern is not expected by most because of the prediction that fracture is more likely to happen through the thinnest part of dentin. The authors explain the mechanism of this counterintuitive phenomenon is that when pressure is applied in a thick-walled vessel, stresses in the wall are of two types: tensile stress in a circumferential direction, and compressive stress in a radial direction. The thin part of the wall will be forced to expand more readily than the thick part of the wall in a radial direction. The asymmetrical expansion creates additional circumferential tensile stresses on the inner surface of the thicker areas, resulting from the outward bending of the thinner part of the dentin wall.

Sathorn *et al.* (2005a, 2005b) further demonstrated another less typical pattern of vertical root fracture: mesio-distal fracture pattern. Canal diameters of 1.0, 1.5, and 2.0 mm were created and the corresponding thinnest parts of dentin were 0.75, 0.5, and 0.25 mm, respectively. As canal diameters were increased in the models, the notional fracture load was correspondingly lower. When the dentin thickness decreases, the stresses are not concentrated specifically on the buccal and lingual aspects in the internal canal wall any more, the external surface stresses in the proximal area increase progressively, and the fracture pattern may change from buccal-lingual to mesio-distal.

However, reduced dentin thickness does not necessarily result in increased fracture susceptibility. Lertchirakarn *et al.* logically speculated that the reduction of the degree of the curvature inside the root canal could reduce fracture susceptibility and that changing the canal shape from oval to round actually relieves internal stresses despite the substantial thinning of proximal dentin. Sathorn. *et al.* proved that elimination of stress-raised areas, as in buccal and lingual extremities of ribbon shaped canals, reduces susceptibility to fracture. These results suggest that the contribution to fracture susceptibility by loss of dentin thickness might be overwhelmed by other morphological parameters, i.e., canal curvature.

In conclusion, stress analysis data obtained by FEA indicates that, with a moderate preparation, dentin thickness is not a determining factor of fracture site or direction. However, with further loss of dentin, an increase of stress concentration occurs, and the vulnerability to root fracture is increased.

3.2 Radius of canal curvature

Knowing the severity of root canal curvature is essential to selecting the instrument and instrumentation technique. What is more, root canal curvature is also a determinant of the prognosis of instrumented teeth, taking a variety of reported complications (apical foramen, creation of ledges, elbows, zips, perforations, instrument fracture, and vertical root fracture) into consideration. Several FEA models have been constructed to discuss the relationship between the radius of canal curvature and stress distribution.

Lertchirakarn *et al.* (2003a) indicated that circumferential tensile stresses were concentrated on the buccal and lingual surfaces of the canal wall, corresponding to areas of greatest canal wall curvature, suggesting that the fracture initiates from the site of greatest curvature of the root canal wall and propagates to the outer root surface. FEA models demonstrated that changing the outer root shape from round to oval, with a round canal, resulted in a smaller

increase in maximum tensile stress than changing the inner canal shape from round to oval, leading to the conclusion that canal curvature is more important than external root morphology, in terms of stress concentration.

Sathorn *et al.* (2005b) constructed an FEA model and modified the root cross-section to a smooth ovoid shaped canal to simulate the root canal preparation, resulting in the reduction of degree of curvature. In the model, the notional fracture load of 40 N was almost double that of the previous model. The stress pattern distribution was much more uniform than in the previous model, and highly localized stress was not found in any area other than the buccal and lingual canal walls. Thus, the fracture susceptibility is reduced.

According to previous studies, it is reasonable to conclude that a prepared root canal with an increased canal curvature radius reduces the stress concentration, and thereby, reduces fracture susceptibility.

3.3 Canal cross-sectional shapes and canal irregularities

Irregular canal cross-sectional shapes and canal irregularities are not only unfavorable factors in the operation of root canal therapy, but also compromise the prognosis of the treated teeth. Canal cross-sectional shapes include circular shape, oval shape, flat shape, and ribbon-shape canals. Canal irregularities include multiple foramina, additional canals, fins, deltas, intercanal connections, loops, 'C-shaped' canals and accessory canals. FEA models are constructed to simulate their influence on stress distribution in the studies described below.

When the canal shape is not circular, the stress distribution becomes asymmetrical. The tendency for highest stresses in a buccolingual direction was greater in the mandibular incisor, with its more pronounced oval root shape and ribbon-shaped canal shape, which suggests that VRF is not the direct outcome of hoop stresses uniformly distributed around the canal, but arises from asymmetrical stress concentrations, and that canal shape and localized irregularities in the canal wall may serve to raise stresses even further to the point of initiation of the fracture. A comparison of models which have the same external root shape showed that changing the canal shape from oval to round actually relieves internal stresses despite the substantial thinning of proximal dentin (Lertchirakarn *et al.* 2003a).

Sathorn *et al.* (2005b) constructed an FEA model with ribbon-shaped root canal and modified the root cross section to a smooth ovoid shaped canal to simulate the root canal preparation. The stress pattern distribution was then much more uniform than in the previous model, and highly localized stress was not found in any area other than the buccal and lingual canal walls. Thus, the fracture susceptibility was reduced.

By FEA, Versluis *et al.* (2006) also indicated that round canals showed lower uniform distributions, whilst oval canals showed uneven distributions with high concentrations at the buccal and lingual canal extensions and greater stresses in the coronal and middle thirds than in the apical third. After root canal preparation, round canals introduced only small circumferential stress increases in the apical half, while oval canals produced substantial reductions where the canal was enlarged to a smooth round shape. Even when fins were not contacted by the instrument, stresses within the root were more evenly distributed and reduced by up to 15%.

Round canal shape has long been proved to have more fracture resistance than that of oval canal shape. Moreover, canal shape as a morphological parameter brings about more profound changes to stress distribution than root shape. Elimination of canal irregularities is certain to provide a more evenly distributed canal stress.

3.4 Canal Taper

The prepared canal diameter has also long been proved to influence the propensity for vertical root fractures. Generally, taper should be sufficient to permit deep penetration of spreaders or pluggers during filling, but should not be excessive to the point where procedural errors occur, and the root is unnecessarily weakened. Holcomb *et al.* (1987) remarked that there must be a point at which increased canal width and taper begin to weaken the root.

To analyze the effect of the change of the tapers on the stress and the distribution on the root canal wall resulting from root canal preparation, Shi *et al.* established a three-dimensional finite element model of maxillary first premolar (single root and type IV root canal), then modified the model according to different root canal tapers after preparation. The maximum stress on the root canal wall and its distribution under vertical or lateral pressure were calculated, revealing that the enlargement of the root canal tapers after preparation can result in the rise of stress on the root canal wall, especially under lateral pressure.

Rundquist *et al.* (1983) constructed three FEA models of a root filled premolar tooth varying only in canal taper. The stress distribution in the root during the occlusal loading and filling was both recorded and compared with each other. During filling, root stress decreases as the canal taper increases, with generated stresses being greatest at the apex and along the canal wall. After root filling is complete and occlusal force is applied, the relationship is reversed. The generated stress is greatest at the cervical portion of the root surface, and increases slightly as taper increases. The authors concluded that it is possible vertical root fractures initiated at the apex are a result of filling forces, whereas vertical root fractures initiated cervically are a manifestation of subsequent masticatory events on the root filled tooth.

However, due to the complexity of relationship of taper and stress distribution (canal taper is in inversely proportionate to the generated stress value under filling forces while directly proportionate to the generated stress value under occlusal forces), and that the relative contribution of filling forces and occlusal forces to VRF remains illusive, further study is required.

3.5 Clinical suggestions

Many morphological factors (dentin thickness, radius of canal curvature, canal shape and size, canal taper, etc.) interact with each other to influence fracture susceptibility and pattern, and any one variable can easily predominate over the others.

Stressing the significance of canal anatomy, Peters *et al.* reported that variations in canal geometry before shaping and cleaning procedures had more influence on the changes that occur during preparation than the instrumentation techniques themselves. However, endodontic and restorative procedures have been addressed as absolute precipitating factors for VRF. On that basis, the results from the previous reports provide an experimental foundation to better understand the morphological parameters affecting stress distribution after root canal preparation and their relative contribution to VRF, which are of great importance in the clinical practice of root canal preparation to prevent VRF.

Canal preparation should be as conservative as practical, consistent with adequate cleaning and shaping. A smoothly rounded canal shape is favorable and can eliminate stress concentration sites, which inevitably result in higher fracture susceptibility. Mechanical instrumentation of root canals can produce craze lines on the root canal wall, which may serve as locations of increased stress (in accordance with stress-concentration theory).

Instrumentation procedures should be undertaken with gentle force, using generous irrigation to avoid crazing. The dentist's goal should be to create a root canal shape that maximizes radius of curvature of the root canal wall. A circular shape minimizes stress concentration areas and will distribute stress more uniformly. Furthermore, procedural pitfalls that create stress concentration sites on the root canal wall, i.e., ledging, gouging, and crazing, should be avoided. Although canal cross-sectional shape seems more important than dentin thickness in stress distribution, removal of root dentin should be minimized. By maintaining dentin thickness as much as possible, especially in the proximal areas or in the thin part of root dentin, stress will be minimized.

4. FEA of Nickel-Titanium Rotary Instruments

In root canal therapy, instrument fracture is a potential consequence of canal instrumentation, especially when the instrument is bound at the tip. FEA has been employed to compare the stresses in a number of Nickel-Titanium Rotary instruments.

4.1 FEA in investigation of the mechanical properties of different NiTi rotary instruments

With the increased use of nickel-titanium (NiTi) rotary instruments for root canal therapy in endodontics, instrument fracture has become more and more prevalent. Extensive research has been carried out on the physical properties and mechanical characteristics of NiTi rotary instruments. Kim *et al.* (2008a) estimated the residual stress thereafter for some nickel-titanium rotary instruments (ProFile, ProTaper, and ProTaper Universal) using a 3-dimensional finite-element package. The simulation in the ProTaper design revealed that there was the greatest pull in the apical direction and the highest reaction torque from the root canal wall, while the least stress in occurred in the ProFile design. Stresses in ProTaper were concentrated at the cutting edge, and the residual stress reached a level that was close to the critical stress for phase transformation in the material. The residual stress was highest in ProTaper (see below for the ProTaper Universal and ProFile design).

Flexibility and fracture properties are determinant for the performance of NiTi rotary instruments. Kim *et al.* (2008b) evaluated geometrical differences between three NiTi instruments which affect the deformation and stress distributions under bending and torsional conditions. ProFile, with a U-shaped cross section, showed the highest flexibility among the three file models. The ProTaper, which has a convex triangular cross-section, was the stiffest file model. In the ProTaper, more force is required to reach the same deflection as the other models, and more torque is needed than other models to achieve the same amount of rotation.

Under torsion, all NiTi files showed the highest stress at their groove area. The ProFile showed the highest von Mises stress value under the same torsional moment whereas the ProTaper Universal showed the highest value under the same rotational angle. Additionally, the assessment of the stress distributions of three NiTi instruments with various cross-sectional configurations under bending or torsional conditional showed that ProTaper has the lowest flexural rigidity of all if a U-shaped groove is incorporated in the middle of each side (Kim *et al.* 2008b, 2009).

4.2 FEA in investigation of the parameters contributing to instrument failure

4.2.1 The radius and the position of the canal curvature

By FEA, Necchi (2008) investigated rotary endodontic instruments and demonstrated the usefulness of the finite element method in simulating the mechanical behaviour of these instruments during root canal preparation. The results indicated that the radius and the position of canal curvature are the most crucial parameters that determined the stress in the instrument, in that higher stress levels are produced by decreasing the radius and moving from the apical to the mid root position. The most demanding working conditions were observed in canals with sharp curves, especially in areas in which the instruments had larger diameters. To prevent the possible damage to instruments and fracture, it is suggested that the instruments should be discarded following their use in such canals.

4.2.2 Cross-sectional design of Nickel-titanium instruments

As NiTi instruments are generally perceived to have high fracture risk during use, new designs with lower fracture risks have been marketed. However, these design variations may also alter the forces distribution on a root during instrumentation and increase the potential dentinal defects that predispose to fracture.

Previous study (Kim *et al.* 2009) has indicated that, in Nickel-titanium instruments with rectangle-based cross-section, higher stress differentials are created during the simulated canal shaping, and higher residual stress and plastic deformation occurs than in instruments with triangle-based cross sections. It has also been shown that different cross-sectional designs affect stress distribution in NiTi instruments during the bending, torsion and simulated shaping of a curved canal.

For three NiTi file designs ProFile (U-shaped cross-section and constant 6% tapered shaft), ProTaper Universal (convex triangular cross-section with notch and progressive taper shaft), and LightSpeed LSX (noncutting round shaft), the stress conditions during the rotary instrumentation in a curved root were also estimated (Kim *et al.* 2010). ProTaper Universal introduced the highest stress concentration in the root dentin and also had the highest tensile and compressive principal strain components in the external root surface. ProTaper Universal had the biggest taper shaft and the calculated stress values from ProTaper Universal approached the strength properties of the dentin. LightSpeed generated the lowest stresses. It can be concluded that the stiffer file designs created higher stress concentrations in the apical root dentin during the shaping of the curved canal, which increases the risk of dentinal defects that may lead to the apical root cracking.

5. FEA of root canal compaction

Teeth with root canal treatment seems to be more susceptible to vertical fracture (VRF) than those with intact pulp. As well, it has been speculated that excessive pressure during the gutta-percha compaction process could lead to root fractures. Meister *et al.* investigated 32 clinical cases of vertical root fractures and postulated that the use of excessive force during canal obturation was the primary cause of these fractures (Meister *et al.* 1980). Studies showed that a spreader load as small as 1.5 kg might produce fracture during lateral compaction in extracted teeth (Holcomb *et al.* 1987, Pitts *et al.* 1983).

FEA was also conducted to estimate the stresses that occurred during compaction procedures. Several results revealed that vertical compaction led to greater stress than

lateral compaction. The maximum stress of vertical condensation was close to the reported tensile strength of dentin (50N), but still lower than the load required to fracture the roots (Cheng *et al.* 2009). The maximum stresses in the root dentin primary emerged during the first gutta-percha increment, afterwards, moved coronally with full condensation of successive gutta-percha cones (Silver-Thorn *et al.* 1999, Rundquist & Versluis 2006). These results lead us to believe that compaction technique when performed skillfully, does not likely create premature root fractures. Thus, obturation should not be regarded as a major cause of VRF, except in very weak roots (Lindauer *et al.* 1989).

Moreover, other factors can be addressed. These factors potentially include the direction and position of force loaded, root canal taper, final canal shape, number of gutta-percha cones, the condition of supporting tissues, damping properties of dental gutta-percha, as well as thermal distribution during compaction.

5.1 The distribution of stress during vertical and lateral compaction

It has been long debated whether excessive application of pressure during compaction could result in root fractures. There have been efforts to understand stresses distributed throughout the curved or straight root canal during the application of different compaction techniques.

For lateral compaction, various standpoints exist for ages, regardless of produced stress magnitude or concentration area. Harvey (1981) noted that throughout condensation, the stress was concentrated in the middle third of the tooth root and moved coronally with full condensation, which coincided with the previous study made by Silver-Thorn (1999). However, Telli *et al.* (1998), consistent with Gimlin *et al.* (1986), showed that lateral loads led to a stress concentration area limited to the apical third of the tooth root. The results above-mentioned were conducted in the straight canal roots. For the curved root canal, the maximum lateral stresses were located in the loading site. The reason why the results appeared such a big difference may account for the diverse of model baseline.

For vertical compaction, a study showed that it led to greater stress than lateral compaction (Cheng *et al.* 2009). The magnitude of stress was close to the reported tensile strength of dentin, which was likely to cause root fracture. It is suggested that inadvertent undue force should not be applied during warm vertical condensation.

5.2 Effect of canal tapers on root stresses during obturation

It has been postulated that the prepared canal diameter may influence propensity for vertical root fractures. It is reasonable to speculate that increasing the taper of the canal preparation by removing more dentin from the canal wall would diminish the structural integrity of the root. Holcomb *et al.* remarked that there must be a point at which increased canal taper begins to weaken the root and might predispose a root to vertical fracture (Holcomb *et al.* 1987). Ricks-Williamson *et al.* found the magnitude of generated radicular stresses to be directly correlated with the simulated canal diameters (1995).

Recently, one investigation was designed to evaluate the effect of different canal tapers on stress distributions to determine whether vertical root fractures occurred at the time of filling or at a later time. The results showed that the stresses tended to be higher closer to the load site, and gradually decreased along the root canal wall until they increased again at the apical constriction during filling; whilst the highest stress levels were obtained during the

first gutta-percha increment and became lower with subsequent increment. However, during the application of the post-filling load, the highest stresses were generated at the external root surface, especially concentrating at the cervical third. Furthermore, with increasing taper, root stresses decreased during root filling but tended to increase with post-filling load. It seems that root fracture derived from at the apical third is likely initiated during obturation, while fracture rooting in the cervical portion is likely caused by occlusal loads. Thus, this emphasizes the need to use moderate compaction forces during the first increment in order to reduce the risk of apical fracture.

5.3 Effect of damping properties on fracture resistance of root filling

It has been well established that the damping properties of the periodontal ligament are the main contributors to reducing the induced strain in a loaded tooth. Yet, other damping materials lie in the teeth, such as dental pulp. Huang *et al.* confirmed that dental pulp was in the nature of the cushioning effect to affect the stress distribution during an impact (Huang *et al.* 2005). A further study (Ou *et al.* 2009) showed that the stress reduction effect in the root filled model was lower than in unprepared intact teeth, but higher than in pulpless teeth. In addition, the stress reduction effect in intact teeth was lower on the loaded than on the opposite side, which may account for why dental pulp can disperse the strain energy and hence reduce the concentrated stress, because of the strain propagation passing through pulp tissue.

It is suggested that the damping properties should be taken into consideration during canal obturation. Dental gutta-percha of a favorable cushioning effect should be better developed in the future.

5.4 Finite element analysis of the thermal distribution

Nowadays, warm vertical compaction is a widely used technique. However, the use of the technique may lead to an unconscious transmission of excessive heat to the surrounding tissues, which may cause irreversible injury to tissues. The use of peak temperature should be well defined.

FEA is the right choice for thermal distribution evaluation, on account of its detection not only for root surface temperature, but also internal distribution of heat. Özgür Er *et al.* established a model of maxillary canine to determine the distribution and level of temperature (2007).

When used with a 200°C initial setting, simulated in a process involving seven stages and lasting for 34 seconds, the maximum temperature lying in gutta-percha is 56.6°C and the periodontal ligament temperature is between 37.3°C and 39.7°C. And the maximum temperature rise was observed at the apical tip of the simulated heat source. It has been reported that alkaline phosphatase is rapidly inactivated at 56°C, which is considered as the critical temperature for bone damage (Matthews & Hirsch 1972). Thus, 200°C as the initiative temperature created no potential harm to the tooth structure and the surrounding tissues.

In brief, although various standpoints exist during filling, those with the highest stresses were in: (a) the apical third; (b) the loading sites; (c) canal with smaller taper; (d) the canal surface; and, (e) during the first gutta-percha increment. In addition, apart from detecting the distribution of stress, FEA also highlights a large variety of filling procedures, and very

accurate simulations can be conducted to investigate what might occur during compaction.

6. Other factors of tooth or periodontal tissue that influence stress distribution

Even though most of the factors contributing to VRF can be prevented by moderate root canal preparation and obturation procedures, still, there are dental and periodontal situations (morphological or pathological), such as oval shape root, proximal concavity, diminishing bone support, internal resorption, root perforation and periapical lesions that are partially responsible for unbalanced stress distribution in the root canal, and which lead directly to root fracture.

6.1 External root morphology

External root morphology, together with root shape are basic factors affecting stress distribution in the root canal wall and external root surface, which further affect the fracture susceptibility of teeth. However, canal shape seemed to be a more substantial factor than the outer root shape.

Lertchirakarn *et al.* (2003a) constructed a series of FEA models. The original model was a simple, thick-walled cylinder with little resemblance to any human tooth root. With progressive changes in inner and outer surface shapes and the reduction in proximal dentin thickness, the final model resembled an idealized, single-canal, mandibular incisor. The results demonstrated that a markedly oval root shape is more susceptible to high stress than a more circular one, regardless of the canal shape, even though the root shape seemed to be of less importance.

Sathorn *et al.* (2005b) also investigated the role of external root morphology, specifically a proximal concavity in mesio-distal fracture susceptibility. In this study, a basic model with a proximal concavity on the mesial and distal root surfaces was created. Both mesial and distal external root surfaces were reduced by 0.2 mm at the mid proximal area, and the resulting root shape was then smoothed to incorporate the concavity into the overall root outline. However, a proximal concavity did not significantly reduce fracture susceptibility as expected, rather it heightened stresses at external proximal surfaces. The concavity in the proximal surfaces itself had very little effect on fracture susceptibility when compared with ordinary roots sharing the same dentin thickness. It was concluded that external root morphology has minor influences on stress distributions of root.

6.2 Teeth with curved roots

Curved roots are another impediment in canal preparation. This is especially true near the apex of maxillary lateral incisors and the palatal root of maxillary first molars. Small, flexible instruments with non-cutting tips negotiate these curves, but larger, stiffer instruments start a ledge that can develop into a perforation. Furthermore, the risk of iatrogenic mishaps increases because root canals tend to be closer to the inner (concave) part of curved roots.

Cheng *et al.* (2009) investigated (by FEA) the stress distribution within roots having curved canals when prepared by three preparation techniques. Three models of preparation by one of crown-down, step-back and reverse-flaring techniques (CDT, SBT and RFT) were established by replacing the inner canal, leaving the outline form unchanged to restrict the

interfering factors. It was concluded that although the prepared canals varied in shape, little difference occurred in stress distribution around the root curvature and apex. The three preparation techniques had similar effects on simulated canals. After appropriate preparation, canals were enlarged and the stress increased to an extent below the tensile strength of dentin.

Lateral and vertical filling techniques were tested under occlusal loading, which revealed that vertical compaction led to greater stress than lateral compaction and excessive vertical compacting force (50 N) was likely to cause root fracture. Vertical compaction also produced a stress concentration area below the compacting level, which might be attributed to the anatomy of curved canal. Comparing the three techniques, RFT led to the least stress under lateral loads. As indicated above, RFT had combined features of both the CDT and SBT models, and as a result, it provided a thicker dentin wall around the curvature, reducing stress under different loadings.

Chatvanitkul *et al.* (2010) created 16 three-dimensional FEA models and estimated stress distribution in various degrees of curved roots with different post and restorations. The results indicated that the stress distribution pattern showed that degrees of root curvature (15, 30, and 45 degrees) had little effect.

The above studies show that degree of root curvature is not decisive to stress distribution; rather, the relevant preparation technique seemed more influential to fracture resistance of the curved root.

6.3 Teeth in pathologic state

Natural teeth with a healthy periodontal support exhibit stress transfer when functional forces are applied to them. These stress patterns show considerable variation during the pathologic state, which may influence both the tooth and supporting alveolar bone.

Telli *et al.* (1999) designed a maxillary canine tooth FEA model, based on a human cadaveric maxilla scanned by CT, in order to investigate the effect of pathological alterations of the dental structures (diminishing bone support, internal resorption, root perforation, periapical lesion). Patterns of stress distribution associated with pathological changes in dental structures were investigated. It was found that, diminishing bone support and internal resorption markedly increased stress magnitudes. However, it was mentioned that these values still remain much below the most frequently reported tensile strength of dentin (50-100 N/mm²). This result led to the conclusion that when warm vertical compaction technique is skillfully performed and unwanted force is not applied, a premature root fracture in a large rooted maxillary anterior tooth with straight root canal anatomy is not likely to occur, even for the unfavourable conditions simulated.

The alveolar ridge's height of tooth was also studied by FEA, in order to simulate the pathological alternations of alveolar ridge and to analyze its effect on the stress distribution on the root canal wall (Hong *et al.* 2002). The height of alveolar ridge was modified by reducing 1/4, 1/3, 1/2 of the height. Lateral and vertical forces were loaded on the models. The modified models represented an increased stress concentration compared to the orifice of root canal, reaching the maximum in the coronal one-third. In models with different heights of the alveolar ridge, when the fixed height was lower, the decrease in stress concentration was less observable. This study indicated that reduction of alveolar ridge height (periodontal disease, periapical disease, etc.) results in increased stress concentration of the root canal, which, in the long run, will lead to root fracture.

Endodontic surgeries adopted to cure relating dental and periodontal pathological states showed considerable variation in stress patterns of the teeth before and after treatment, which may influence both the tooth and supporting alveolar bone.

Uensal E *et al.* (2002) applied FEA methods to evaluate variations in the stress under functional loads on first molars with periodontal furcation involvement, which were treated either with by root resection or root separation. This study used a two dimensional mathematical model of a mandibular first molar that was subjected to either a root separation or a root resection procedure. An evenly distributed dynamic load (600 N) was applied on two buccal cusps and distal fossae of the molar in centric occlusion. It was found that in the root resection model the stress values were maximum on the centre of rotation, and compressive stresses increased towards the middle of the cervical line. For the root separation model, the maximum shear stress values were observed in the distal portion, and a uniform stress distribution was observed in the mesial portion. Shear stress values for bone increased towards the centre in the bifurcation area.

Comparing to other factors that bring about VRF, the investigation of the influence of pathologic state seems scattered and the conclusions more diverse; moreover, the validation process of the FEA results remains absent. This situation is possibly the result of the many dental and periodontal states. More reliable FEA models are imperative to deal with these variations, and promote studies in this area.

7. The combination of two methods, FEA and fracture strength test, in stress analysis of vertical root fracture

7.1 Fracture strength test

Fracture strength test is an engineering discipline focusing on how and why fractures occur, and ultimately, the way to prevent or at least minimize the chance of fracture occurring. To get the general knowledge about dentin fracture mechanics and tooth fracture resistance, it is necessary to understand the biomechanical properties of dentin (Table 1).

	Young's modulus(GPa)	Poisson's ratio	References
Enamel	84.10	0.33	(Kampoosiora <i>et al.</i> 1994)
Dentin	18.30	0.31	
Pulp	2.07×10^{-3}	0.45	
Periodontal ligament	68.90×10^{-3}	0.45	
Cortical bone	10.00	0.30	
Cancellus bone	0.25	0.30	(Lakes 2002)
Restorative material	3.00	0.24	
Gutta-percha	9.30×10^{-4}	0.40	

Table 1. The values of Young's modulus of dental tissues

A study has revealed that the hardness values decreased in dentin closer to the pulp; and the authors concluded that the reduction in hardness was due to the increased tubule density and dentin porosity (Pashley *et al.* 1985). A subsequent study, however, demonstrated that the intertubular dentin near the pulp is highly likely to be less mineralized (Kinney *et al.* 1996).

Young's modulus, known as tensile strength, is another property of dentin. It is different

even in the same tooth at different distances from the DEJ. Studies have shown that the Young's modulus of dentin tends to decrease with the distance from the DEJ toward the pulp (Kinney *et al.* 1996). But studies still have not found any difference of Young's modulus between vital and non-vital dentin.

Moisture content is another aspect that defines tooth mechanical properties. It is reported that decreased moisture content corresponded to increased tooth fracture susceptibility compared with vital teeth (Johnson *et al.* 1976). Therefore, reduced moisture content is considered a common reason for tooth vertical fracture in pulpless teeth. In teeth with pulp extirpation, fracture resistance can be reduced by 9% because of the progressive loss of moisture with time.

It is widely accepted that fracture can be classified as brittle and ductile and it involves crack formation and propagation produced by an imposed stress. To study the fracture susceptibility of roots, a specific testing machine was used in experiments to load roots to the point of fracture. However, because of the large variation of root shape, size and canal shape, size, the information about fracture susceptibility can be very difficult to assess statistically. Moreover, it is difficult to study the effect on fracture susceptibility, because of the uncertainty and irregularity of root and canal as well as the fact that no exact morphologies of roots and canals are available.

Studies aiming to remedy the defects of experimental methods, introduced a three-dimensional computerized numerical method to precisely evaluate and calculate the stress concentration and magnitude. FEA is a computerized method. It can solve complex problems by dividing complex structures (non-geometrical shape, e.g. tooth) into many small-interconnected simple structures (geometrical shape), which are called finite elements. It can directly display the location of stress concentration areas and the intensity of the stress. Given the above-mentioned advantages, FEA can be used to predict fracture patterns and fracture susceptibility. Variations of root shape and loads can be easily incorporated into the calculation to make the results more accurate and authentic. Combined with fracture strength test, FEA can provide comprehensive and convincing information about VRF for practitioners to take better preventive measurements.

7.2 FEA and fracture strength test in root canal therapy

As there is a high occurrence of VRF in endodontically treated teeth, endodontic procedures have been considered as a frequent cause of VRF.

One study, involving combined fracture strength testing and FEA to compare the preparation techniques of hand files and rotary Ni-Ti, demonstrated that the fracture load was almost identical, but the fracture pattern differed, and the FEA models correlated very well with the observed fracture pattern, demonstrating a reliable predictability for VRF (Sathorn *et al.* 2005a).

Another study by the same researchers found that dentin thickness, curvature of the external proximal root surface, canal size, and shape all interact in influencing fracture susceptibility and pattern of fracture (Sathorn *et al.* 2005b). The more dentin removed, the greater the fracture susceptibility. Canal preparation should be as conservative as practical, consistent with adequate cleaning and shaping. A smoothly rounded canal shape is favorable and can eliminate stress concentration sites, to reduce the fracture susceptibility. Studies have also revealed that canal wall curvature was a major influential factor in stress concentration and fracture pattern (Lertchirakarn *et al.* 2003a).

Besides having an application in endodontic procedures, FEA and fracture strength testing have also been used to evaluate different kinds of posts and crowns. For endodontically-treated teeth, it is pivotal to strengthen the fracture resistance. One study revealed that endodontically treated premolars whose coronal hard tissue were severely damaged, obtained higher fracture resistance with the computer-aided design/computer-aided manufacturing ceramic endocrown restoration compared with classical crown configuration (Lin *et al.* 2010). Another study showed that the combination of a fiber post and composite resin core with a full cast crown is most beneficial for the remaining tooth structure under the conditions of vertical and oblique loadings (Hayashi *et al.* 2006). The subsequent study of pulpless teeth with a combination of a fiber post and a composite resin core, compared to restorations using a metallic post, showed superior fracture resistance against both static and fatigue loadings, an approach therefore recommended in restoring pulpless teeth (Hayashi *et al.* 2008). However, a recent study has indicated that the endodontically-treated tooth without coronal tooth structure restored with titanium post showed higher fracture strength values as opposed to the composite post or no-post approach (Ozcan & Valandro 2009). It is reported that, in older patients, teeth receiving root canal treatment with posts are more prone to VRF, especially in those with low dentin thickness (Mireku *et al.* 2010). Consequently, clinicians should pay more attention to avoid VRF in older patients. A recent study using FEA confirmed the hypothesis that lack of an effective bonding between root and post increases the risk of VRF in upper premolars restored with endodontic posts (Santos *et al.* 2009), so an effective bonding can integrate the posts to the root canal. On the contrary, the conditions involving negative bonding will produce a torque force to form higher stress concentration areas, which will cause detrimental VRF to the root. A comparison of a fiber post core and a conventional cast post core system using conventional fracture strength test showed no obvious difference between the two post systems, but the finite elemental stress-analysis method revealed that stress was accumulated within the cast post core system. The stress transmitted to supportive structures and the tooth was low, because the post core system transferred the stress to the supportive structures and the tooth while stress accumulation within the post system was low (Eskitascioglu *et al.* 2002). This reminds the clinician to evaluate tooth conditions carefully before restoration, and to take special care with supporting tissues in their clinical activities. In endodontically-treated, single-rooted teeth with approximal cavities, fracture strength testing showed no advantageous fracture resistance in post groups (Heydecke *et al.* 2001). This conclusion may be different from others'. A study recommended the application of zirconia posts with ceramic cores as an alternative to cast posts and cores (Heydecke *et al.* 2002). The maxillary central incisors restored with a cast post and core demonstrated more vertical root fractures (Pontius & Hutter 2002). The preservation of both internal and external tooth structure is of utmost importance when restoring endodontically-treated teeth.

Experimental fracture strength tests were conducted and teeth had a significantly lower fracture resistance with stainless steel posts in place. The finite element model, from the point of stress distributions, confirmed the results. The failure of stainless steel posts was due to a worse mechanical performance and a high stress concentration because of the significant difference between the elastic modulus of the steel and the surrounding materials (Barjau-Escribano *et al.* 2006). This study again emphasized the importance of the similarity of the elastic modulus of the post, dentin and core, which determines the biomechanical

performance of post systems. A study aiming to test the fracture strength in endodontically-treated premolars with a bonded restoration, full-coverage crown or onlay failed to discover significant difference (Steele & Johnson 1999).

Here, a problem we should note is that many studies are interested in single root teeth, rather than multi-root teeth. VRFs in molars are not infrequent, and happen even in non-treated, intact teeth, and often occur bilaterally. The high masticatory occlusal force is blamed for the occurrence of such unfavorable events. Because of the limitations in current research methods and the diversifications of root morphology, it is very difficult to simulate the real shape and pattern of the teeth, and the mechanical analysis is complex. Construction of a vivid multi-root model is an urgent task, and a formidable challenge, in order to provide theoretical supports for preventing VRF in molars.

8. The future challenge of FEA

Conventional technologies, the two-dimensional analysis of tooth and the strength tests, are unable to accurately determine subtle changes in teeth. In addition, direct assessment of the mechanical properties of teeth through mechanical experiment often brings errors and significant uncertainty to clinicians. This is because mechanical measurements are quite sensitive to friction between the sample and load patterns, and the result is largely dependent on the size and shape of the samples. Therefore, mechanical strength tests may not effectively detect small or even large changes of teeth. To overcome these technological shortcomings, FEA was introduced in measuring the mechanical properties of teeth. However, mastication forces are dynamic, and the magnitude is diverse at different times (Salis *et al.* 2006). From a clinical aspect, the results of monotonic analysis are questionable, not only is the monotonic load not representative of the clinical occlusal loads (Goto *et al.* 2005, Qing *et al.* 2007), but also the corresponding stress to dynamic and static loads are different. Static analysis ignores the stress-damping and stress-transmission effect of the soft tissues; thereby, an unexpected error is unavoidable.

Further, the root canal system is complex, which makes it difficult to completely know the root canal configuration and the stress distribution during endodontic treatment. It is well known that the accessory canals are impossible to clean mechanically. The high incidence of multiple canals and different canal types in premolars and molars, frequency of C-shaped canals in distal root of mandibular second molars, location of apical foramen, level of bifurcation, root canal ramifications, transverse anastomoses, apical deltas and any other variant anatomy are all factors preventing clinicians from successful treatment. FEA based on CT scanning can help to obtain a detailed picture of the complex root canal system, and establish the models to analyze the mechanical property and stress distribution before treatment commencement. This real-time simulation before any treatment undoubtedly facilitates the choice of the best therapeutic strategy. Doing so will protect the tooth structure and minimize the risk of improper operational approaches.

Recently, high-resolution microfocus computed tomography (CT) has been used successfully in endodontics, especially in measuring the diameters of the canal system. In previous studies (Barjau-Escribano *et al.* 2006, Steele & Johnson 1999), microfocus CT (micro-CT) produced valid root canal details in three dimensions, providing a visualized method to estimate canal preparation techniques. Micro-CT can directly compare different restoration conditions with the model of the unaltered tooth structure. Nowadays, with

technological developments, advanced high-resolution CT will be widely applied in clinic. It will exactly detect the subtle variations and dynamic changes of each specific tooth. We can optimistically foresee that the exponential development of commercial dental CT-scanners, computer processing power, and friendly interface will make this approach even faster and more automated, offering the rapid model construction of patient-specific simulations of tooth structure.

Micro-CT also brings great challenges to FEA. Using micro-CT, a precise FEA model of a curved canal was established (Cheng *et al.* 2007). Furthermore, finite element (FE) models can be generated through the conversion of micro-images of the root canal system obtained from micro-CT to simulate real mechanical tests, which will largely eliminate the experimental errors and accurately provide an evaluation model of the root mechanical properties. It is true that small differences may exist between the reality and the finite element environment, but it is the best approach available to reveal the inaccessible stress distribution within the tooth-restoration complex, and the propensity that this numerical method has for increasing the power to imitate reality is irresistible. Therefore, FEA is a useful and crucial tool in endodontics.

Accurate dynamic FEA model is indispensable to investigate the stress concentrations in teeth under dynamic occlusal loads. To generate the model, the numerical analysis method of FEA is proposed to simulate the staged stress variation process. The complex time and special effect of dynamic occlusal loads poses a serious problem in analyzing the stress distribution. As for endodontically-treated teeth, using FEA and micro-CT will definitely help to create the best treatment strategy to keep the teeth in function as long as possible. In conclusion, to gain detailed information about root canal system, the method of three-dimensional modeling based on CT data is a promising approach. The use of FEA based on micro-CT images allows the changes in the microstructural and mechanical properties of teeth. This method has, to a large extent, helped endodontists to uncover the tooth strength properties and to develop effective therapeutic strategies, while the traditional method is far less useful. Moreover, developments in FEA will enable it to be even more practical and accurate.

9. Limitations of FEA

The use of FEA in dental research provides a qualitative and comparative studies method for various stress problems. For the study of root canal therapy, FEA simulates and analyzes the stress within the root canal, on the root canal wall. This kind of study is superior to other techniques. FEA can take different and repeatable experiments on a single or a few typical subjects. It also brings in variables that cannot be approached by other means. In recent years, for the research on root canal therapy, FEA has most often dealt with: the stress and root fracture resistance of different preparation techniques; the stress analysis of obturation; and, the stress analysis of root fracture and vertical fracture.

Although FEA serves as a helpful and multifunctional technique, the limitations of FEA should not be ignored. As a computer solution, it does not necessarily reveal how the stresses are influenced by important variables such as materials properties and geometrical features. When FEA is used to simulate the tooth, the available model assumes the dentin is isotropic, linear elastic, and uniform with a tissue Young's modulus and Poisson's ratio. However, previous studies have shown that the hardness of dentin decreases from the outer

surface to the dental pulp cavity, which makes the Poisson's ratio and Young's modulus not always the same within the dentin. It is also reported that dentin has its own anatomical shapes and structures. Thus, the Young's modulus varies according to the distance from the pulp. Simulating the true structure of the teeth to obtain an objective result is still a challenge for FEA.

In addition, the complicated geometry of root canals and uncertainties about their mechanical properties make it a necessity that calculated values must be corroborated by experimental measurements at certain points. These developmental defects within the teeth are often ignored in establishing an FEA model. However, the defects often potentially initiate the tooth fracture.

Moreover, most FEA experiments assume that the force distributed on the canal surface is uniform. Gutta-percha was simulated to behave like a perfect fluid, distributing the load around the canal wall equally and uniformly. In reality, this seldom happens. A pointed force is more likely the real situation of stress in the root canal under force of preparation or with condensation during obturation.

10. References

- Barjau-Escribano A, Sancho-Bru JL, Forner-Navarro L, Rodríguez-Cervantes PJ, Pérez-González A, Sánchez-Marín FT (2006) Influence of Prefabricated Post Material on Restored Teeth: Fracture Strength and Stress Distribution. *Operative Dentistry* 31:47-54.
- Chatvanitkul C, Lertchirakarn V (2010) Stress distribution with different restorations in teeth with curved roots: a finite element analysis study. *Journal of Endodontics* 36:115-118.
- Chen J, Yue L, Wang JD, Gao XJ (2006) The correlation between the enlargement of root canal diameter and the fracture strength and the stress distribution of root. *Chinese Journal of Stomatology* 41:661-663
- Cheng R, Zhou XD, Liu Z, Hu T (2007) Development of a Finite Element Analysis Model With Curved Canal and Stress Analysis. *Journal of Endodontics* 33:727-731.
- Cheng R, Zhou XD, Liu Z, Yang H, Gao QH, Hu T (2009) Finite element analysis of the effects of three preparation techniques on stresses within roots having curved canals. *International Endodontic Journal* 42:220-226.
- Damaschke T, Steven D, Kaup M, Ott KH (2003) Long-term survival of root-canal-treated teeth: a retrospective study over 10 years. *Journal of Endodontics* 29:638-643.
- Er O, Yaman SD, Hasan M (2007) Finite element analysis of the effects of thermal obturation in maxillary canine teeth. *Oral Medicine, Oral Pathology, Oral Radiology and Endodontics* 104:277-286.
- Eskitascioglu G, Belli S, Kalkan M (2002) Evaluation of Two Post Core Systems Using Two Different Methods (Fracture Strength Test and a Finite Elemental Stress Analysis). *Journal of Endodontics* 28:629-633.
- Fuss Z, Lustig J, Tamse A (1999) Prevalence of vertical root fractures in extracted endodontically treated teeth. *International Endodontic Journal* 32:283-286.
- Gimlin DR, Parr CH, Aguirre-Ramirez GA (1986) comparison of stresses produced during lateral and vertical condensation using engineering models. *Journal of Endodontics* 12:235-240.

- Goto Y, Nicholls JL, Phillips KM, Junge T (2005) Fatigue resistance of endodontically treated teeth restored with three dowel-and-core systems. *Journal of Prosthetic Dentistry*. 93:45-50.
- Harvey TE, White JT, Leeb JJ (1981) Lateral condensation stress in root canals. *Journal of Endodontics* 7:151-155.
- Hayashi M, Takahashi Y, Imazato S, Ebisu S (2006) Fracture resistance of pulpless teeth restored with post-cores and crowns. *Dental Materials* 22:477-485.
- Hayashi M, Sugeta A, Takahashi Y, Imazato S, Ebisu S (2008) Static and fatigue fracture resistances of pulpless teeth restored with post-cores. *Dental Materials* 24:1178-1186.
- Heydecke G, Butz F, Strub JR (2001) Fracture strength and survival rate of endodontically treated maxillary incisors with approximal cavities after restoration with different post and core systems: an in-vitro study. *Journal of Dentistry* 29:427-433.
- Heydecke G, Butz F, Hussein A, Strub JR (2002) Fracture strength after dynamic loading of endodontically treated teeth restored with different post-and-core systems. *Journal of Prosthetic Dentistry* 87:438-445.
- Hong J, Xia WW, Xiong HG (2002) Analysis the Effect on the Stress of Wall of Root Canal by Different Proportion between Crown and Root. *Shanghai Journal of Stomatology* 11:213-215.
- Hong J, Xia WW, Zhu YQ (2003) Analysis the Effect on the Stress of Root Canal Wall of Pulpless Teeth by Different Degree of Root Canal Preparation. *Shanghai Journal of Stomatology* 12:24-26.
- Honlcomb JQ, Pitts DL, Nicholls JL (1987) Further investigation of spreader loads required to cause vertical root fracture during lateral condensation. *Journal of endodontics* 13:277-284.
- Huang HM, Ou KL, Wang WN, Chiu WT, Lin CT, Lee SY (2005) Dynamic finite element analysis of the human maxillary incisor under impact loading in various directions. *Journal of Endodontics* 31:723-7.
- Johnson JK, Schwartz NL, Blackwell RT (1976) Evaluation and restoration of endodontically treated posterior teeth. *Journal of the American Dental Association* 93: 597-605.
- Kampoosiora P, Papavasiliou G, Bayne SC, Felton DA (1994) Finite element analysis estimates of cement microfracture under complete veneer crown. *Journal of Prosthetic Dentistry* 71: 435-441.
- Kawai K, Masaka N (2002) Vertical root fracture treated by bonding fragments and rotational replantation. *Dental Traumatology* 18:42-45.
- Kim HC, Cheung GS, Lee CJ, Kim BM, Park JK, Kang SI (2008) Comparison of Forces Generated During Root Canal Shaping and Residual Stresses of Three Nickel-Titanium Rotary Files by Using a Three-Dimensional Finite-element Analysis. *Journal of Endodontics* 34:743-747.
- Kim HC, Kim HJ, Lee CJ, Kim BM, Park JK, Versluis A (2009) Mechanical response of nickel-titanium instruments with different cross-sectional designs during shaping of simulated curved canals. *International Endodontic Journal* 42:593-602.
- Kim HC, Lee MH, Yum J, Versluis A, Lee CJ, Kim BM (2010) Potential Relationship between Design of Nickel-Titanium Rotary Instruments and Vertical Root Fracture . *Journal of Endodontics* (ARTICLE IN PRESS)

- Kim TO, Kim BM, Park JK, Hur B, Kim HC (2008) Stress distribution of three NiTi rotary files under bending and torsional conditions using a mathematic analysis. *The Journal of Korean Academy of Conservative Dentistry* 33:323-331.
- Kim TO, Cheung GS, Lee JM, Kim BM, Hur B, Kim HC (2009) Stress distribution of three NiTi rotary files under bending and torsional conditions using a mathematic analysis. *International Endodontic Journal* 42:14-21.
- Kinney JH, Balooch M, Marshall SJ, Marshall GW, Weihs TP (1996) Hardness and Young's modulus of human peritubular and intertubular dentine. *Archives of Oral Biology* 41: 9-13.
- Lakes RS (2002) Composite Biomaterials. In: Park JB, Bronzino JD, ed. *Biomaterials Principles and Applications*, 1st edn, Boca Raton: CRS Press, pp.79-93.
- Lertchirakarn V, Palamara JE, Messer HH (2003a) Patterns of vertical root fracture: factors affecting stress distribution in the root canal. *Journal of Endodontics* 29:523-528
- Lertchirakarn V, Joseph E (2003b) Finite Element Analysis and Strain-gauge Studies of Vertical Root Fracture. *Journal of Endodontics* 29:529-534.
- Lin CL, Chang YH, Chang CY, Pai CA, Huang SF (2010) Finite element and Weibull analyses to estimate failure risks in the ceramic endocrown and classical crown for endodontically treated maxillary premolar. *European Journal of Oral Sciences* 118:87-93.
- Lindauer PA, Campbell AD, Hicks ML, Pelleu GB (1989) Vertical root fractures in curved roots under simulated clinical conditions. *Journal of Endodontics* 15:345-349.
- Matthews LS, Hirsch C (1972) Temperatures measured in human cortical bone when drilling. *The Journal of Bone and Joint Surgery-American* 54:297-308.
- Meister F, Lommel TJ, Gerstein H (1980) Diagnosis and possible causes of vertical root fractures. *Oral Surgery, Oral Medicine, and Oral Pathology* 49:243-253.
- Mireku AS, Romberg E, Fouad AF, Arola D (2010) Vertical fracture of root filled teeth restored with posts: the effects of patient age and dentine thickness. *International Endodontic Journal* 43:218-225.
- Necchi S, Taschieri S, Petrini L, Migliavacca F (2008) Mechanical behaviour of nickel-titanium rotary endodontic instruments in simulated clinical conditions: a computational study. *International Endodontic Journal* 41:939-949.
- Ou KL, Chang CC, Chang WJ, Lin CT, Chang KJ, Huang HM (2009) Effect of damping properties on fracture resistance of root filled premolar teeth: a dynamic finite element analysis. *International Endodontic Journal* 42:694-704.
- Ozcan M, Valandro LF (2009) Fracture Strength of Endodontically-treated Teeth Restored with Post and Cores and Composite Cores Only. *Operative Dentistry* 34:429-436.
- Pashley D, Okabe A, Parham P (1985) The relationship between dentin microhardness and tubule density. *Dental Traumatology* 1: 176-179.
- Pitts DL, Matheny HE, Nicholls JI (1983) An in vitro study of spreader loads required to cause vertical root fracture during lateral condensation. *Journal of Endodontics* 9:544-550.
- Pontius O, Hutter JW (2002) Survival Rate and Fracture Strength of Incisors Restored with Different Post and Core Systems and Endodontically Treated Incisors without Coronoradicular Reinforcement. *Journal of Endodontics* 28:710-715.

- Qing H, Zhu ZM, Chao YL, Zhang WQ (2007) In vitro evaluation of the fracture resistance of anterior endodontically treated teeth restored with glassfiber and zircon posts. *Journal of Prosthetic Dentistry* 97:93-98.
- Reinhardt RA, Krejci RF, Pao YC, Stannard JG (1983) Dentin stress in posts reconstructed with diminishing bone support. *Journal of Dental Research* 62:1002-1008.
- Ricks-Williamson LJ, Fotos PG, Goel VK, Spivey JD, Rivera EM, Khera SC (1995) A three-dimensional finite-element stress analysis of an endodontically prepared maxillary central incisor. *Journal of Endodontics* 21:362-327.
- Rundquist BD, Versluis A (2006) How does canal taper affect root stresses? *International Endodontic Journal* 39:226-237.
- Salis SG, Hood JAA, Stokes ANS, Kirk EEJ (2006) Patterns of indirect fracture in intact and restored human premolar teeth. *Dental Traumatology* 3:10-14.
- Santos AFV, Tanaka CB, Lima RG, Esposito COM, Ballester RY, Braga RR, Meira JBC (2009) Vertical Root Fracture in Upper Premolars with Endodontic Posts: Finite Element Analysis. *Journal of Endodontics* 35: 117-120.
- Sathorn C, Palamara JEA, Messer HH (2005a) A comparison of the effects of two canal preparation techniques on root fracture susceptibility and fracture pattern. *Journal of Endodontics* 31:283-287
- Sathorn C, Joseph EA, Harold H (2005b) Effect of Root Canal Size and External Root Surface Morphology on Fracture Susceptibility and Pattern: A Finite Element Analysis. *Journal of Endodontics* 31:288-292.
- Silver-Thorn MB, Joyce TP (1999) Finite Element Analysis of Anterior Tooth Root Stresses Developed During Endodontic Treatment. *Journal of Biomechanical Engineering; Transactions of the ASME* 121:108-115.
- Steele A, Johnson BR (1999) In vitro fracture strength of endodontically treated premolars. *Journal of Endodontics* 25:6-8.
- Telli C, Gülkan P (1998) Stress analysis during root canal filling by vertical and lateral condensation procedures: three-dimensional finite element model of a maxillary canine tooth. *British Dental Journal* 185:79-86.
- Telli C, Gülkan P, Raab W (1999) Additional studies on the distribution of stresses during vertical compaction of gutta-percha in the root canal. *British Dental Journal* 187:33-37.
- Trope M, Rosenberg ES (1992) Multidisciplinary approach to the repair of vertically fractured teeth. *Journal of Endodontics* 18:460-463.
- Unsal E, Eskitasiloglu G, Soykan E, Walsh TF (2002) Finite element analysis of forces created by root separation and resection modeling. *Journal of Oral Sciences* 44:79-84.
- Versluis A, Messer HH, Pintado MR (2006) Changes in compaction stress distributions in roots resulting from canal preparation. *International Endodontic Journal* 39:931-939.
- Walton RE (2002) Longitudinal Tooth Fractures. In RE Walton, M Torabinejadeds. *Principles and practice of endodontics*, 3th edn; pp:513-518. Philadelphia: Saunders.
- Wilcox LR, Roskelley C, Sutton T (1997) The relationship of root canal enlargement to finger-spreader induced vertical root fracture. *Journal of Endodontics* 23:533-534.
- Yang JB, Liu TJ, Li JY (2004) Initial clinic research on curved canal preparation by reverse flaring technique. *West China Journal of Stomatology* 22:123-5.

Finite element simulation. Applications in Orthopaedic Surgery and Traumatology

Antonio Herrera^{a,c}, Luis Gracia^b, Elena Ibarz^b, Juan J. Panisello^{a,c}, José Cegoñino^b, Jesús Mateo^{a,c}, Javier Rodríguez-Vela^{a,c} and Sergio Puértolas^b

^a*Medicine School. University of Zaragoza (Spain)*

^b*Engineering Faculty. University of Zaragoza (Spain)*

^c*Miguel Servet University Hospital, Zaragoza (Spain)*

1. Introduction

Research in different fields concerning Orthopaedic Surgery and Traumatology requires a methodology that allows, at the same time, a more economic approach and the possibility of reproducing in a easy way different situations. Such a method could be used as a guide for research on biomechanics of the locomotor system, both in healthy and pathologic conditions, along with the study of performance of different prostheses and implants. To that effect, the use of simulation models, introduced in the field of Bioengineering in recent years, can undoubtedly mean an essential tool to assess the best clinical option, provided that it will be accurate enough in the analysis of specific physiological conditions concerning a certain pathology.

The Finite Element Method (FEM) was originally developed for solving structural analysis problems relating to Mechanics, Civil and Aeronautical Engineering. The paternity of this method is attributed to Turner, who published his first, historic, job in 1956 (Turner et al., 1956). In 1967 Zienkiewicz OC published the book "The finite element method in structural and continuum mechanics" (Zienkiewicz, 1967) which laid down mathematical basis of the method. Other fundamental contributions to the development of Finite Element Method (FEM) took place on dates nearest (Imbert, 1979; Bathe, 1982; Zienkiewicz & Morgan, 1983; Hughes, 1987).

Finite element (FE) simulation has proved to be specially suitable in the study of the behaviour of any physiological unit, despite its complexity. Nowadays, it has become a powerful tool in the field of Orthopaedic Surgery and Traumatology, helping the surgeons to have a better understanding of the biomechanics, both in healthy and pathological conditions. FE simulation let us know the biomechanical changes that occur after prosthesis or osteosynthesis implantation, and biological responses of bone to biomechanical changes. It also has an additional advantage in predicting the changes in the stress distribution around the implanted zones, allowing to prevent future pathologies derived from an unsuitable positioning of the prostheses or its fixation. Simulation also allows us to predict the behavior of orthopedic splints, utilized for the correction of deformities, providing the

recovering force-displacement and angle-moment curves that characterize the mechanical behavior of the splint in the overall range of movement.

2. Methodology for the finite element analysis of biomechanical systems

One of the most significant aspects of biomechanical systems is its geometric complexity, which greatly complicates the generation of accurate simulation models. Classic models just suffered from this lack of geometrical precision, present even in recent models (Guan et al., 2006; Little et al., 2007), which challenged, in most studies, the validity of the results and their extrapolation to clinical settings.

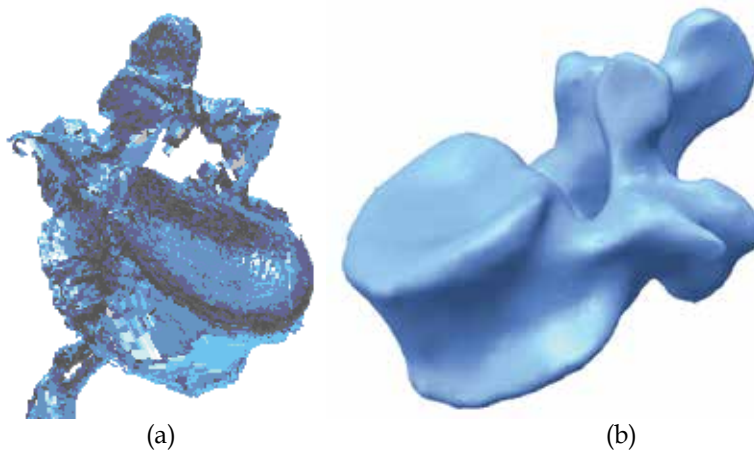


Fig. 1. 3D scanning of a vertebra: a) Original without processing; b) Final after processing



Fig. 2. 3D laser scanner



Fig. 3. Anatomical model of the lumbar spine

Currently, there are methodologies developed over recent years that avoid such problems, allowing the generation of models with the desired precision in a reasonable time and cost is not excessive. Thus, the use of scanners together with three-dimensional images obtained by CT allow making geometric models that combine high accuracy in the external form with an excellent definition of internal interfaces. The method requires not only appropriate software tools, capable of processing images, but also its compatibility with the programs used later to generate the finite element model. For example, in Fig. 1a is shown the initial result obtained by a three-dimensional laser scanner Roland Picza (Fig. 2), from an anatomical model of the lumbar spine Somso brand QS-15 (Fig. 3).

After processing by Dr. Picza 3 and 3D Editor programs, we get the final result in Fig. 1b, which shows the geometric precision obtained. In these models, the characterization of the internal structure is made by 3D CT, from images like that shown in Fig. 4. An alternative to the above procedure is the use of 3D geometrical reconstruction programs, for example, MIMICS (Mimics, 2010). In any case, the final result is a precise geometrical model which serves as a basis for the generation of a finite elements mesh.

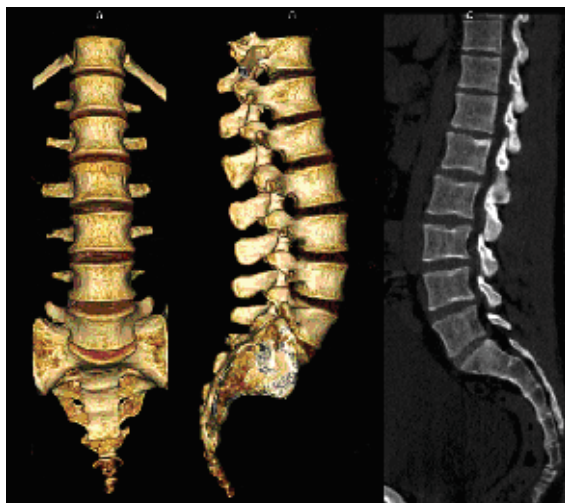


Fig. 4. Volume rendering 3D reconstruction and sagittal multiplanar reconstruction of the lumbar spine

In view of the difficulties experienced in living subjects, FE simulation models have been developed to carry out researches on biomechanical systems with high reproducibility and versatility. These models allow to repeat the study as many times as desired, being a non-aggressive investigation of modified starting conditions. However, work continues on the achievement of increasingly realistic models that allow to put the generated results and predictions into a clinical setting. To that purpose it is mainly necessary the use of meshes suitable for the particular problem, as regards both the type of elements and its size. It is always recommended to perform a sensitivity analysis of the mesh to determine the optimal features or, alternatively, the minimum necessary to achieve the required accuracy. In Fig. 5 is shown a FE mesh of a lumbar vertebral body, using tetrahedron type elements. It can be seen that the element size allows to depict, with little error, the geometry of the vertebra, compared with Fig. 1b.

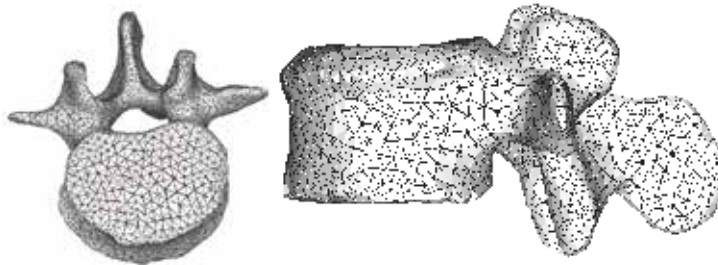


Fig. 5. Meshing of a lumbar vertebra



Fig. 6. Meshing of proximal fémur with stem

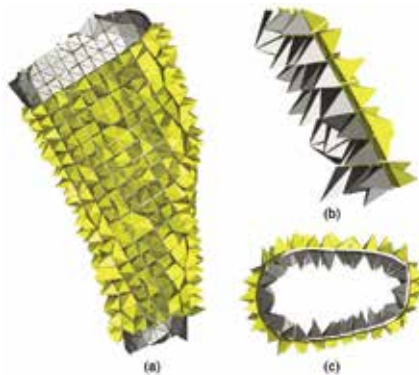


Fig. 7. Contact interface femur-stem

A key issue in FE models is the interaction between the different constitutive elements of the biomechanical system, especially when it comes to conditions which are essential in the behaviour to be analyzed. Thus, in Fig. 6 is shown a FE mesh of the proximal femur, with a cementless stem in place. The biomechanical behaviour of this type of implant depends basically on the conditions of contact between the stem and bone, so that the correct simulation of the latter determines the validity of the model. In Fig 7 can be seen the stem-femur contact interface, defined by the respective surfaces and the frictional conditions needed to produce the press-fit which is achieved at surgery.

Finally, in FE simulation models is essential the appropriate characterization of the mechanical behaviour of the different materials, usually very complex. So, the bone exhibits an anisotropic behaviour with different responses in tension and compression (Fig. 8). Moreover, it varies depending on the bone type (cortical or cancellous) and even along different zones in the same specimen, as in the vertebrae (Denozière & Ku, 2006). This kind of behaviour is reproducible in a reliable way in the simulation, but it leads to an excessive computational cost in global models. For this reason, in most cases, and specially in long bones, a linear elastic behaviour in the operation range concerning strains and stresses is considered.

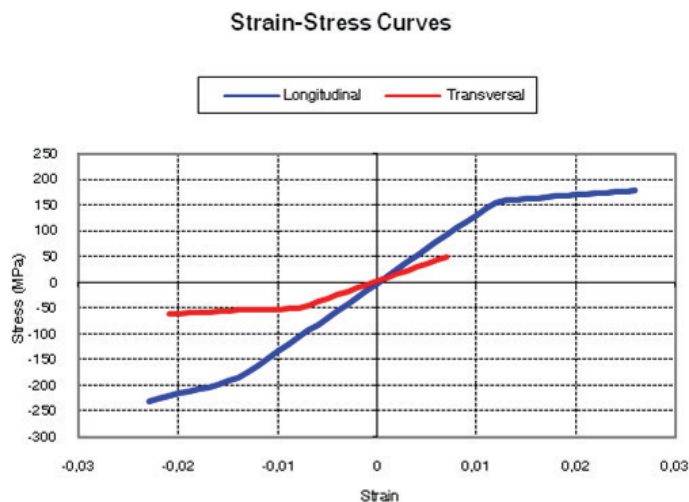


Fig. 8. Strain-stress curves for cortical bone

In soft tissues, the behaviour is even more complex, usually as a hyperelastic material. This is the case of ligaments (Fig. 9), cartilages and muscles, also including a reologic effect with deferred strains when the load conditions are maintained (viscoelastic behaviour). A special case arises in the intervertebral discs, where nucleus and annulus present totally different features: while the nucleus behaves as an incompressible fluid, the annulus could be considered as a two-phase material with a flexible matrix and a set of fibers with only tension hyperelastic behaviour.

This inherent complexity to the different biological tissues, reproducible in reduced or local models, is very difficult to be considered in global models as the used to analyse prostheses and implants, because the great amount of non-linearities do the convergence practically

unfeasible. On the other hand, it leads to a prohibitive computational cost, only possible to undertake by supercomputers.

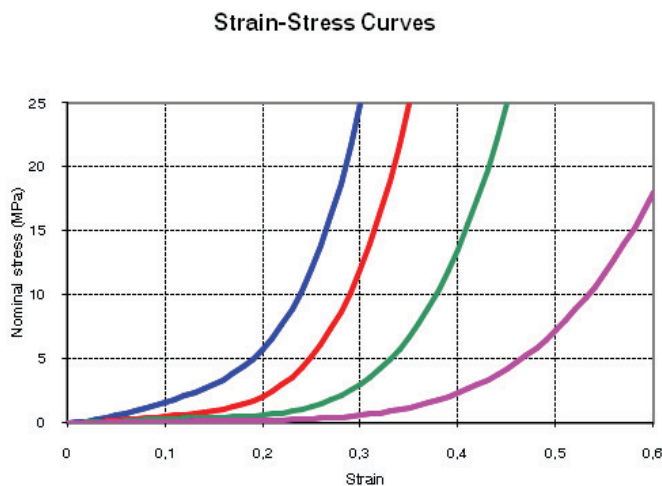


Fig. 9. Strain-stress curves for vertebral ligaments

3. Application to the behaviour of hip prostheses

Bone is living tissue that undergoes a constant process of replacement of its structure, characterized by bone reabsorption and new bone formation, without changing their morphology. This process is called bone remodeling. On the other hand, bone adapts its structure, according to Wolff's Law, to the forces and biomechanical loads that receives (Buckwalter et al., 1995). In a normal hip joint, loads from the body are transmitted to the femoral head, then to the medial cortical bone of femoral neck towards the lesser trochanter, where they are distributed by the diaphyseal bone (Radin, 1980).

The implantation of a cemented or cementless femoral stem produced a clear alteration of the physiological transmission of loads, as these are now passed through the prosthetic stem, in a centripetal way, from the central marrow cavity to the cortical bone (Marklof et al., 1980). These changes of the normal biomechanics of the hip bone leads to a phenomenon called adaptive remodeling (Huiskes et al., 1989), since bone has to adapt to the new biomechanical situation. Remodeling is a multifactorial process depending on both mechanical and biological factors. Mechanical factors are related to the new distribution of loads caused by implantation of the prosthesis in the femur, the physical characteristics of the implant (size, implant design and alloy), and the type of anchoring in the femur: metaphyseal, diaphyseal, hybrid, etc. (Summer & Galante, 1992; Sychter & Engh, 1996; Rubash et al., 1998; McAuley et al., 2000; Gibson et al., 2001; Glassman et al., 2001). Biologics are related to age and weight of the individual, initial bone mass, quality of primary fixation and loads applied to the implant. Of these biological factors, the most important is initial bone mass (Sychter & Engh, 1996).

Different models of cementless stems have tried to achieve perfect load transfer to the femur, mimicking the physiological transmission from the femoral calcar to the femoral shaft. The main objective was to avoid stress-shielding, since in absence of physiological transmission of loads, and lack of mechanical stimulus in this area, causes a proximal bone atrophy.

Cemented stem fixation is achieved by the introduction of cement into bone, forming a bone-cement interface. Inside the cement mantle a new interface is made up between cement and stem. It might seem that the cement mantle enables better load distribution in the femur; however the design, material and surface of prostheses, play an important role in transmission and distribution of charges, influencing bone remodelling (Ramaniraka et al., 2000; Li et al., 2007)

Long-term follow-up of different models of cementless stems have shown that this is not achieved, and to a greater or lesser extent the phenomenon of stress-shielding is present in all the models, and therefore the proximal bone atrophy. It is interesting to know, in cemented stems, not only the stress-shielding and subsequent proximal bone atrophy, but also the long-term behavior of cement-bone and stem-cement interfaces. This requires long-term studies monitoring the different models of stems.

FE simulation allows us to study the long term biomechanical behavior of any type of stem cemented or cementless, and predict the impact of biomechanics on the femur, with its consequent effects on bone remodelling. So, we have developed FE models to study the biomechanical behavior of cemented and cementless stems. Our models have been validated with long-term DEXA studies of patients who were treated with different types of femoral stems (Herrera et al., 2007; Herrera et al., 2009).

The development of the model of a healthy femur is crucial to make accurate the whole process of simulation, and to obtain reliable results. A femur from a 60 years old man, died in traffic accident, has been used to build the model. Firstly, each of the parts necessary to set the final model were scanned using a three dimensional scanner Roland Picza brand. As a result, we get a cloud of points which approximates the scanned geometry. These surfaces must be processed through the programs Dr.Picza-3 and 3D-Editor. This will eliminate the noise and performs smooth surfaces, resulting in a geometry that reliably approximates to the actual geometry.



Fig. 10. a) FE model of healthy femur, b) Coronal section of healthy femur model

CT scans and 3D-CT reconstructions were taken from the femur to determine the geometry of the cancellous bone, allowing a perfect model of this part of the bone. For a precise geometry, splines are plotted according to the tomograms and then the cancellous surface is

modelled. This area represents the separation between cortical and cancellous bone. The meshing is performed by using I-deas program (I-deas, 2007), which creates two groups of elements (cancellous and cortical). Taking cancellous bone elements as start point, a third group of elements is selected by applying the properties of bone marrow. The mesh is based on tetrahedral solid elements with linear approximation, obtaining a total of 408,518 elements (230,355 elements for cortical bone, 166,220 elements for cancellous bone and 11,943 elements for bone marrow) (Fig. 10).

Different publications (Evans, 1973) were consulted to obtain the properties of bone material. Table 1 summarizes the mechanical properties values used in biological materials, which have been simplified to consider bone as isotropic and linear elastic material.

MATERIAL	ELASTIC MODULUS (MPa)	POISSON RATIO	MAXIMUM COMPRESSION STRESS (MPa)	MAXIMUM TENSION STRESS (MPa)
CORTICAL BONE	20000	0,3	150	90
CANCELLOUS BONE	959	0,12	23	
BONE MARROW	1	0,3		

Table 1. Mechanical properties of the healthy femur model

The main features of each of the boundary conditions are:

1.- Clamped in the middle of the femoral shaft

The middle zone has been clamped instead of distal zone because middle zone is considered enough away from proximal bone (Fig. 11). This model can be compared with other that have been clamped at a distal point, since the loads applied practically coincided with the femoral axis direction thus reducing the differences in final values.

2.- Hip muscles Loads

Forces generated by the abductor muscles are applied on the greater trochanter, in agreement with most authors' opinion (Weinans et al., 1994; Kerner et al., 1999). Generally, muscle strength generated in the hip joint is 2 times the body weight, and this produces a reaction strength in the femoral head that accounts for 2.75 times the body weight. However, when the heel impacts to the ground, and in double support stage of the gait, the load increases up to 4 times the body weight. The latter case, being the worst one, has been considered to impose the boundary conditions. It has also been considered a body weight of 79.3 kg for cementless stems, and 73 kg for cemented stems. Those were the average values obtained from the clinical sample to be contrasted with the simulation results. The load due to the abductor muscles, accounting for 2 times the corporal weight, is applied to the proximal area of the greater trochanter, at an angle of 21 degrees, as shown in Fig. 11.

3.- Reaction strength on the femoral head due to the body weight.

As already mentioned, we have studied the case of a person to 79.3 kg in cementless stems, and 73 kg in cemented stems, in the worst case of double support or heel impact stages of

the gait. The resultant force on the femoral head would be worth 4 times the body weight (Fig. 11).

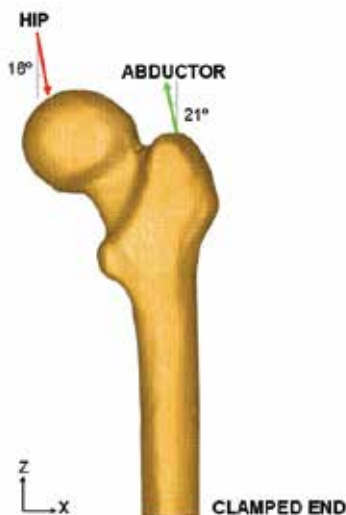


Fig. 11. Boundary conditions applied in the healthy femur model

The models generated using I-deas are calculated by means of Abaqus 6.7, and the postprocessing of the results was performed by Abaqus Viewer (Abaqus, 2009).

To generate the models from different stems, these were scanned to obtain its geometry. We studied two cementless stems. Both of them were anatomically shaped, metaphyseal anchored and coated with HA in their metaphyseal zone (ABG I and ABG II). The size used has been similar in both models, however the alloy, geometry, length, thickness and distal diameter were different. For cemented stems models, we chose the cemented anatomical stem ABG, and the Versys straight, polished stem. After obtaining the geometry of the different stems, several cadaver femurs were operated on in order to implant each of the prostheses, in the same way as one would carry out a real hip replacement.

Those operated femurs were scanned a second time to use them as a reference in the positioning of the prostheses. We employ, for every model, three meshes generated by the I-deas program: healthy femur, femoral stem, and operated femur. The mesh of healthy femur and the mesh of the operated femur were superimposed, then the healthy femoral epiphysis was removed in a identical way as it is done during surgery, so as to insert the prosthesis. Afterward the stem was positioned in the femur, always taking the superimposed mesh of the operated femur as a base (Fig. 12).

In the case of cementless stems, the previous process for modeling the cadaveric femur was repeated only for cortical bone. While the cancellous bone was modeled again in such a way that it fitted perfectly to contact with the prostheses. The Abaqus 6.7 program was utilized for calculation and simulation of the previously generated models, and the Abaqus Viewer was used for viewing the results. Union between the stems and the cancellous bone were not considered, but contact conditions were defined with a constant 0.5 friction coefficient, simulating the perfect press-fit setting. The final model with ABG-I stem comprises a total of 60401 elements (33504 for cortical bone, 22088 for cancellous bone and 4809 for ABG-I stem). The final model with the ABG-II stem is made up of 63784 elements (33504 for cortical bone,

22730 for cancellous bone and 7550 for ABG-II stem). Fig. 13 shows both FE models obtained for cementless prostheses.

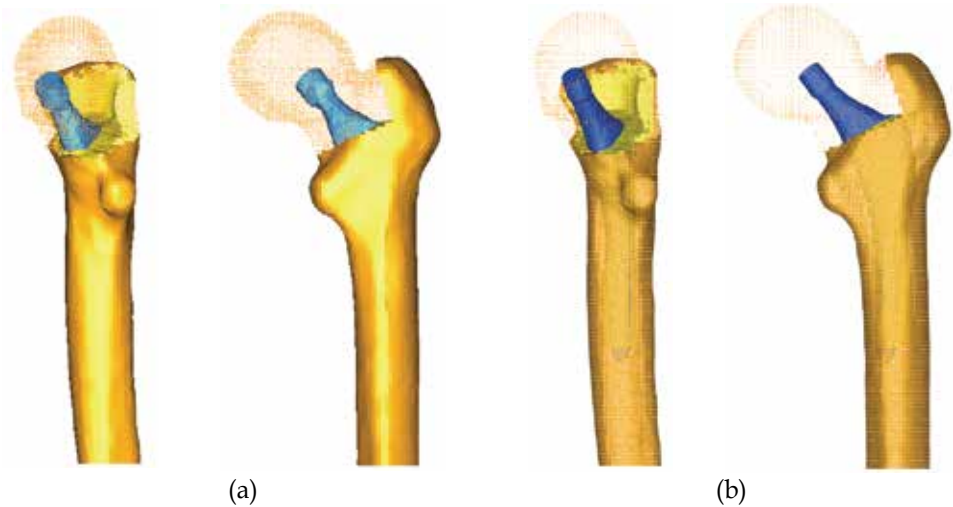


Fig. 12. Removal of the femoral head and positioning of the cementless stems: (a) ABG-I and (b) ABG-II.

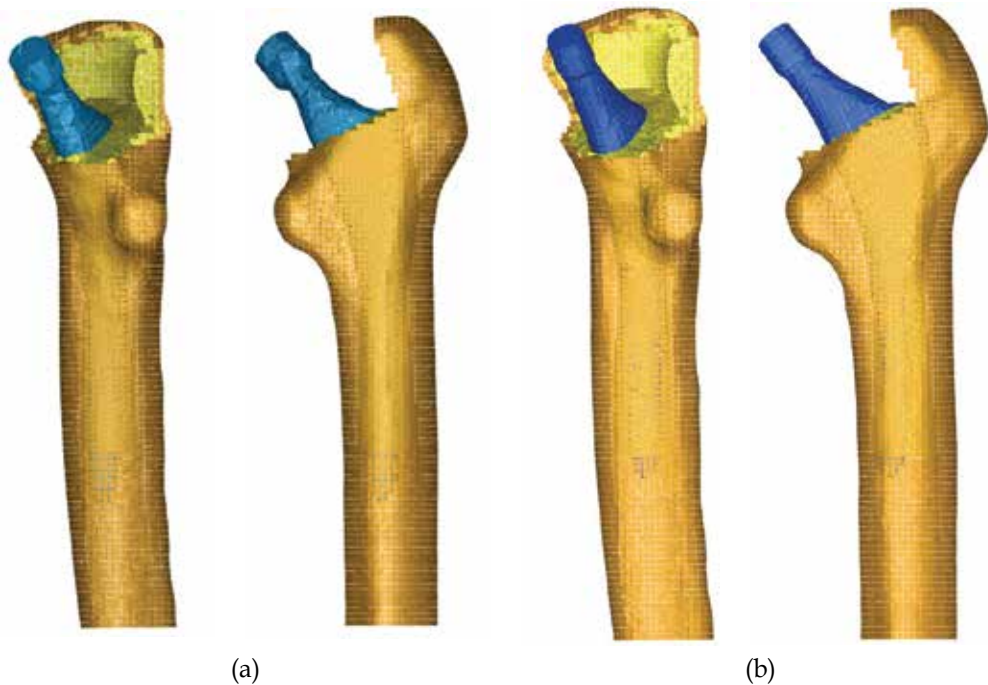


Fig. 13. FE models of the femur with cementless prosthesis: (a) ABG-I and (b) ABG-II

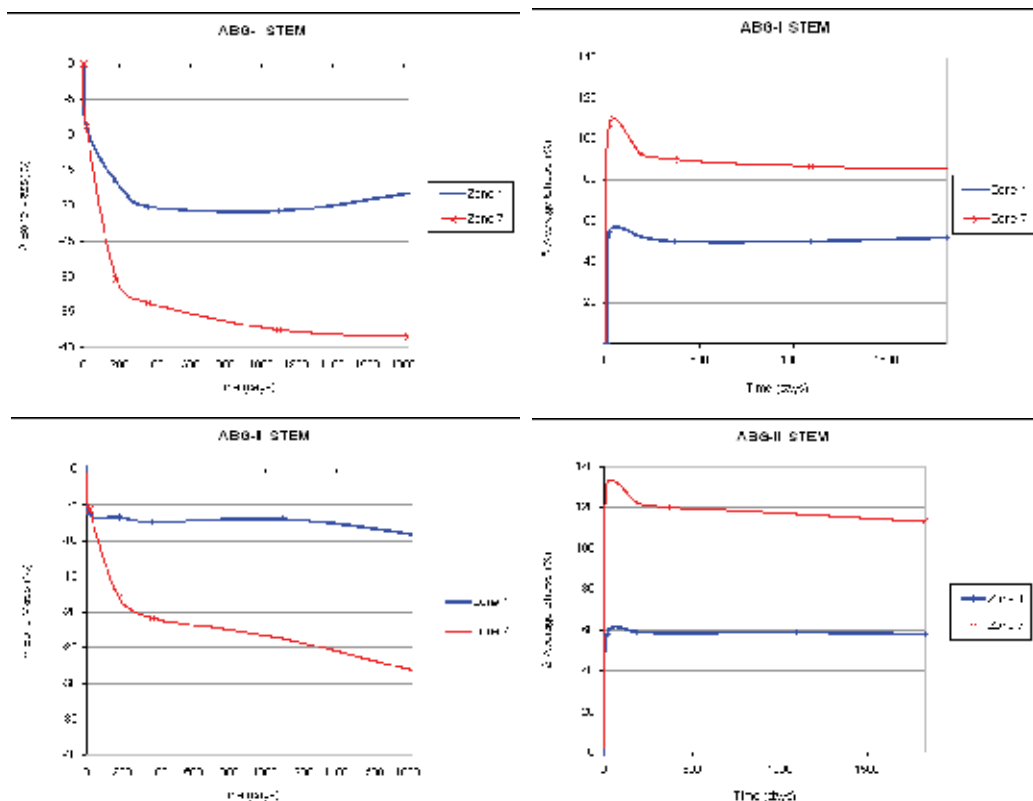


Fig. 14. BMD and average von Mises stress evolution

Calculation was performed using the program Abaqus 6.7. Both prostheses have been simulated with the same mechanical properties, thus, the result shows the influence of stem geometry on the biomechanical behavior. Fig. 14 shows, for each cementless stem model, the change (%) in BMD and average von Mises stress corresponding to the Gruen proximal zones (1 and 7), being the most representative concerning stress-shielding, and taking as reference the pre-operative data. It can be confirmed, for both stems, that the maximum decreasing in BMD is achieved in zone 7. This reduction in BMD is bigger in the ABG-I than in ABG-II stem.

In the case of cemented stems the process of modelling was similar, varying the surgical cut in the femoral neck of the healthy femur. Each stem was positioned into the femur, always taking the superimposed mesh of the operated femur as a base (Fig. 15). In the previous process of modelling, on the cadaveric femur, only the cortical bone was used. The cancellous bone was modelled again taking into account the cement mantle surrounding the prosthesis and the model of stem (ABG or Versys), so as to obtain a perfect union between cement and cancellous bone. The cement mantle was given a similar thickness, in mm, which corresponds to that usually achieved in patients operated on, different for each of the stem models studied and each of the prosthesis, so that the simulation model be as accurate as possible.

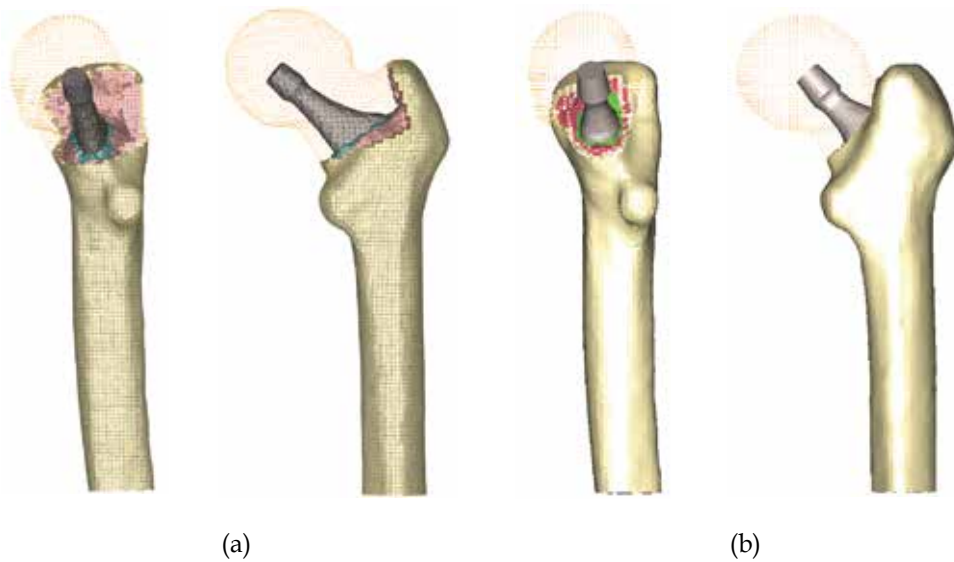


Fig. 15. Removal of the femoral head and cemented prosthesis positioning: (a) ABG-cemented and (b) Versys

In models of cemented prostheses it is not necessary to define contact conditions between the cancellous bone and the stem. In this type of prosthesis the junction between these two elements is achieved by cement, which in the EF model should simulate conditions of perfect union between cancellous bone-cement and cement-stem. It has also been necessary to model the diaphyseal plug that is placed in actual operations to prevent the spread of the cement down to femoral medullary canal. Fig. 16 shows the longitudinal sections of the final models.

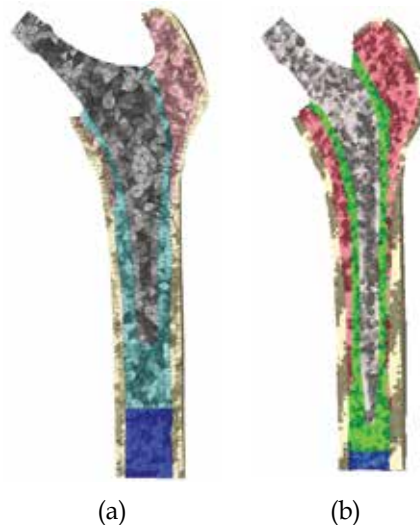


Fig. 16. Longitudinal section of the FE models with cemented femoral prostheses: (a) ABG-cemented and (b) Versys.

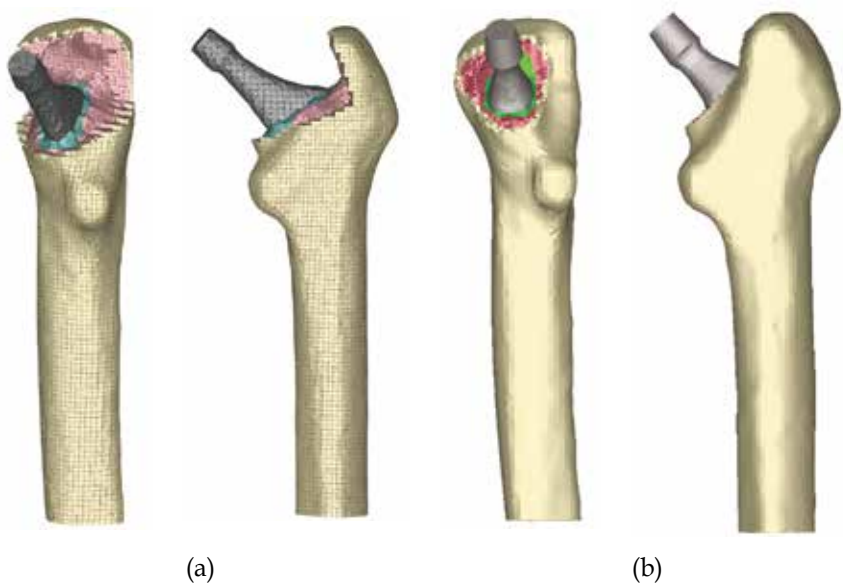


Fig. 17. FE model with cemented femoral prostheses: (a) ABG-cemented and (b) Versys.

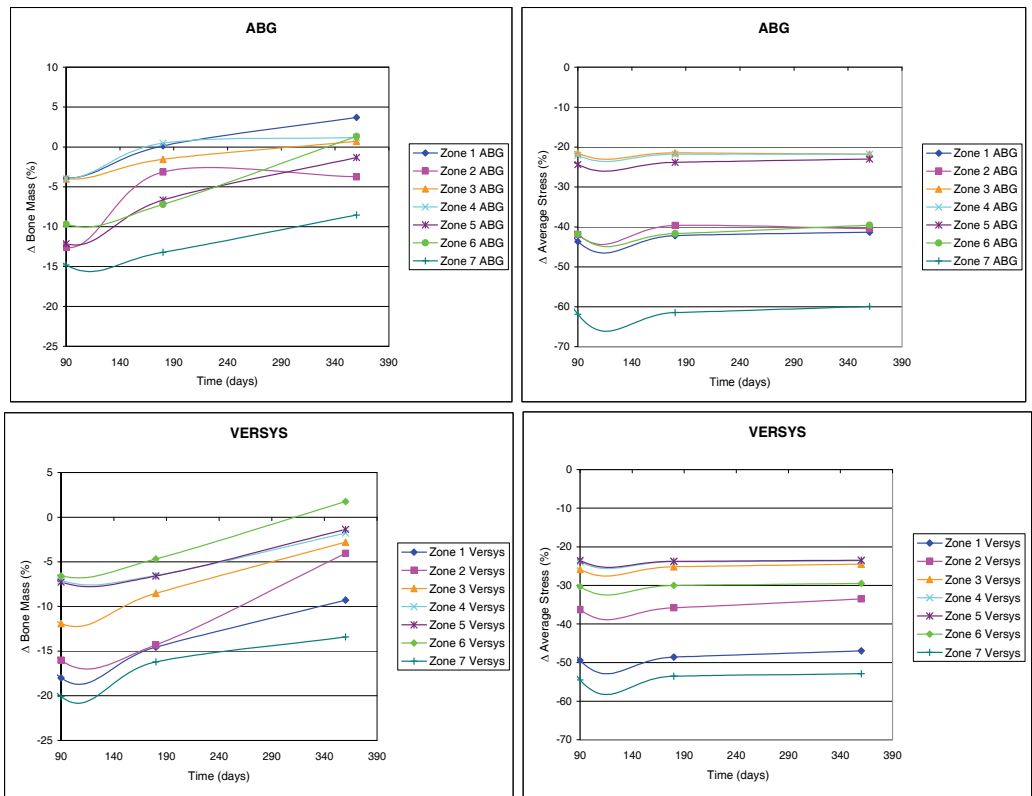


Fig. 18. BMD and average von Mises stress evolution

Both models were meshed with tetrahedral solid elements linear type, with a total of 74192 elements in the model for ABG-cemented prosthesis (33504 items cortical bone, cancellous bone 17859, 6111 for the ABG stem-cement, cement 13788 and 2930 for diaphyseal plug), and 274651 in the model for prosthetic Versys (119151 items of cortical bone, cancellous bone 84836, 22665 for the Versys stem, 44661 for the cement and 3338 in diaphyseal plug). In Fig. 17 are shown both models for cemented stems.

Calculation was performed using the program Abaqus 6.7. Both prostheses have been simulated with the same mechanical properties, thus, the result shows the influence of stem geometry on the biomechanical behavior. Fig. 18 shows the variation (%) of bone mass and average von Mises stress (%) in each of the Gruen zones for each of the models of cemented prostheses, with reference to the preoperative time. It can be seen that for both stems, the maximum decrease in bone mass occurred in Zone 7. This decrease in bone mass is greater in the Versys model than in the ABG stem

Prior to the development of our FE models several long-term studies of bone remodeling after the implantation of two different cementless stems, ABG I and ABG II, were performed (Panisello et al., 2006; Panisello et al., 2009a; Panisello et al., 2009b). These studies were performed using DEXA, a technique that allows an accurate assessment of bone density losses in the different Gruen zones (Fig. 19). We take as a reference to explore this evolution, the postoperative value obtained in control measurements and those obtained from contralateral healthy hip. New measurements were made at 6 months, one year and 5 years after surgery. The ABG II stem is an evolution of the ABG I, which has been modified both in its alloy and design. The second generation prosthesis ABG-II is manufactured with a different titanium alloy from that used in the ABG-I. The prosthetic ABG-I stem is made with a Wrought Titanium alloy (Ti 6Al-4V) of which elasticity modulus is 110 GPa. Meanwhile, the TMZF alloy, which is used on the ABG-II stem, has a Young's modulus of 74-85 GPa, according to the manufacturer information, using a mean value of 79.5 GPa in the different analyses. On the other hand, the ABG II stem has a new design with less proximal and distal diameter, less length and the shoulder of stem has been redesign to improve osteointegration in the metaphyseal area.

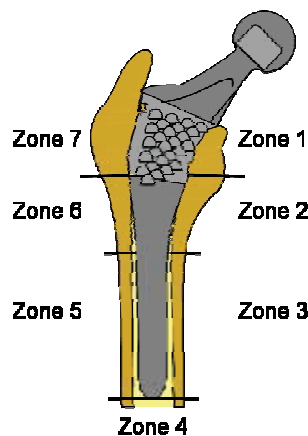


Fig. 19. Gruen zones

In our DEXA studies, directed to know the loss of bone mass in the different zones of Gruen caused by the stress-shielding, we found that ABG II model produces less proximal bone

atrophy in post-operative measurements, for similar follow-up periods. In the model ABG II, we keep finding in studies with DEXA a proximal bone atrophy, mainly in zones 1 and 7 of Gruen, but with an improvement of 8.7% in the values obtained in ABG I series. We can infer that improvements in the design of the stem, with a narrower diameter in the metaphyseal area, improve the load transfer to the femur and therefore minimizes the stress-shielding phenomenon, resulting in a lower proximal bone atrophy, because this area receive higher mechanical stimuli. These studies for determination of bone mass in Gruen zones, and the comparative study of their postoperative evolution during 5 years have allowed us to draw a number of conclusions: A) Bone remodeling, after implantation of a femoral stem, is finished one year after surgery B) Variations in bone mass, after the first year, are not significant.

The importance of these studies is that objective data from a study with a series of patients, allow us to confirm the existence of stress-shielding phenomenon, and quantify exactly the proximal bone atrophy that occurs. At the same time they have allowed us to confirm that the improvements in the ABG stem design, mean in practice better load transfer and less stress-shielding phenomenon when using the ABG II stem.

DEXA studies have been basic to validate our FE models, because we have handled real values of patients' bone density, which allowed us to measure mechanical properties of real bone in different stages. Through computer simulation with our model, we have confirmed the decrease of mechanical stimulus in femoral metaphyseal areas, having a higher stimulus in ABG II type stem, which corresponds exactly with the data obtained in studies with DEXA achieved in patients operated with both models stems.

In the case of cemented stems, densitometric studies were performed with two different types of stem: one straight (Versys, manufactured in a cobalt-chromium alloy) and other anatomical (ABG, manufactured in forged Vitallium patented by Stryker Howmedica). It was carried on the same methodology used in the cemented stems series, but postoperative follow-up was only one year long. Densitometric studies previously made with cementless stems allow us to affirm that bone remodeling is done in the first postoperative year, a view shared by most of the authors. So, we accept that bone mineral density values obtained one year after surgery can be considered as definitive. As in cementless models, densitometric values have been used for comparison with those obtained in the FE simulation models. Our studies confirmed that the greatest loss of bone density affects the area 7 of Gruen (Joven, 2007), which means that stress-shielding and atrophy of metaphyseal bone also occurs in cemented prostheses. This phenomenon is less severe than in non-cemented stems, therefore we can conclude that the load transfer is better with cemented stems than with cementless stems. The findings of proximal bone atrophy, mainly in the area 7, agree with those published by other authors (Arabmotlagh et al., 2006; Dan et al., 2006). We have also found differences in the rates of decrease in bone density in the area 7 of Gruen, which were slightly lower in the anatomical ABG stem than in the Versys straight stem. This also indicates that the prosthesis design has influence in the remodeling process, and that mechanical stimuli are different and related to the design.

4. Application to the lumbar spine

The spine is a complex anatomical structure that has triplanar movements, maintains the erect posture of the individual and supports a significant load. In its central part forms the

spinal canal to contain the nervous structures, therefore has to combine the flexibility to perform movements and to maintain stability and protect nervous structures. The spine changes its mechanical properties depending on the loads, therefore behaves as a viscoelastic structure (Yaszemski et al., 2002). For these special features the study of its biomechanics, in its three areas, is a very complex matter. And it is very difficult to reproduce it for in vivo or in vitro studies. Of the whole of spine, lumbar spine has been widely studied, showing in many papers, a large variability of results.

Biomechanics of the lumbar spine has been studied in cadaveric specimens (Panjabi et al., 1994). But lack of flexibility makes difficult to reproduce the range of motion presented by living persons. In vivo studies has been made by various methods (radiographic, CTA, IMR, TV and computer, electrogoniometer, inclinometer, etc.). The results are extremely variable, even for the same person throughout the day (Ensink et al., 1996) and also have different values on account of age and existing pathology (Sullivan et al., 1994). Animal spines have also been used for these biomechanical studies, despite major differences with the human lumbar spine (Kettler et al., 2007).

Because of the difficulties to research with living persons, their variability with mixed results published, the problems arisen with in vitro studies and differences between human and animal column, we have developed simulation models, using finite elements (FE). This model allows to research on lumbar spine, in physiological conditions, to simulate different load conditions and study the impact on biomechanics. We can also simulate the disc degeneration, to a greater or lesser degree, and study the impact on adjacent elements to degenerated disc. Finally, the model may be useful to test different fixation systems, as a pedicular screw, a interbody device or rigid fixations compared with the dynamics.

By using the methodology described in §1, it is possible to obtain the geometrical model corresponding to the S1-L5 functional unit (Fig. 20).

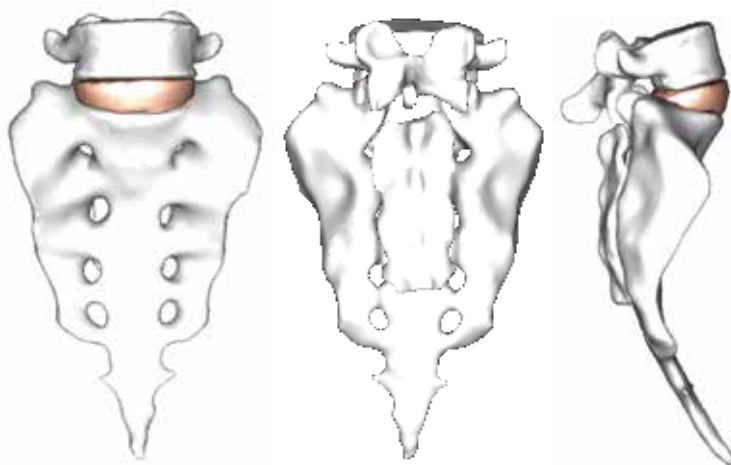


Fig. 20. Geometrical model corresponding to the S1-L5 functional unit

The mesh of the vertebrae is made by means of tetrahedrons with linear approximation in the I-deas program (I-deas, 2007) with a size thin enough to allow a smooth transition from the zone of exterior cortical bone to the zone of interior cancellous bone; this transition was obtained by means of statistical averages from CTs of vertebrae in healthy individuals. Disc

meshes are essential for a correct reproduction of the biomechanical behaviour of the functional unit analysed. In order to do this, each disc is divided into nucleus pulposus and annulus fibrosus with commonly accepted dimensions (White & Panjabi, 1990). Each part is meshed separately so that mesh sizes should match each other and with the vertebrae, getting the complete FE model (Fig. 21).

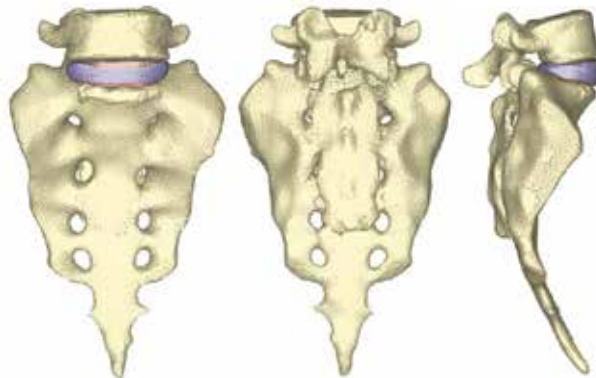


Fig. 21. FE model corresponding to the S1-L5 functional unit

Concerning material properties, those of bone were taken from literature. About discs, nucleus pulposus behaves like a non-compressible fluid, which upon being compressed expands towards the exterior tractioning the fibers of the annulus. Fibers of the annulus show a hyperelastic behaviour, but only in tension. The correct interaction between the different elements (vertebrae, discs and ligaments) is essential. Conditions of union between the vertebral body and the intervertebral disc have been established, as it is the most representative of the real anatomy. Finally, contact conditions have been established between the different apophysis which provides a global stability. In order to verify the effectiveness of the fixation, flexion-extension movement has been analyzed as the most representative (Fig. 22). As boundary conditions displacements in the alas of sacrum have been prevented.



Fig. 22. Flexion-extension movement of the S1-L5 functional unit

Muscle and ligament forces are adjusted to the appropriate values to obtain the required movement ranges. So, the deformed shapes shown in Fig. 23 are achieved. Calculation and post-processing are carried out using the Abaqus program (Abaqus, 2009).

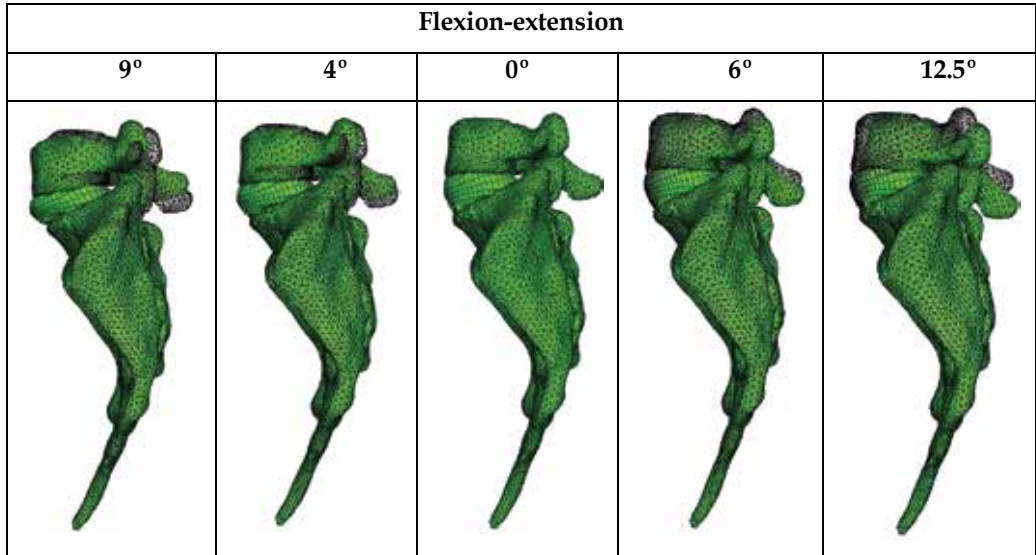


Fig. 23. Deformed shapes of the S1-L5 functional unit for the flexion-extension movement

Once the healthy model has been adjusted and validated, it is possible to simulate new conditions corresponding to different levels of discal degeneration. So, after decreasing the mechanical properties of annulus and fibers in the disc, according with the mechanical damage theory, we can measure the changes produced in the different movements. The next step is the study of several types of fixations (rigid and dynamic). Changes induced by surgery must be reproduced in the FE model. To that effect, the holes for the pedicular screws must be generated by removing the elements intersected by the screws, and resetting the mesh along the new interfaces. In the same way, pedicular wings must be trimmed to allow the introduction of the screws. With the screws in place, the struts are implemented to obtain the final models (Figs. 24 & 25).

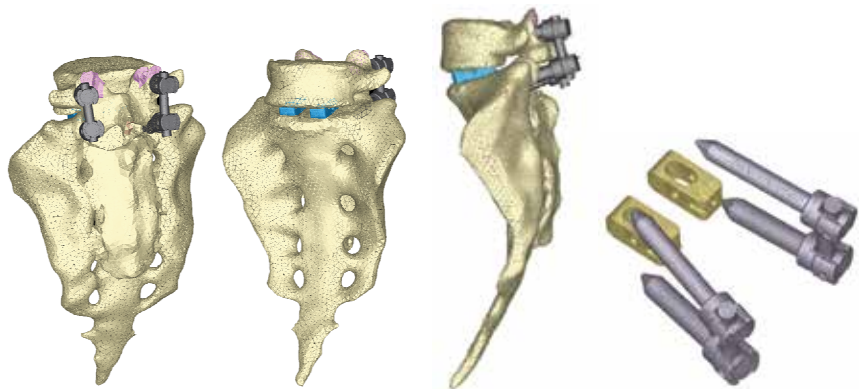


Fig. 24. FE model corresponding to the S1-L5 functional unit with rigid fixation

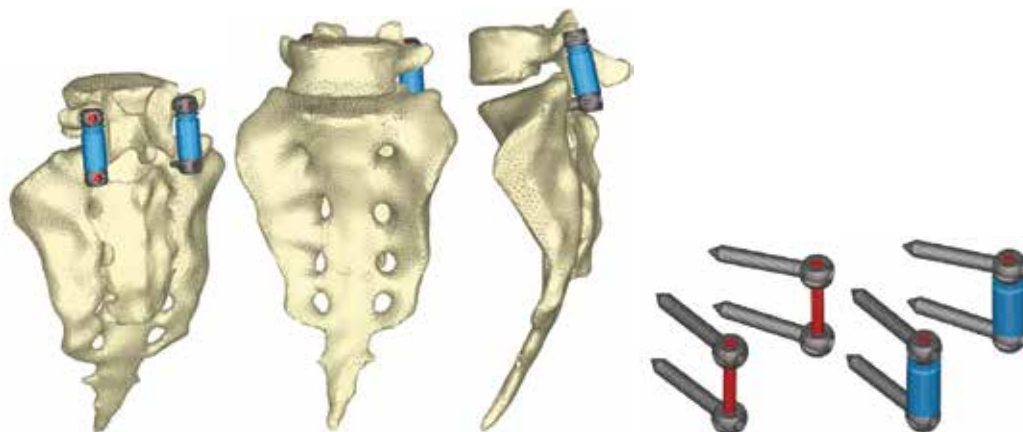


Fig. 25. FE model corresponding to the S1-L5 functional unit with dynamic fixation

From those new models the study of the flexion-extension movement can be reproduced in the new conditions, verifying the changes produces as consequence of the fixations. The results allow comparing the performance of the different models, and moreover it is possible to analyse the changes in the local stress distribution around the screws detecting points of possible future pathologies due to the alterations produces by the presence of the fixations. So, in Figs. 26 and 27 the von Mises stress distribution, both in the healthy and in the implanted models, are shown. Comparing both Figs., a high stress concentration around the screw root is observed, increasing the maximum value till 400% with respect to the healthy model.

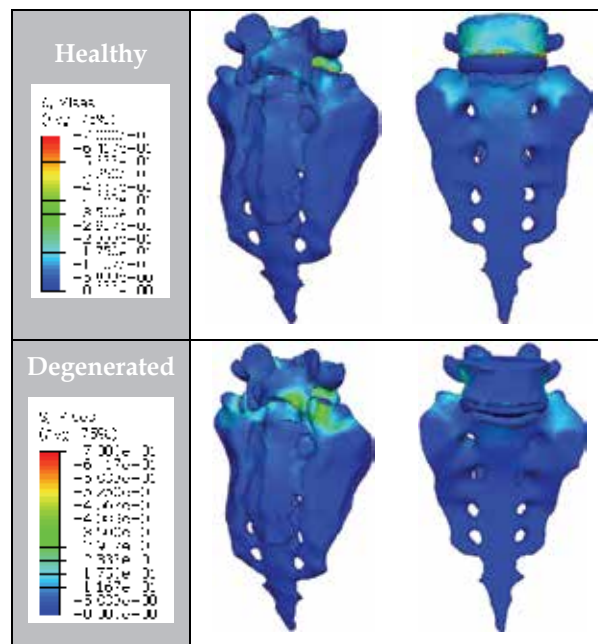


Fig. 26. Von Mises stresses for the healthy and degenerated models

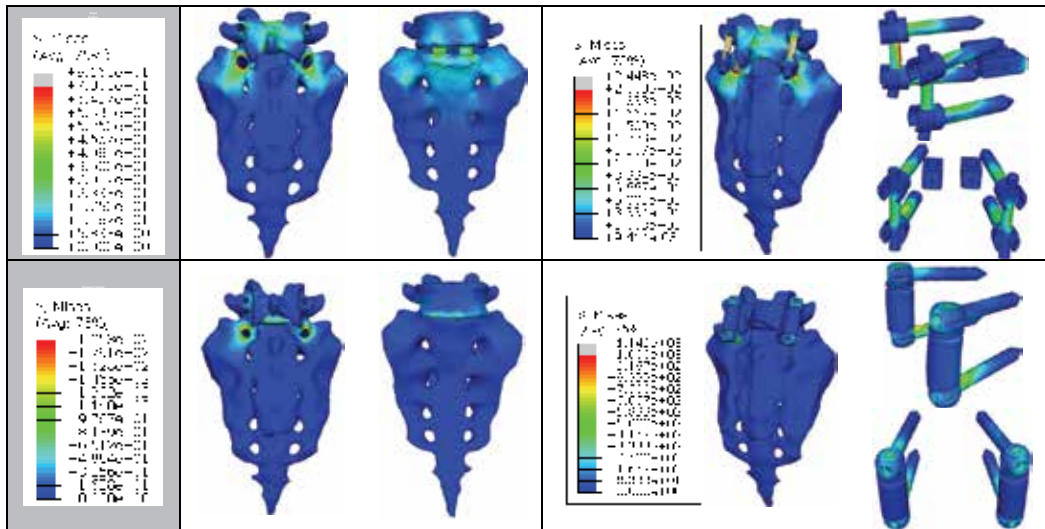


Fig. 27. Von Mises stresses for the implanted models

Future research concerns the development of the complete lumbar spine model, from S1 to L1, including ligaments and cartilages, in order to get a better approximation to the real biomechanical behaviour.

5. Application to splints for hand therapy

In the pathology of the hand is often the existence of joint stiffness, usually post-traumatic, causing vicious positions and loss of mobility of the joints. Also rheumatic disease that affects the joints of the hand, can cause deformity and changes in normal joint alignment. The worst situation occurs when these alterations produce one flexion stiffness of the joints

In the treatment of articular rigidities the usual therapy consists of the application of splints, which utility is guaranteed by different clinic studies that demonstrated its efficacy in most of the pathologies (Prosser, 1996).

The mission of the orthopedics splints is relieve the strength they have or to keep a constant tension, in a determinate mode in the joints and stimulate, that way, the hisitic changes which allow the stretching of the capsule and joint structures until the deformity is corrected. Among all the different pathologies affecting the hand , due to traumas or sickness, one of the more frequent is the contracture by flexion of the proximal interphalangeal joint (AIP). This joint and its mobility are one of the most important factors in the hand functionality.

The use of materials with shape memory and superelastic behaviour like NiTi for the manufacture of orthopedic splints is a possibility with great expectations for the future. The experience of the authors in the design of other devices based on the NiTi alloy (Puértolas et al., 2002; Lahoz et al., 2002; Domingo et al., 2005; Domingo et al., 2006), makes it possible to carry out the proposed design of a finger splint for the treatment of the contracture in flexion of the PIP joint. About the material, NiTi is an equiatomic alloy of nickel and titanium (commercially known as Nitinol), discovered in the U.S. Naval Ordenance Laboratory (Buehler & Wiley, 1965). It belongs to a group of materials with shape memory (SMA). Basically, these alloys have the

attribute of being able to recover a previously defined form when the material is subjected to an adequate thermal treatment; associated to this behaviour, the material has a super elasticity which leads to the property of withstanding large elastical deformations with relatively low tensions. This property is due to the change of phase which the material undergoes when it is subjected to tension (Auricchio & Petrini, 2002).

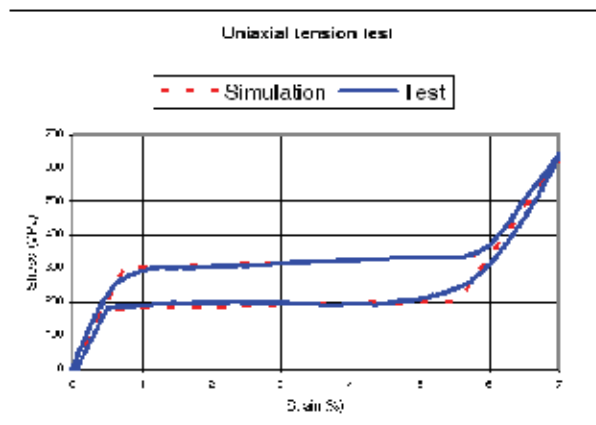


Fig. 28. Stress-strain curve for loading and unloading process corresponding to the NiTi alloy at 22° C

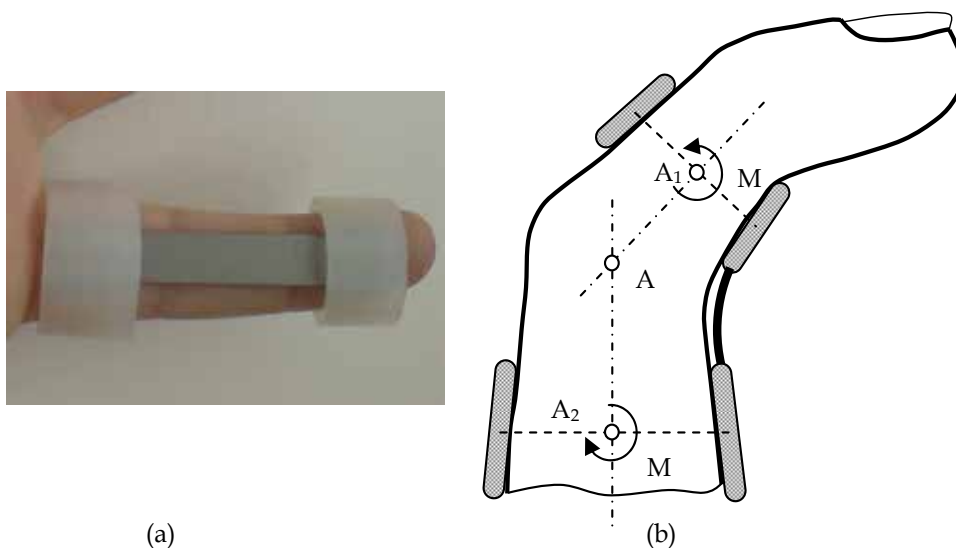


Fig. 29. a) Prototype of the designed NiTi splint; b) Force transmission mechanism for the designed splint

The first step corresponds to characterization of the basic behaviour of the material, because the mechanical properties of NiTi are very sensitive to little changes in the alloy composition. For this reason, it is necessary to carry out tension-compression uniaxial standard test in the overall range of strains, obtaining the whole behaviour curve both for

loading and unloading processes. With this purpose, a universal machine INSTRON 5565 was used to testing standard test specimens obtained from shells 1 mm width corresponding to a reference alloy (50.8% Ni, 49.2% Ti) supplied by the company Memory Metalle GBMH. This alloy exhibits a superelastic behaviour at environment temperature. In every test a complete loading-unloading cycle with displacement control is reached till a maximum strain of 7% at 22° C. In Fig. 28 the tension curve is shown, evaluating from that the different mechanical parameters: elastic moduli for martensite and austenite phases (E_A , E_M), respectively, phase transition stresses (σ_s^{AM} , σ_f^{AM} , σ_s^{MA} , σ_f^{MA}) and maximum strain (ε_L). The proposed design uses a thin plate of NiTi, which is fixed onto the finger by means of rings, which are responsible for transmitting the recovering force (Fig. 29). The mechanism provides a practically unidimensional bending performance, such that the device presents a mechanical response close to the intrinsic material behaviour. This means that the moment-angle curve of the splint has a similar shape to the material tension-deformation curve.

Parameter	Description	Value
E_A	Austenite Young Modulus	52650 MPa
ν_A	Austenite Poisson Ratio	0.33
E_M	Martensite Young Modulus	38250 MPa
ν_M	Martensite Poisson Ratio	0.33
ε_L	Maximum Transformation Strain	6%
σ_s^{AM}	Transformation Activation Stress (A→M)	300 MPa
σ_c^{AM}	Transformation Completion Stress (A→M)	340 MPa
σ_s^{MA}	Transformation Activation Stress (M→A)	200 MPa
σ_c^{MA}	Transformation Completion Stress (M→A)	180 MPa
T_0	Reference Temperature	22 °C
C^{AM}	$\frac{\partial \sigma_{s,c}^{AM}}{\partial T}$	6.7 MPa/°C
C^{MA}	$\frac{\partial \sigma_{s,c}^{MA}}{\partial T}$	6.7 MPa/°C

Table 2. Material properties (NiTi)

The action of the splints is directly related to the rigidity, and in the proposed design the rigidity is directly related to the width, the thickness and the length, although all of these geometric factors work in an uneven way. An increase in width supposes a linear growth in the recovering force and a better finger support. However, the most important factor used to control the force exerted by the splint is the plate thickness. The device is very sensitive to thickness change, presenting a cubic rate influence. Hence, the greater the thickness, the greater the effect of straightening and the smaller the risk of breakage although it is more

difficult to bend the splint and fit it in the volar zone of the injured finger. On the contrary, if the thickness is reduced so is the straightening effect and the risk of breakage increases, although it is easier to bend the splint and fit it on the finger.

To obtain a design which transmits a force adequate for the recovery of the original position of the finger, a finite elements simulation for a plate of these dimensions, 80x10x1 mm, is carried out. For the behaviour of the material a proprietary developed user subroutine is used, based on Auricchio's models (Auricchio & Petrini, 2002), in the Abaqus program (Abaqus, 2009). Previously, an adjustment of parameters from the results of the tensile test is carried out (Fig. 28). The different parameters used in the simulation are gathered in Table 2. As for the boundary conditions, initially a displacement of 1 mm in the centre of the plate is applied to later apply the eccentric compression until reaching the maximum curvature, moment in which the load is removed and a free restoration is produced.

Fig. 30. shows the von Mises stress maps in both faces (top and bottom) of the shell for different curvatures in the whole range of deformation, while Fig. 31 shows the martensite fraction maps for the same levels of deformation as Fig. 30.

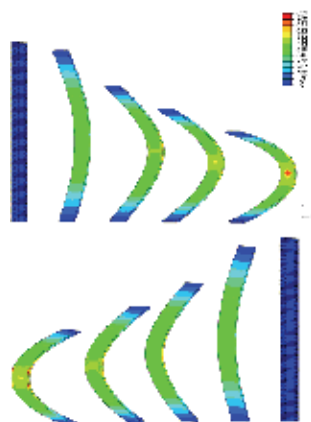


Fig. 30. Von Mises stress maps

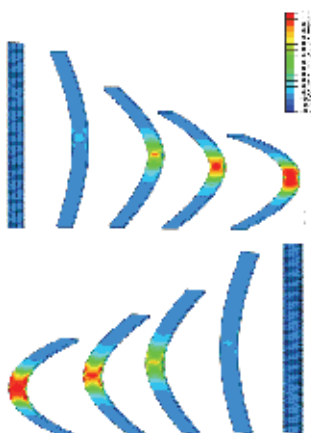


Fig. 31. Martensite fraction maps

Once the model was adjusted, a parametric study was carried out with different values of length, width and thickness for the plate. Figs. 32 and 33 show the Recovering Force-Angle curves for the different lengths and widths of the plate, with a fixed thickness of 1mm, and the Recovering Force-Angle curves for different lengths and thicknesses of the plate, with a fixed width of 6mm. A practically constant recovering force value can be observed over a wide range of angles that vary between 30° and 150° . This makes it possible to define a characteristic value of recovering force which is ascribed to an angle flexion of 80° . The same occurs in all of the cases analysed.

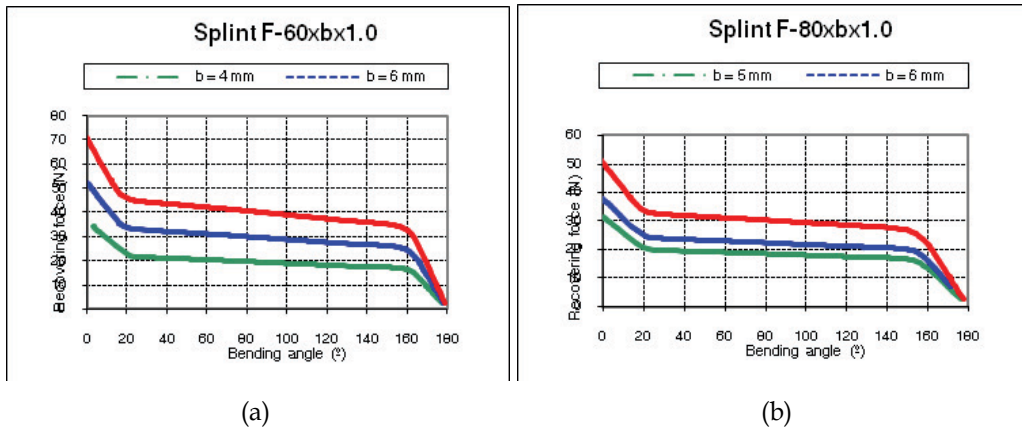


Fig. 32. Recovering force-Angle curves: (a) Length 60.0 mm, width variable, thickness 1.0 mm; (b) Length 80.0 mm, width variable, thickness 1.0 mm

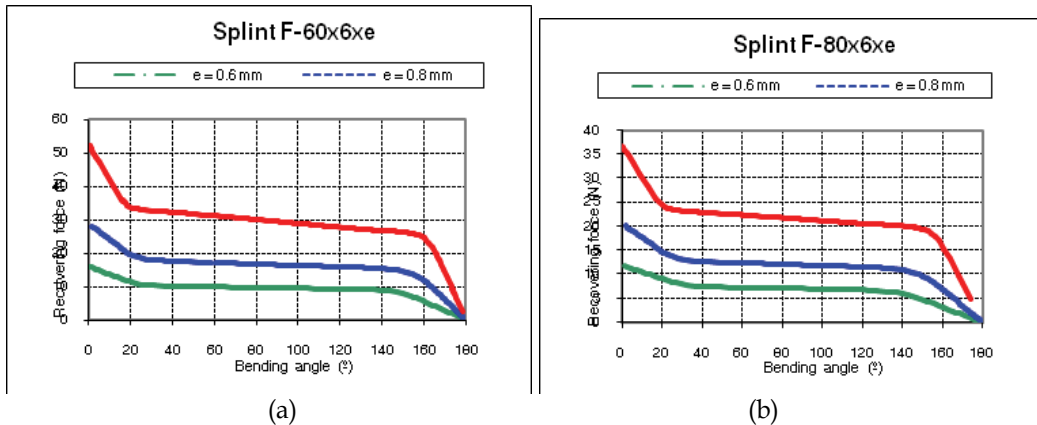


Fig. 33. Recovering force-Angle curves (a) Length 60.0 mm, width 6.0 mm, thickness variable; (b) Length 80.0 mm, width 6.0 mm, thickness variable

In order to design a NiTi prosthesis with a recovering force equivalent to the usual in commercial orthosis, it is necessary to adjust the rigidity to bending to the value of this. A parametric simulation has been carried out, covering a range of forces from 5 to 30 N, with lengths and widths capable of adapting to fingers of different sizes. One main advantage of the proposed designs is that they achieve a practically constant recovering force in a range

of angles from 20° to 150°. This makes possible a therapeutic device that practically covers the total recovery of the PIP joint. In Fig. 34, comparative values of simulation versus experimental results are presented.

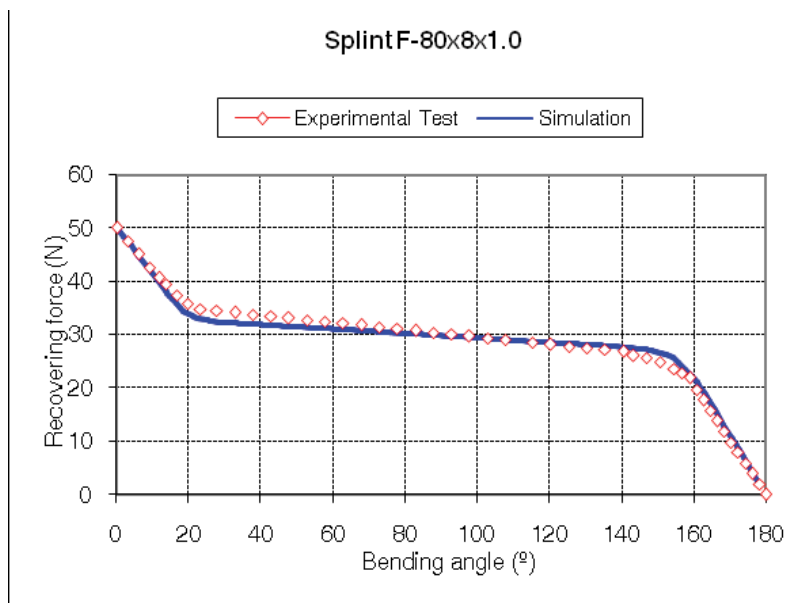


Fig. 34. Recovering force-Angle curves for the 80x8x1.0 mm splint. Experimental test versus simulation

Apart from the properties of NiTi, the biomechanical behaviour of the proposed splint compared to the commercial ones analysed is totally different due to its design. Hence, the mechanism to transmit forces onto the finger in the Bort class of splints (and in the majority of those that exist on the market) is based on equilibrium in a simple bending situation, with action (F_1 , F_2) and reaction (F_3) in the different parts of the splint (Fig. 35). On the contrary, in the developed prototype the transmission mechanism is based on equilibrium in pure bending, transmitting both torques on the fixation rings (Fig. 29), through the local equilibrium of forces in the fixation rings.

The common bending mechanism has two important drawbacks: firstly, for a linear behaviour spring the ratio between recovering force and angle is constant in its whole length. Hence, as recovery is produced the torque transmitted on the joint decreases significantly, even though the distance from the point where the force is applied F_1 on the joint, point A in Fig. 35 becomes progressively larger. Its increase does not compensate the loss of force, for which it is necessary to change the splint for another with a different force calibration. Moreover, the forces involved in the equilibrium have components that can generate compression or traction on the joint itself, possibly increasing the damage to the joint.

However, in the mechanism of pure bending as it is directly transmitting torques, the effect on the joint is always the same in all of the recovering range. In addition forces are not generated on joints; the forces are generated at local level on the fixation rings to give rise to the torques transmitted, acting on zones that are away from the joint and without damaging

effects. Thus the recovering moment is constant in all of the length A_1 - A - A_2 (Fig. 29b), without producing undesired effects on the joint. Given that the splint reproduces in the graph moment-angle the basic behaviour of NiTi in the tensile test, the recovering moment is practically constant in the all of the range where the splint works. Then the whole treatment is possible with only one splint, and without the need of progressive replacements as the joint recovers.

No other splint available on the market offers this property since all of them are based on materials and mechanisms in which global result is that of a cuasi-linear behaviour. This makes it impossible to obtain the curves with a practically null slope in the recovery stage like the one presented here.

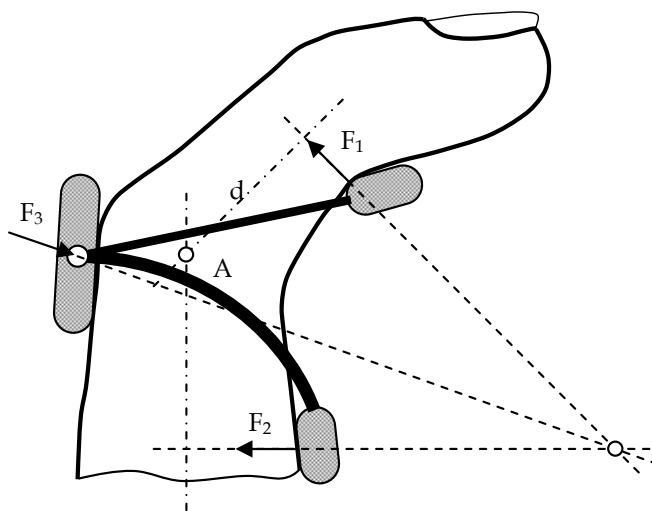


Fig. 35. Force transmission mechanism for Bort type splints.

Another talking point is the optimum time for the use of straightening splints, which Flower defined as TERT (Total End Range Time) (Flower & LaStayo, 1994), having checked that the longer it is worn daily the better the results (Flower & LaStayo, 1994; Flower, 2002). However, Flower himself concludes that in addition to the time of use, the force application parameters are fundamental in attaining a good correction of the deformity (Flower, 2002). Due to its comfort the proposed design makes it possible to wear permanently. On the other hand, permanent action dynamic orthosis, regularly used in orthopaedics for the PIP joint, are difficult to fit, above all at the level of the proximal phalange despite the therapist being able to choose the size.

On the contrary, the designed splint in this work improves the initial adjustment and make it easier to use for both the patient and the specialist, with no difficulties in its fitting. The proposed design also avoids the harmful effect of pressure on the back of the joint (Li, 1999) which is produced in the usual static and dynamic systems. The advantages offered against the most frequently used commercial models can be summarised as:

- Better control of the recovering force.
- Maintaining the splint over long periods of time without replacement, since its effect remains unalterable over a wide period of recuperation.
- Ease of use for both the patient and specialist.

- Ease of producing custom made designs for each patient.
- Significant economic saving in the treatment.

For all of the above mentioned, the proposed design is highly competitive compared to those used presently. The future research concerns the design of alternative geometries and the improvement of the fixation system.

6. References

- ABAQUS. (2009). Web site, <http://www.simulia.com/>.
- Arabmotlagh, M.; Sabljic, R. & Rittmeister, M. (2006). Changes of the biochemical markers of bone turnover and periprosthetic bone remodelling after cemented arthroplasty. *J Arthroplasty*, 21,1, 129-34. ISSN: 0883-5403
- Auricchio, F. & Petrini, L. (2002). Improvements and algorithmical considerations on a recent three-dimensional model describing stress-induced solid phase transformations. *Int J Numer Meth Eng*, 55, 1255-1284. ISSN: 0029-5981
- Bathe, K.J. (1982). *Finite element procedures in engineering analysis*, Prentice-Hall, New Jersey. ISBN-10: 0133173054
- Buckwalter, J.A.; Glimcher, M.J.; Cooper, R.R. & Recker, R. (1995). Bone biology, *J Bone Joint Surg*, 77A, 1276-1289. ISSN: 0021-9355
- Buehler, W.J. & Wiley, R.L. (1965). *Nickel-base alloys*, U.S. Patent 3.174.851.
- Dan, D.; Germann, D.; Burki, H.; Hausner, P.; Kappeler, U.; Meyer, R.P.; Klaghofer, R. & Stoll, T. (2006). Bone loss after total hip arthroplasty. *Rheumatol Int*, 26, 9, 792-8. ISSN: 0172-8172
- Denozière, G. & Ku, D.N. (2006). Biomechanical comparison between fusion of two vertebrae and implantation of an artificial intervertebral disc. *J Biomechanics*, 39,766-755. ISSN: 0021-9290
- Domingo, S.; Puértolas, S.; Gracia, L. & Puértolas, J.A. (2006). Mechanical comparative analysis of stents for colorectal obstruction. *Minim Invasiv Ther*, 15, 6, 331-338. ISSN: 1364-5706
- Domingo, S.; Puértolas, S.; Gracia, L.; Mainar, M.; Usón, J. & Puértolas, J.A. (2005). Design, manufacture and evaluation of a NiTi stent for colon obstruction. *Bio-Med Mater Eng*, 15, 5, 357-365. ISSN: 0959-2989
- Ensink, F.B.; Saur, P.M.; Frese, K.; Seeger, D. & Hildebrandt, J. (1996). Lumbar range of motion: influence of time of day and individual factors on measurements. *Spine*, 21, 11, 1339-43. ISSN: 0362-2436
- Evans, F.G. (1973). *Mechanical properties of bone*, Evans F.G. (Ed), Ed. Springfield, Illinois.
- Flower KR. (2002). A proposed decision hierarchy for splinting the stiff joint, with an emphasis on force application parameters. *J Hand Ther*, 15, 158-162. ISSN: 0894-1130
- Flower, K.R. & LaStayo, P. (1994). Effect of total end range time on improving passive range of motion. *J Hand Ther*, 7, 150-157. ISSN: 0894-1130
- Gibbons, C.E.R.; Davies, A.J.; Amis, A.A.; Olearnik, H.; Parker, B.C. & Scott, J.E. (2001). Periprosthetic bone mineral density changes with femoral components of different design philosophy. *In Orthop*, 25, 89-92. ISSN: 0341-2695
- Glassman, A.H.; Crowninshield, R.D.; Schenck, R. & Herberts, P. (2001). A low stiffness composite biologically fixed prostheses. *Clin. Orthop*, 393, 128-136. ISSN: 0009-921X

- Guan, Y.; Yoganandan N.; Zhang, J.; Pintar, F.A.; Cusick, J.F.; Wolfla, Ch. E. & Maiman, D.J. (2006). Validation of a clinical finite element model of the human lumbosacral spine. *Med Bio Eng Comput*, 44, 633-641. ISSN: 0140-0118
- Herrera, A.; Panisello, J.J.; Ibarz, E.; Cegoñino, J.; Puertotas, J.A. & Gracia, L. (2009). Comparison between DEXA and Finite Element studies in the long term bone remodelling of an anatomical femoral stem. *J Biomech Eng*, 31, 4, 1004-13. ISSN: 0148-0731
- Herrera, A.; Panisello, J.J.; Ibarz, E.; Cegoñino, J.; Puertotas, J.A. & Gracia, L. (2007). Long term study of bone remodelling after femoral stem: a comparison between DEXA and finite element simulation. *J Biomechanics*, 40, 16, 3615-25. ISSN: 0021-9290
- Hughes, T.J.R. (1987). *The finite element method*, Prentice-Hall, New Jersey. ISBN-10 013317025X
- Huiskes, R.; Weinans, H. & Dalstra, M. (1989). Adaptive bone remodeling and biomechanical design considerations for noncemented total hip arthroplasty. *Orthopedics*, 12, 1255-1267. ISSN: 0147-7447
- I-DEAS. (2007). Web site, <http://www.ugs.com/>.
- Imbert, F.J. (1979). *Analyse des structures par élément finis*, Cepadues Edit, Toulouse. ISBN: 2854280512
- Joven, E. (2007). *Densitometry study of bone remodeling in cemented hip arthroplasty with stem straight and anatomical*. Doctoral Degree Dissertation, University of Zaragoza.
- Kerner, J.; Huiskes, R.; van Lenthe, G.H.; Weinans, H.; van Rietbergen, B.; Engh, C.A. & Amis, A.A. (1999). Correlation between pre-operative periprosthetic bone density and post-operative bone loss in THA can be explained by strain-adaptive remodelling. *J Biomechanics*, 32, 695-703. ISSN: 0021-9290
- Kettler, A.; Liakos, L.; Haegerle, B. & Wilke, H.J. (2007). Are the spines of calf, pig and sheep suitable models for pre-clinical implant tests?. *Eur Spine J*, 16, 12, 2186-92. ISSN: 0940-6719
- Lahoz, R.; Gracia, L. & Puértolas, J.A. (2002). Training of the two-way shape memory effect by bending in NiTi alloys. *J Eng Mater-T*, 124, 4, 397-401. ISSN: 0094-4289
- Li, C. (1999). Force analysis of the Belly Gutter and Capener splints. *J Hand Ther*, 12, 337-343. ISSN: 0894-1130
- Li, M.G.; Rohrl, S.M.; Wood, D.J. & Nivbrant, B. (2007). Periprosthetic Changes in Bone Mineral Density in 5 Stem Designs 5 Years After Cemented Total Hip Arthroplasty. No Relation to Stem Migration. *J Arthroplasty*, 22, 5, 698-91. ISSN: 0883-5403
- Little, J.P.; Adam, C.J.; Evans, J.H.; Pettet, G.J. & Pearcy, M. J. (2007). Nonlinear finite element analysis of annular lesions in the L4/5 intervertebral disc. *J Biomechanics*, 40, 2744-2751. ISSN: 0021-9290
- Marklof, K.L.; Amstutz, H.C. & Hirschowitz, D.L. (1980). The effect of calcar contact on femoral component micromovement. A mechanical study. *J Bone Joint Surg*, 62A, 1315-1323. ISSN: 0021-9355
- McAuley, J., Sychterz, Ch. & Ench, C.A. (2000). Influence of porous coating level on proximal femoral remodeling. *Clin Orthop Relat R*, 371, 146-153. ISSN: 0009-921X
- MIMICS. (2010). Web site, www.materialise.com.

- Panisello, J.J.; Herrero, L.; Herrera, A.; Canales, V.; Martínez, A.A. & Cuenca, J. (2006). Bone remodelling after total hip arthroplasty using an uncemented anatomic femoral stem: a three-year prospective study using bone densitometry. *J Orthop Surg*, 14, 122-125. ISSN: 1022-5536
- Panisello, J.J.; Canales, V.; Herrero, L.; Herrera, A.; Mateo, J. & Caballero, M.J. (2009). Changes in periprosthetic bone remodelling alter redesigning an anatomic cementless stem. *Inter Orthop*, 33, 2, 373-380. ISSN: 0341-2695
- Panisello, J.J.; Herrero, L.; Canales, V.; Herrera, A.; Martínez, A.A. & Mateo, J. (2009). Long-term remodelling in proximal femur around a hydroxiapatite-coated anatomic stem. Ten years densitometric follow-up. *J Arthroplasty*, 24, 1, 56-64. ISSN: 0883-5403
- Panjabi, M.M.; Oxland, T.R.; Yamamoto, I. & Crisco, J.J. (1994). Mechanical behavior of the human lumbar and lumbosacral spine as shown by three-dimensional load-displacement curves. *J Bone Joint Surg Am*, 76, 3, 413-24. . ISSN: 0021-9355
- Prosser, R. (1996). Splinting in the management of proximal interphalangeal joint flexion contracture. *J Hand Ther*, 9, 378-386. ISSN: 0894-1130
- Puértolas, J.A.; Pérez-García, J.M.; Juan, E. & Ríos, R. (2002). Design of a suture anchor based on the superelasticity of the Ni-Ti alloy. *Biomed Mater Eng*, 12, 3, 283-289. ISSN: 0959-2989
- Radin, E.L. (1980). Biomechanics of the Human hip. *Clin Orthop Relat R*, 152, 28-34. ISSN: 0009-921X
- Ramaniraka, N.A.; Rakotomanana, L.R. & Leyvraz, P.F. (2000). The fixation of the cemented femoral component .Effects of stem stiffness, cement thickness and roughness of the cement-bone surface . *J Bone Joint Surg*, 82Br, 297-303. ISSN: 0301-620X
- Rubash, H.E.; Sinha, R.K.; Shanbhag, A.S. & Kim, S.Y. (1998). Pathogenesis of bone loss after total hip arthroplasty. *Orthop Clin N Am*, 29(2), 173-186. ISSN: 0030-5898.
- Sullivan, M.S.; Dickinson, C.E. & Troup, J.D. (1994). The influence of age and gender on lumbar spine sagittal plane range of motion. A study of 1126 healthy subjects. *Spine*, 19, 6, 682-6. ISSN: 0362-2436
- Sumner, D.R. & Galante, J.O. (1992). Determinants of stress shielding :designing versus materials versus interface. *Clin Orthop Relat R*, 274, 202-12. ISSN: 0009-921X
- Sychter, C.J. & Engh, C.A. (1996). The influence of clinical factor on periprosthetic bone remodeling. *Clin Orthop Relat R*, 322, 285-292. ISSN: 0009-921X
- Turner, M.J.; Clough, R.W.; Martin, M.C. & Topp, L.J. (1956). Stiffness and deflection analysis of complex structures. *J Aeronautical Society*, 23, 9, 805-823.
- Weinans, H.; Huiskes, R. & Grootenboer, H.J. (1994). Effects of fit and bonding characteristics of femoral stems on adaptative bone remodelling. *J Biomech Eng*, 116, 4, 393-400. ISSN: 0148-0731
- White, A.A. & Panjabi, M.M. (1990). *Clinical Biomechanics of the Spine*, Lippincott Williams & Wilkins, Philadelphia. ISBN-10: 0397507208
- Yaszemski, M.J.; White III, A.A., Panjabi, M.M. (2002). Biomechanics of spine, In: *Orthopaedic Knowledge Update: Spine 2*, Fardon, D.F. and Garfin, S.R. (Ed.), 17-26, Amer. Acad. Orth. Surgeons, Illinois
- Zienkiewicz, O.C. & Morgan, K. (1983). *Finite element and approximation*, John Wiley & Sons, New York. ISBN 10: 0471982407
- Zienkiewicz, O.C. (1967). *The finite element method in structural and continuum mechanics*, Prentice-Hall, New Jersey. ASIN: B000HF38VG

Finite Element Analysis in Orthopaedic Biomechanics

Daniel Kluess, Jan Wieding, Robert Souffrant,
Wolfram Mittelmeier and Rainer Bader
*University of Rostock, Department of Orthopaedics
Rostock, Germany*

1. Introduction

Based on the evolution of today's personal computers and workstations, numerical simulation gained an important significance in biomechanical research. A common tool in numerical simulation is the finite element method (FEM), which can be applied in various ways, e.g. in structure mechanics, thermodynamics or acoustics. The field of orthopaedic biomechanics utilises two major numerical approaches: musculo-skeletal research is based on multi-body-dynamics, dealing with kinematics of the skeletal systems and muscle activity. The second, structure-mechanical point-of-view deals with stress and strain analysis of bone, joints (natural and artificial) and load-bearing implants. Finite-element-analysis (FEA) is the preferred method for this second group of numerical problems occurring in orthopaedic biomechanics.

Due to the resolution of the calculated mechanical parameters, e.g. stress and strain, the results get more accurate the more elements the analysed structure is divided into. Taking a closer look at literature reveals how the development of faster and also affordable computers accelerated the finer mesh density of FE-models over the past years. A FE-model of the pelvis, published in 1995 by Dalstra et al. (Dalstra, Huiskes et al. 1995) and frequently cited, consisted of 2,602 elements. Five years later, Garcia et al. (Garcia, Doblare et al. 2000) reported results gained from a FE-model of the pelvis consisting of 6,425 elements. In 2006, in conjunction with the development of higher-performance workstations, Manley et al. (Manley, Ong et al. 2006) presented their pelvis model using almost 300,000 finite elements. However, even the finest mesh density and accuracy of results is useless if the orthopaedic researchers fail to draw clinical relevance out of their analyses. Naturally, despite very few exceptions, the FE analysts are engineers, who need to make their results understandable to a clinician not familiar with numerical simulation.

What is the key to a fruitful communication between engineers and clinicians? What enlivens the circle between clinical input and provision of analytical results? At first, both the clinician and the engineer have to get used to the other one's subject-specific language. Each should be familiar with fundamental terms in order to enable a successful communication. Further, the clinician needs to know the potentials of numerical analysis. It

unquestionably is beneficial if the clinician can also question the engineer's results, which necessitates the clinician's basic understanding of the pitfalls in numerical analysis.

On the other hand, the engineer needs the clinician's input in terms of e.g. boundary conditions, geometry and material. By following these guidelines, a curriculum of the 'input of relevant clinical questions' and 'provision of solutions with analytical background', as depicted in Fig. 1, can be achieved.

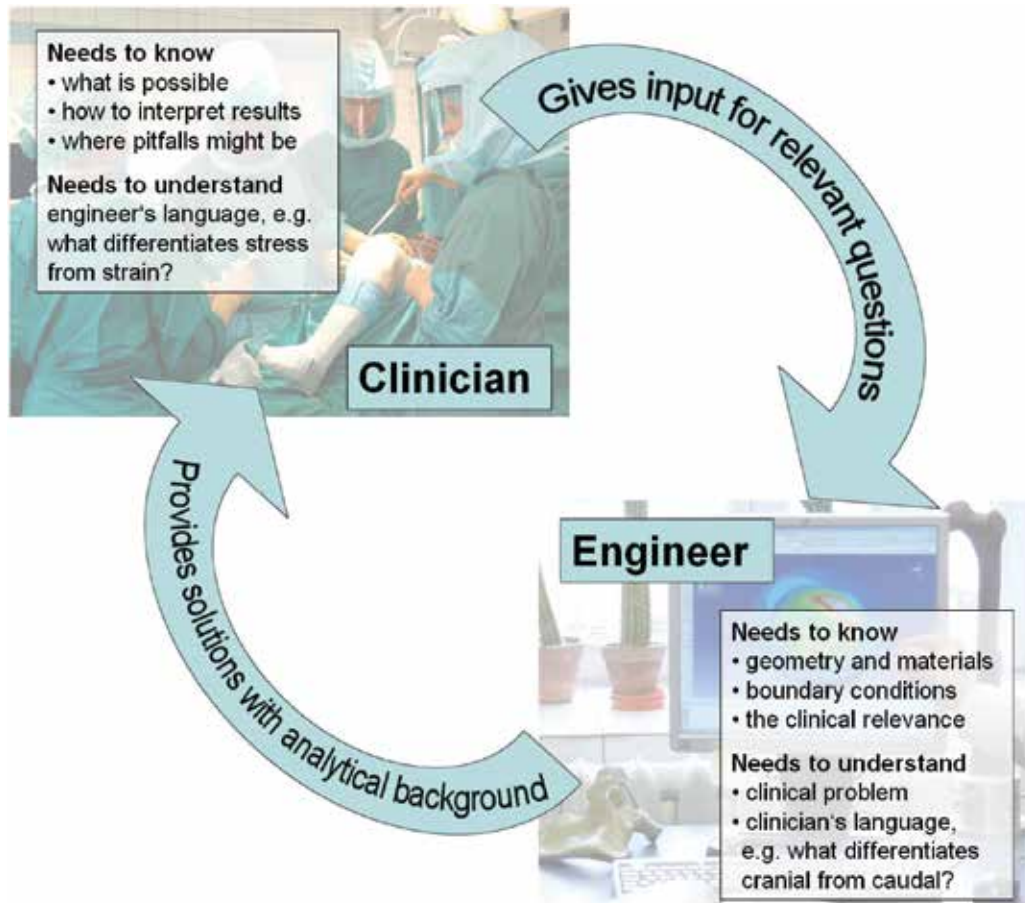


Fig. 1. Communication and knowledge scheme in numerical orthopaedic biomechanics between clinician and engineer

In orthopaedic biomechanics, as well as in related areas of expertise, such as traumatology research or dental biomechanics, modelling of the problem necessitates consideration of bone, implants and both combined, called the implant-bone-compound. As it is the most common implant in orthopaedic surgery, we want to take a closer look at the total hip replacement (THR).

There are about 500,000 THR surgeries in Europe (150,000 in Germany) each year (Aldinger, Jung et al. 2005), with growing numbers throughout the world. Implantation of an artificial hip joint has become a standard procedure in the past decades and the enhanced

performance of today's implants draws attention to younger and more active patients. The major indication for THR is arthrosis, a degenerative disease of the hip joint, which is caused by destruction of the hip joint cartilage.

The modern THR consists of the acetabular component and the femoral component (Fig. 2). The acetabular component usually is a metal shell which holds an insert, mainly made of polyethylene or ceramics. The metal shell is either impacted (press-fit) or screwed (threaded) in the prepared acetabulum. The femoral component consists of the stem and the ball-head. The prosthetic head can be made of either ceramics or cobalt-chromium. The prosthetic stem is, in the majority of THR, made of titanium alloy (Ti-6Al-4V). The mainly used bearings are ceramic-on-ceramic, ceramic-on-polyethylene and metal-on-polyethylene. Despite the good results achieved with THR, there is still a number of serious complications that require further research on implant design, implant material, bearing material, coating and many more.



Fig. 2. Total hip replacement (THR) displayed in the surrounding bone with cut view of the proximal femur

The finite-element-method is used frequently in implant development, helping to answer unresolved questions related to clinical complications. Presently, various approaches to generate models of the implant-bone-compound in THR are published (Spears, Pfeleiderer et al. 2001; Thompson, Northmore-Ball et al. 2002; Kaku, Tsumura et al. 2004; Oki, Ando et al. 2004; Manley, Ong et al. 2006). However, these models each have its issues in being reproducible by the reader. In any case, these issues occur because it was not the publisher's focus to describe the methodology in detail, but to present their results.

In the following chapter of this paper, we firstly want to present a convenient modus operandi of generating FE-models of the implant-bone-compound which can be reproduced by other scientists. We developed an approach starting from computed tomograms of the patient and corresponding CAD-models of the implant. The algorithm is aimed at predicting the stress and strain states in the surrounding bone stock and in the implant itself and has the potential to predict relative micromotion.

Furthermore, we show an example of how this modelling approach can be applied. Therefore, a finite-element study of a newly developed implant for acetabular cup revision (exchange of the acetabular component after implant loosening) is demonstrated.

2. How to generate finite-element-models of the implant-bone-compound

Opposed to other scientific disciplines in engineering and design, the analyzed structures in Orthopaedic Biomechanics are not man-made, but of biological origin. Hence, when bone is analyzed, there are no exactly defined angles, curves and distances, but patient-specific morphology which is highly inhomogeneous and which changes over lifetime, depending on physiological loads, health, age and nutrition. Bone morphology is mainly determined by genetic factors, but also by mechanics, as discovered by the German orthopaedic surgeon Julius Wolff (Wolff 1892). Wolff's law states that bone has the ability to adapt to mechanical loads, i.e. the external and internal structure of bone is transformed depending on the load occurring in the bone. Especially with regard to implant technology and arthroplasty, bone transformation plays an important role. If the biomechanical distribution of forces in and around the treated joint is reconstructed inappropriately during surgery, or if the design of the implant is improper, so-called 'stress shielding' can occur. If the forces are mainly transferred by the implant, adjacent bony regions get minimally loaded and are subsequently degraded. This is another problem that can be solved by stress and strain analysis of the implant-bone-compound.

Modern imaging methods like computed tomography (CT) and magnetic resonance tomography (MRT) allow reconstruction of biological structures for computational finite-element-analysis, whereby CT is the method of choice for bone. The three-dimensional reconstruction is the basis for the presented approach.

On the other hand, the implants to be analyzed in the finite-element-model can be designed using CAD-software, imported from the manufacturer's data or can also be three-dimensionally reconstructed from scan data, e.g. laser-scanning point clouds.

Starting from CT-data of bone and CAD-data of the implant, our approach to model the implant-bone-compound is demonstrated in Fig. 3. A stepwise description of the single modules follows.

The algorithm is based on a number of software packages used (Table 1). The mentioned software modules are not mandatory; however, during development of this algorithm, we found no relevant issues in interfaces and data transfer using these packages.

In the following subchapters, a detailed description of the single steps included in the algorithm is given.

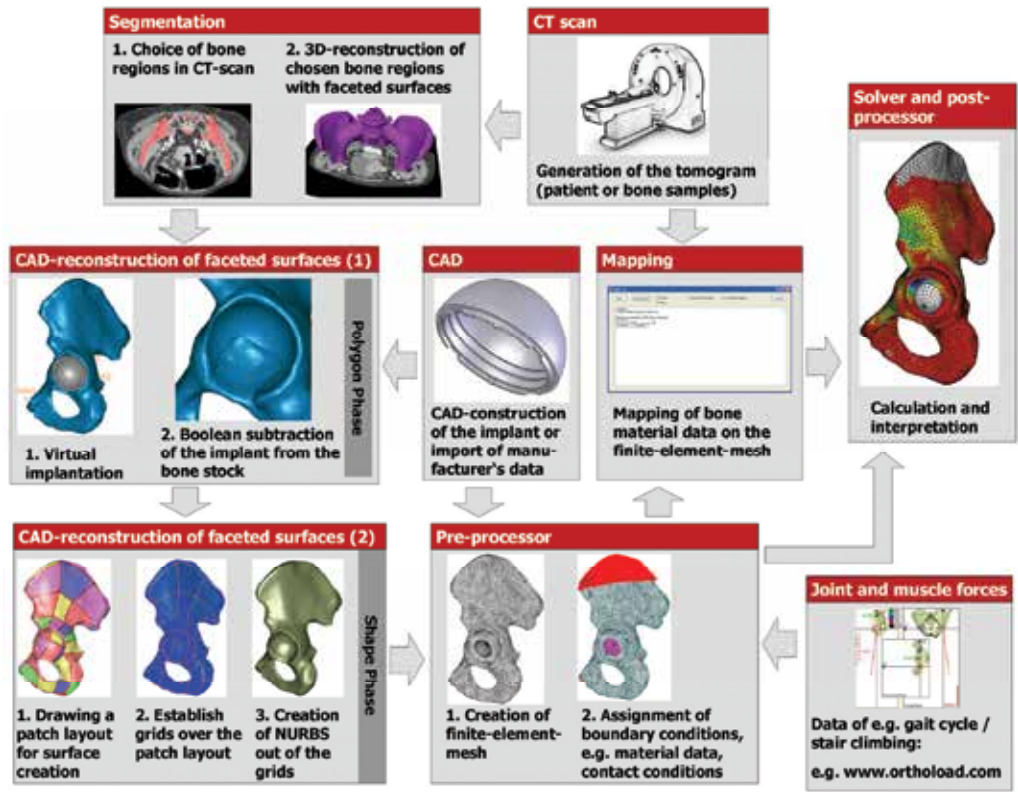


Fig. 3. Approach to analyze the implant-bone-compound via finite-element-methods

Purpose	Software	Manufacturer / Distributor
Segmentation	AMIRA	Mercury Computer Systems Inc., MA, USA
CAD-reconstruction of faceted surfaces	GEOMAGIC STUDIO	Raindrop Geomagic, NC, USA
CAD	SOLIDWORKS	Dassault Systèmes, RI, USA
Mapping	AMAB	Self-developed program, University of Rostock, Germany
Pre-processor	MSC/PATRAN	Marc Schwendler Corp., CA, USA
Solver	Simulia/ ABAQUS	Dassault Systèmes, RI, USA
Post-processor	Simulia/ ABAQUS Viewer	Dassault Systèmes, RI, USA

Table 1. Software used for generating the presented finite-element-models (not mandatory, other software packages, e.g. CAD-software, are considered equivalent)

2.1 Three-dimensional reconstruction of bone morphology

The reconstruction of bone morphology is based on a stack of CT-slices showing the designated bone. In our approach, we make use of the correlation of bone stiffness and attenuation (Snyder and Schneider 1991; Rho, Hobatho et al. 1995). The level of attenuation is measured in Hounsfield Units (HU), which are typically used in the DICOM-files of the CT-scans. HU are normalized in the way that air has a value of -1,000 and water has a value of 0. Bone usually has HU of 250-3,000. The DICOM-files which include all of the sectional slices calculated by the CT scanner are imported into the software AMIRA. After the bony structures are labeled in all slices of the CT-scans, an automatic three-dimensional reconstruction of the bone is performed creating triangulated surfaces (Fig. 4).

2.2 Virtual implantation and CAD-transformation of faceted surfaces

After the bone is imported in the software GEOMAGIC, the virtual implantation can be undertaken. It might be necessary to adjust the position of the bone in order to achieve a definite implant position. There are recommendations for placing coordinate systems at bones and joints given by the International Society of Biomechanics (Wu, Siegler et al. 2002; Baker 2003; Wu, van der Helm et al. 2005). At this point it is very important to record the 4*4 transformation matrix, so the adjustment can be undone at a later stage using the inverse of the transformation matrix..

When the bone is adjusted properly, the CAD-file of the implant can be imported in GEOMAGIC using the IGES interface. The bone stock can be prepared for virtual implantation by subtracting the implant geometry from the bone morphology. Additionally, press-fit cavities can be created by downscaling the implant before subtraction. Since the bone is existent in GEOMAGIC's polygon phase, the implant also has to be transformed to a polygon model before the Boolean subtraction is performed. Implant placement plays a major role in order to generate an adequate model of the implant-bone-compound. In clinical application, implant placement is determined via the surgery manual, the instruments (e.g. cutting guide) as well as accurately defined angles. For example, the acetabular cup in total hip replacement is placed according to lateral abduction, also known as inclination, as well as anteversion. Nevertheless, an orthopaedic surgeon should be consulted in order to verify an adequate virtual implantation.

After the cavity is created in the bone, the next important step is to convert the faceted polygon surfaces to analytical non-uniform rational b-spline surfaces (NURBS). NURBS are analytically defined surfaces which are based on control points. The calculation and definition of NURBS is a major advantage compared to discrete tessellated surfaces. On the one hand, this allows a better performance of CAD modelling, on the other hand, the analytical surfaces can be distributed such that an automatic hexahedral meshing operation can be facilitated (Fig. 5).

When the faceted surfaces are converted into NURBS properly, the bone can be transferred to a pre-processor using the IGES interface. Provided the surfaces form a waterproof body, PATRAN is able to convert these surfaces to one solid. Then the solid can be meshed automatically with tetrahedral finite elements.

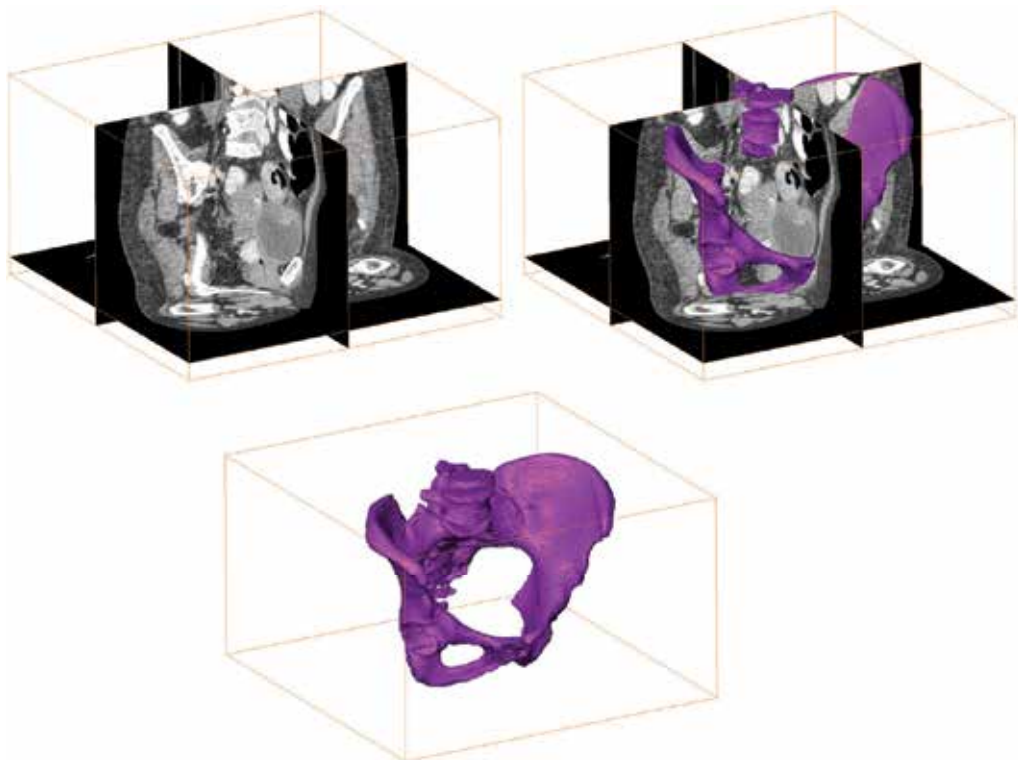


Fig. 4. Three-dimensional reconstruction of bone from CT-scan data using AMIRA. Bone morphology is exported as a tessellated surface.

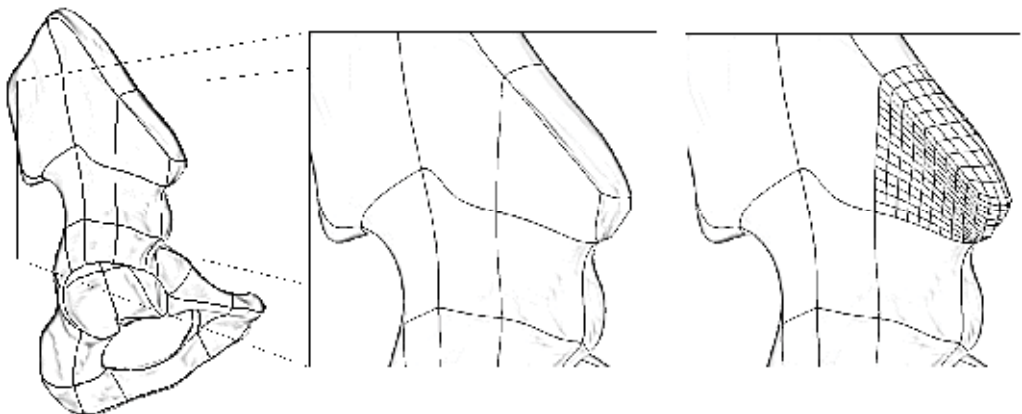


Fig. 5. Distribution of NURBS-surfaces such that quadrilateral surfaces are oppositely located enables generation of solids with five or six faces. Such solids can be meshed with hexahedral elements automatically.

2.3 Mapping of material data on the finite-element-mesh

Mapping of material data from the CT slices onto the finite-element-mesh is a common procedure in finite-element-analysis nowadays (Thompson, Northmore-Ball et al. 2002;

Anderson, Peters et al. 2005; Manley, Ong et al. 2006; Schultze, Kluess et al. 2007). The most common technique is to map the HUs onto the elements and to assign a high number of different material definitions corresponding to equivalent ranges of HU. However, our approach is to handle the HU as temperatures and to assign a temperature-dependent material model (Zacharias 2001; Schultze, Kluess et al. 2007). Hence, the mapping is not proceeded element-wise, but node-wise. Instead of defining several material models for the different ranges of HU, only one material definition depending on the assigned temperatures is needed, resulting in a cost-effective and time-saving calculation. This also enables consideration of cortical bone with a constant Young's modulus as well as consideration of trabecular bone with a HU-dependent modulus using the same temperature-dependent material definition.

There are both material models for trabecular bone which depend directly on the HU (Snyder and Schneider 1991; Rho, Hobatho et al. 1995) and material models which depend on the apparent density respectively the Calcium equivalence (Snyder and Schneider 1991; Dalstra, Huiskes et al. 1993; Keller 1994). To correlate the HU from the CT scans with apparent density and Calcium equivalence, a bone mineral density phantom must be CT-scanned together with the relevant bone.

Using a Toshiba tomograph (Aquilion 32, TOSHIBA Medical Systems GmbH, Germany) at 120 kV and 300 mAs, we found a linear correlation between HU and Calcium density (Fig. 6).

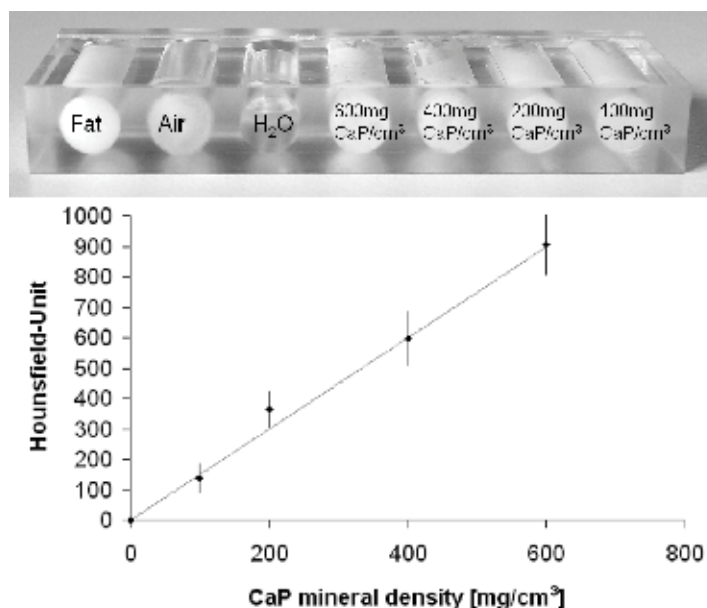


Fig. 6. Top: Custom made bone mineral density phantom. Bottom: Correlation between HU and Calcium mineral density determined by using the bone mineral density phantom.

With our custom-made mapping algorithm (AMAB, University of Rostock), each nodal coordinate of the FE-mesh is retrieved in the stack of CT-slices by detecting firstly the coronal height (axis z) and secondly the location in the coronal plane (axis x and y). The HU

in the corresponding location in the CT slices are averaged within a definite zone surrounding the nodal coordinates. The software allows adjustment of the averaging zone by means of height and width. A weighting algorithm is included which increases the weight of the HU in closer proximity to the FE-node's location during averaging. Besides the density distribution in cancellous bone, the varying thickness of pelvic cortical bone can be represented with high accuracy.

2.4 Applying boundary conditions

The application of boundary conditions in biomechanical FEA is based on assumptions including forces and pressures acting in the human body, as well as displacements and symmetry boundary conditions based on simplifications in the model. The major source of acting loads in the musculoskeletal system for orthopaedic biomechanics is the telemetric in-vivo measurement using instrumented implants (Bergmann, Graichen et al. 1993; Bergmann, Deuretzbacher et al. 2001; D'Lima, Patil et al. 2006). Some of the data gathered by Bergmann et al. is now available online at <http://www.orthoload.com>.

Furthermore, the calculation of acting muscle forces by means of inverse dynamics as a numerical approach yields growing potential in musculoskeletal biomechanics (Rasmussen, Damsgaard et al. 2003). Future steps in musculoskeletal biomechanics will have as their goal the implementation of inverse dynamics with regard to kinematics into finite-element-methods in a combined model.

3. Application of the finite-element-method for preclinical analysis of an endoprosthetic implant

Following the presentation of how to generate finite-element-models of the implant-bone-compound, an example of the application of this algorithm is given. Therefore, we take a closer look at the clinical problem first.

3.1 Clinical background and objective

Despite the technical advances in THR the number of revisions (exchange of a failed implant) in the United States doubled from 1990 till 2002 (Kurtz, Mowat et al. 2005). The major indication for revision of approx. 75 % of failed THR is aseptic loosening (Malchau, Herberts et al. 2002). The reasons why, in turn, revisions fail are instability (35%) and again, aseptic loosening (30%) (Springer, Fehring et al. 2009). The occurrence of failed THRs displaying massive deficiencies in acetabular bone stock is enhanced by the increasing number of total hip arthroplasties.

Numerous implant systems with different fixation principles, specially designed for acetabular cup revision, have been developed so far and have revealed fair to good clinical results. The cemented systems include acetabular roof rings (Siebenrock, Trochslar et al. 2001; Gerber, Pisan et al. 2003; Yoon, Rowe et al. 2003) and antiprotrusio cages (Peters, Curtain et al. 1995; Wachtl, Jung et al. 2000; Perka and Ludwig 2001; Weise and Winter 2003; Gallo, Rozkydal et al. 2006), which are usually implanted in combination with autografts. Cementless acetabular revision systems include bilobed components (DeBoer and Christie 1998; Chen, Engh et al. 2000), oval shaped cups (Götze, Sippel et al. 2003; Hauray, Raeder et al. 2004; Civinini, Capone et al. 2007), pedestal cups (Perka, Schneider et al. 2002; Schoellner

2004) and oversized acetabular components (Dearborn and Harris 2000; Whaley, Berry et al. 2001). Most of the non-cemented acetabular revision systems are designed for a small range of defects and do not provide the modularity necessary to adapt to the wide range of patient-individual bone morphology and defect sizes seen at reoperation.

The presented example of a finite-element-study in Orthopaedic Biomechanics forms the basis in the development of a new acetabular cup revision system capable of providing a cementless solution. The development is aimed at patient individual situations for enhanced fixation of the cup by additional polyaxial and angular stable fixation pegs as well as a modular adaptable lateral flap. The body of the revision cup is oblong shaped to fit the acetabular defect without sacrificing too much bone in the anterior and posterior column. The lateral flap is available in different angles to fit the patient's pelvic morphology individually. An outstanding feature of the fixation pegs is the ability to apply the pegs right after the cup is impacted into the bone. Consequently, the orthopaedic surgeon can decide whether to use one, two or no fixation pegs based on the objective primary stability established by press-fit of the cup and the optional additional fixation using the flap. Moreover, the fixation pegs provide angular stable fixation with the cup within a rotation angle of 16° .

The objective using the finite-element-method is to predict the biomechanical performance of the new implant system in terms of micromotion and stress states in the bone. The biomechanical analysis is carried out by modeling different combinations of fixation pegs and the lateral flap used with the implanted revision cup system and applying loads corresponding to the gait cycle and stair climbing. The calculated micromotion in the implant-bone-compound gives a prediction of how the primary stability of the newly developed acetabular cup revision system can be enhanced using additional fixation.

3.2 Methods

The CAD prototype of the newly developed acetabular cup revision system is shown in Fig. 7. The lateral flap is attached using a dove-tail in combination with two securing screws and locking rings. The implant system contains three different flaps with ascending blade angles to account for the patient-individual pelvic morphology. The flaps are anatomically shaped to provide optimum fit onto the underlying bone. Each lateral flap is manufactured with three drill holes for the application of bone screws.

In order to analyze the biomechanical performance of the additional fixation elements, a FE-model of the pelvis with the implanted revision cup system was modelled using our methodology from the previous chapter. Peak loads from walking and from stair climbing were applied. The morphology of a hemi-pelvis was reconstructed from high resolution computed-tomography (CT) (slice distance 0.3 mm). The hemi-pelvis was oriented in a Cartesian coordinate system, with the x -axis connecting the right and left spinae iliacae anterior superior (SIAS), the y -axis touching the pubic symphysis and pointing cranially and the z -axis pointing anteriorly.

An acetabular bone defect of D'Antonio Type III (D'Antonio, Capello et al. 1989) was created in the computer model of the hemi-pelvis. Afterwards, the osteolysis-modified bone structure was imported into the software GEOMAGIC (Raindrop Geomagic, NC, USA) for virtual implantation of the revision system. For that purpose, the CAD-model of the revision cup system was oriented at the pelvic coordinate system and placed in 45° lateral abduction and 20° anteversion.



Fig. 7. CAD-prototype of the newly developed acetabular cup revision system with modular adaptable lateral flap and fixation pegs

Subsequently, the acetabular bone stock was prepared by Boolean subtraction of the implant volume from the pelvic volume. The hemi-pelvis was meshed with 10-node tetrahedral finite elements, the acetabular revision system as well as the liner were meshed with 8-node hexahedral finite elements and the prosthetic head was modeled as an analytical rigid surface (Preprocessor: Patran, MSC, CA, USA). A stick-slip contact formulation with a friction coefficient of $\mu = 0.6$ was assigned between implant and bone. The assembled FE-mesh is shown in Fig. 8. Constraints were applied at the sacroiliac joint (fully constrained) and the pubic symphysis (symmetry boundary conditions).

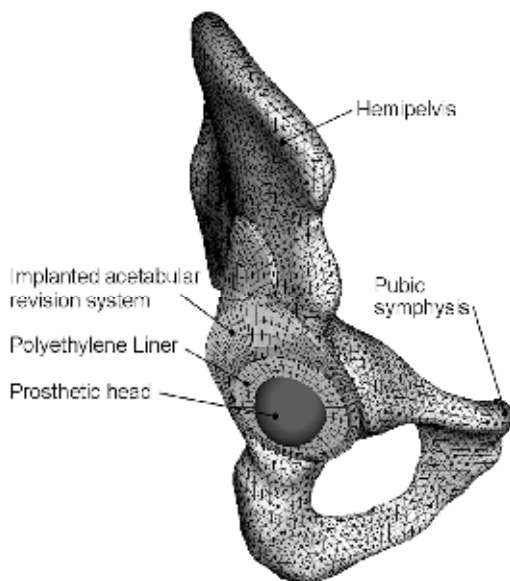


Fig. 8. Assembled finite-element-mesh of the acetabular cup revision system and the surrounding bone. Instead of meshing, the prosthetic head was modeled as an analytical rigid surface.

The material model of the bone was based upon the bone mineral distribution derived from the CT-scans. Our custom-made bone mineral density phantom was CT-scanned together with the analyzed pelvis. The Hounsfield Units (HU) in 100 voxels of each of these chambers in the CT scans were measured in order to gain the analytical relationship between the attenuation, measured in HU, and the bone mineral density, measured in Calcium-Equivalence (ρ_{Ca-Eq}):

$$HU = 1,4997 \rho_{Ca-Eq} \quad (1)$$

Using equation (1), the material model for pelvic trabecular bone from Dalstra et al. (Dalstra, Huiskes et al. 1993) could be applied. Calcium-Equivalence ρ_{Ca-Eq} was converted to Apparent Density ρ_{app} :

$$\rho_{app} = \frac{\rho_{Ca-eq}}{0,626} \quad (2)$$

This conversion is followed by the correlation of Young's modulus of trabecular bone E_{trab} and Apparent Density ρ_{app} (Dalstra, Huiskes et al. 1993):

$$E_{trab} = 2017,3 \rho_{app}^{2,46} \quad (3)$$

In the FE-model, Apparent Density ρ_{app} was mapped from the CT scans onto each node and treated mathematically as temperatures. Using a temperature-dependent material model, the distribution of stiffness in trabecular and subchondral bone was calculated as a function of Apparent Density ρ_{app} respectively Calcium-Equivalence ρ_{Ca-Eq} . Areas with $\rho_{Ca-Eq} > 700$ were assigned a constant Young's modulus of $E_{cort} = 8.5$ GPa and treated as cortical bone (Klueess, Souffrant et al. 2009).

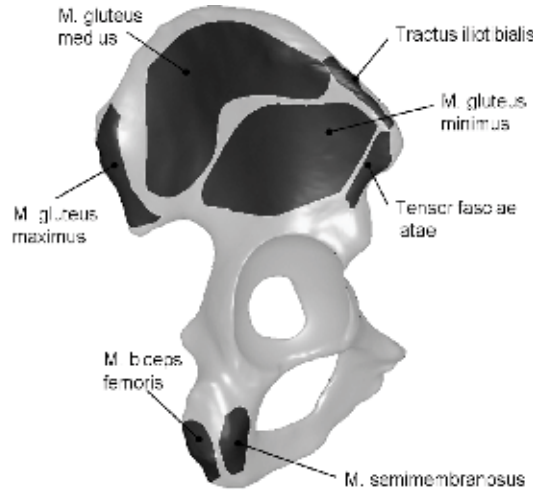


Fig. 9. Attachment areas of simulated muscle forces in the FE-model acting at the hip joint.

The muscle forces acting at the hip joint as well as the hip force resultant during normal gait and during stair climbing were extracted from telemetric measurements (Bergmann, Graichen et al.

1993; Bergmann, Deuretzbacher et al. 2001) and corresponding calculations (Heller, Bergmann et al. 2001). The acting muscles were concentrated into separate groups according to Fig. 9. The amounts of the distributed force vectors in the FE-model are summarized in Tab. 2, given in percentage of the patient's bodyweight (80.0 kg respectively 748.4 N). The hip force resultant was applied in the reference node of the analytical rigid surface located in the center of the reconstructed hip joint.

Muscles	Normal gait			Stair climbing		
	F _x (%BW)	F _y (%BW)	F _z (%BW)	F _x (%BW)	F _y (%BW)	F _z (%BW)
Abductors (Mm. glutei maximus, medius, minimus)	-58.0	-86.5	-4.3	-70.1	-84.9	-28.8
M. biceps femoris	-0.6	-4.1	0.5	-	-	-
Tensor fasciae latae	-7.2	-13.2	-11.6	-3.1	-2.9	-4.9
M. semimembranosus	-17.6	-57.4	-0.8	-	-	-
Tractus iliotibialis	-	-	-	-10.5	-12.8	3
Hip force resultant F _R	54	229.2	-32.8	59.3	236.3	-60.6

Table 2. Applied muscle forces and hip force resultant during normal gait and during stair climbing. The amounts are given in percentage of the patient's bodyweight (%BW) (Bergmann, Graichen et al. 1993; Bergmann, Deuretzbacher et al. 2001; Heller, Bergmann et al. 2001)

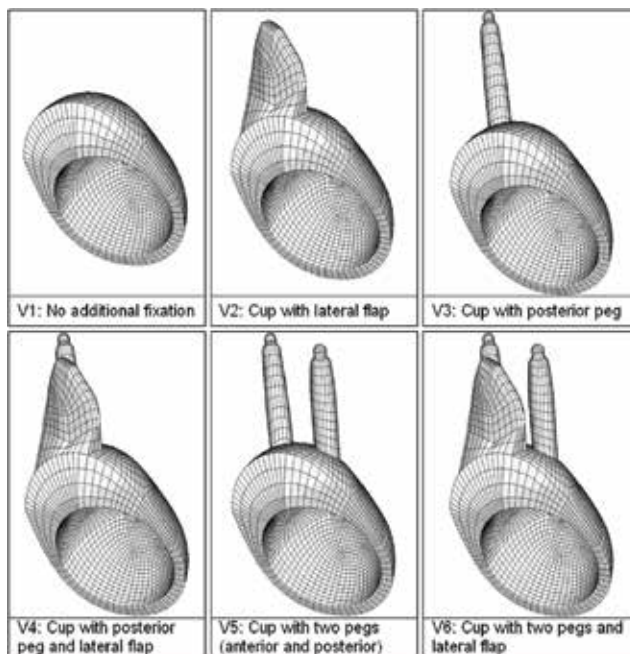


Fig. 10. Variations of the acetabular cup revision system concerning application of the lateral flap and the anterior end posterior fixation pegs. Each variant was virtually implanted into the hemipelvis and meshed with hexahedral elements.

In order to evaluate the influence of the fixation elements, six FE-models of the implant-bone-compound were generated with different variants of the modular acetabular cup revision system. Therefore, different combinations of the lateral flap and the anterior and posterior fixation pegs were virtually implanted, according to Fig. 10.

The structural response of the implant-bone compound to the applied forces was calculated nonlinearly using the solver Simulia/ABAQUS V 6.7 (Dassault Systèmes, RI, USA). The micromotion, being defined as the relative motion between implant and adjacent bone, was evaluated element-wise at the cup-bone interface.

3.3 Results

The micromotion of the cup revision system in the surrounding bone stock was evaluated using the relative tangential node displacements in the contact surface. The postprocessor ABAQUS enables the prediction of tangential displacements (CSLIP) in the two perpendicular directions t_1 und t_2 throughout the whole surface of the implant bed. The maximum amounts of micromotion u_{mic} were calculated in each finite element n by equation (4):

$$u_{mic}(n) = \sqrt{[CSLIP1(t_1, n)]^2 + [CSLIP2(t_2, n)]^2} \quad (4)$$

The results of the FE-analyses were plotted as a greyscale-figure for visualization. The micromotion under maximum loads during gait considering different combinations of fixation elements applied to the acetabular revision system are shown in Fig. 11.

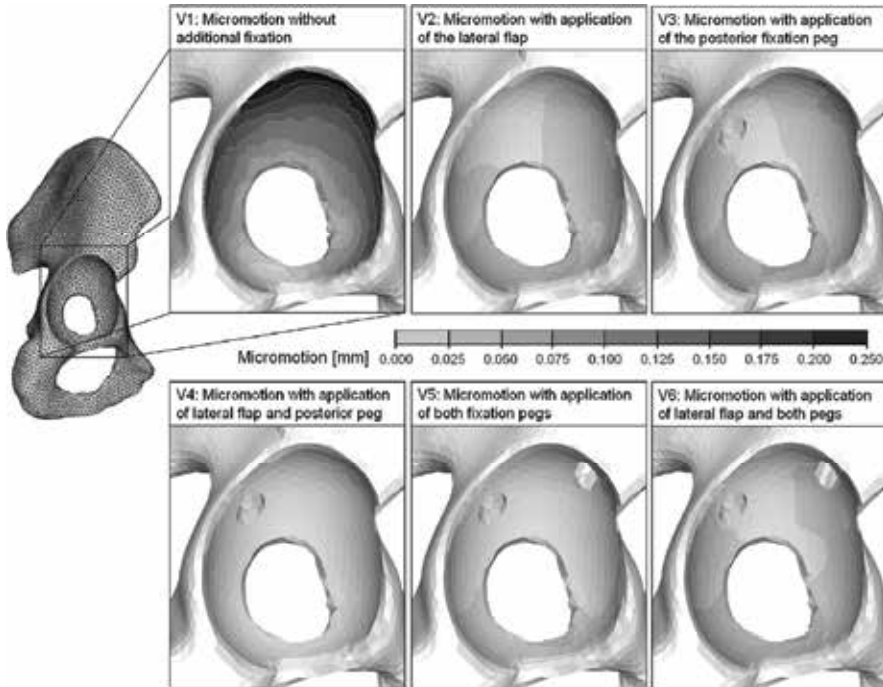


Fig. 11. Greyscale plot of the micromotion in the acetabular bone stock with a D'Antonio defect type III under the maximum loads and muscle forces during gait. Simulating the application of additional fixation elements, the micromotion was reduced significantly.

The interpretation of the results included firstly the maximum micromotion in the implant-bone interface, secondly the percentage of the contact surface with less than 50 μm micromotion.

The calculation of the revision cup without additional fixation (V1) showed high motion of the implant amounting up to 264 μm . Allowable micromotion of < 50 μm was found in the caudal acetabulum only within 5.7 % and 6.1 % of the whole contact surface during gait and during stair climbing. The application of the lateral flap (V2) resulted in a significant reduction of micromotion, especially in the cranial acetabulum. However, the caudal bone bed showed motion of up to 200 μm . Using the posterior fixation peg instead of the flap (V3) decreased the maximum micromotion to values below 100 μm . The areas of micromotion < 50 μm amounted 79 % of the whole contact area during gait and 84 % during stair climbing. When the posterior peg and the flap were used (V4), these areas increased up to 87 % and 98 %. The highest reduction of micromotion was reached using both fixation pegs (V5). In this case, the calculation showed no micromotion higher than 50 μm during gait and only 1 % of the contact area with micromotion higher than 50 μm during stair climbing. When the lateral flap was attached in addition to both pegs (V6), the amount of micromotion slightly increased.

Besides the evaluation of micromotion, the FE-analysis gives information about the expected stress states in the periprosthetic bone during load. Hence, bone regions with sparse amounts of stress which are subject to bone resorption can be identified. An adequate force transmission from the implant to the surrounding bone should enhance bone ongrowth as a result of reconstructed hip biomechanics. In the undesirable case of stress shielding, bone resorption could induce implant loosening. The present analysis was interpreted especially with regard to the bony region around the fixation pegs.

The greyscale plots in Fig. 12 show the stress states within the bone stock of the posterior fixation peg before and after implementation. The stress state without the peg clearly shows the cortical load transmission according to the sandwich-architecture of flat bones. If the fixation peg is inserted, the von-Mises stresses in the trabecular bone between the cortical walls increase up to 5.0 MPa. These stresses are an indicator for possible bone formation in the cranial part of the acetabulum as a result of stress-induced remodeling.

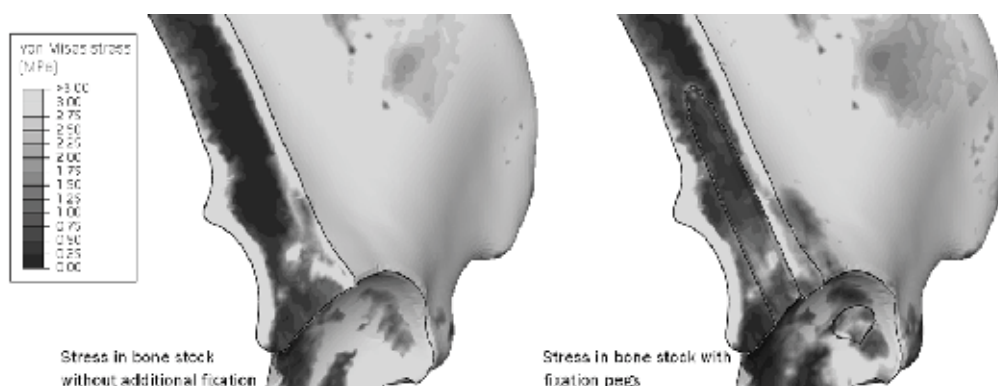


Fig. 12. Cross-sectional view of the von Mises stresses in the posterior ilium under maximum loads during gait. Stress plot with fixation pegs (right) shows higher internal stress in the trabecular bone than calculated without pegs (left).

4. Conclusion

Finite-element-analyses of the implant-bone-compound gained growing accuracy with the fast-paced development of computer workstations. There are a number of reasons why a biomechanical problem should be addressed and solved numerically using FEA:

1. The problem cannot be analyzed clinically:
 - Preclinical stage of implant development
 - Variation of a parameter which does not yet exist in implants on the market
 - Clinical study might not be sufficient to answer specific questions
 - Ethical reasons (e.g. dislocation modes of endoprostheses cannot be conducted in vivo)
2. The problem cannot be analyzed experimentally:
 - Analyzing many different parameters necessitates manufacturing of many prototypes and might be very cost-intensive
 - Boundary conditions such as muscle forces cannot be realized in an experimental setup
 - Experiments on human specimen may not be reproducible due to variations of biologic properties
3. Results from a clinical or experimental study cannot be interpreted
 - A numerical model can help interpreting clinical and experimental results

We presented an approach of how finite-element-models of the implant-bone-compound can be generated, and we gave an example of an FE-analysis of a new implant specifically designed to withstand loads and micromotion in large bone defects due to revision surgery. There is a variety of specifications made for modeling e.g. the pelvis in numerical analyses. The inhomogeneous material properties of trabecular bone as well as the varying thickness of the pelvic compacta should be considered (Anderson, Peters et al. 2005). For economic reasons, a hemipelvis can be modeled under symmetrical boundary conditions (Dalstra, Huiskes et al. 1995; Spears, Pfeleiderer et al. 2001; Kaku, Tsumura et al. 2004; Oki, Ando et al. 2004; Anderson, Peters et al. 2005; Bachtar, Chen et al. 2006). The mapping of trabecular bone material data based on computed tomography scans is widely accepted (Dalstra, Huiskes et al. 1993; Dalstra, Huiskes et al. 1995; Anderson, Peters et al. 2005; Manley, Ong et al. 2006). Nevertheless, a dependency of compact bone stiffness and mineral content has not been proved yet.

Our approach to apply a temperature-dependent material model in combination with Hounsfield-Units treated mathematically as temperatures resulted in a good agreement of numerically and experimentally analyzed principal strains. The requirements for consideration of varying cortical thickness and inhomogeneous trabecular bone stiffness were satisfied. By assigning temperatures, the duration of the calculation was significantly decreased compared to assigning different Young's moduli to discrete trabecular volumes. The finite-element-mesh of bone, if meshed with tetrahedral elements, should consist of tetrahedral elements with quadratic interpolation, i.e. at least 10 nodes per element. Besides the considerable increase in accuracy of such elements, the mapping of HU is realized on a higher node density.

It must be noted that development of finite-element-models in Orthopaedic Biomechanics goes far beyond the examples given in this article. Of course, different anatomical regions are being analyzed apart from the hip joint. There are numerous models of the knee joint,

spine, shoulder and many joints and bony regions in the human body (Fig. 13). Additionally, other disciplines close to orthopaedics deal with similar numerical approaches and models, e.g. traumatology and dental implantology.

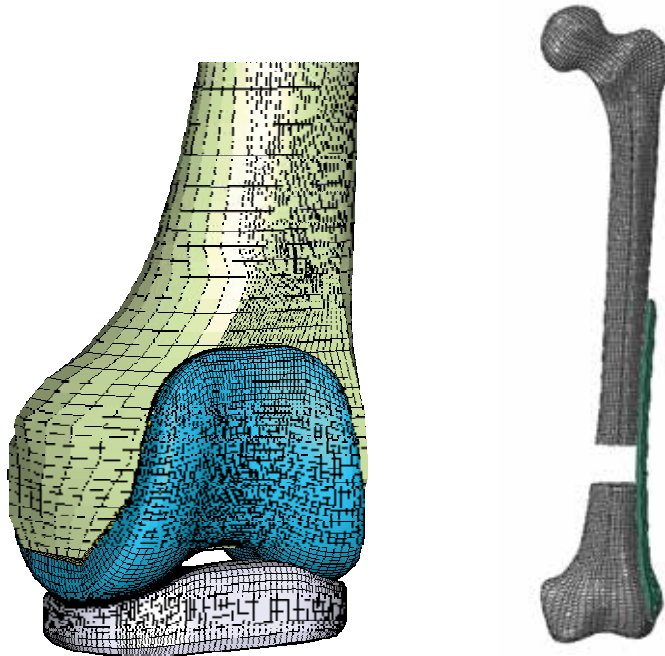


Fig. 13. Finite-element-models of a knee replacement with proximal femur (left) and an osteosyntheses with a segmental femoral bone defect (right)

But also the techniques in numerical simulation are varying. While the models presented in this section are all calculated statically, one also uses dynamic models for special questions. As an example, the impaction of implant components during surgery is in the focus of biomechanical modeling, thereby considering velocity-dependent damping forces as well as acceleration-dependent inertial forces.

Beyond that, more research groups concentrate on predicting biological effects in bone, e.g. the above mentioned bone remodeling initially described by Wolff's law.

In conclusion, finite-element-analysis yields a large potential to help in the development of sophisticated implants, surgery techniques and materials. But as depicted before (Fig. 1), a lively circle between clinical input and provision of analytical results conducted by the clinicians and the engineers is the basic requirement for fruitful research.

5. References

- Aldinger, P. R., A. W. Jung, et al. (2005). Der zementfreie CLS-Titangeradschaft - Langzeitergebnisse, Indikationen und Limitationen. *Akt Traumatol* 35: 320-327.
- Anderson, A. E., C. L. Peters, et al. (2005). Subject-specific finite element model of the pelvis: development, validation and sensitivity studies. *J Biomech Eng* 127(3): 364-73.

- Bachtar, F., X. Chen, et al. (2006). Finite element contact analysis of the hip joint. *Med Biol Eng Comput* 44(8): 643-51.
- Baker, R. (2003). ISB recommendation on definition of joint coordinate systems for the reporting of human joint motion-part I: ankle, hip and spine. *J Biomech* 36(2): 300-2; author reply 303-4.
- Bergmann, G., G. Deuretzbacher, et al. (2001). Hip contact forces and gait patterns from routine activities. *J Biomech* 34(7): 859-71.
- Bergmann, G., F. Graichen, et al. (1993). Hip joint loading during walking and running, measured in two patients. *J Biomech* 26(8): 969-90.
- Chen, W. M., C. A. Engh, Jr., et al. (2000). Acetabular revision with use of a bilobed component inserted without cement in patients who have acetabular bone-stock deficiency. *J Bone Joint Surg AM* 82: 197.
- Civinini, R., A. Capone, et al. (2007). Acetabular revisions using a cementless oblong cup: five to ten year results. *Int Orthop*.
- D'Antonio, J. A., W. N. Capello, et al. (1989). Classification and management of acetabular abnormalities in total hip arthroplasty. *Clin Orthop Relat R* 420: 96-100.
- D'Lima, D. D., S. Patil, et al. (2006). Tibial forces measured in vivo after total knee arthroplasty. *J Arthroplasty* 21(2): 255-62.
- Dalstra, M., R. Huiskes, et al. (1993). Mechanical and textural properties of pelvic trabecular bone. *J Biomech* 26(4-5): 523-35.
- Dalstra, M., R. Huiskes, et al. (1995). Development and validation of a three-dimensional finite element model of the pelvic bone. *J Biomech Eng* 117(3): 272-8.
- Dearborn, J. T. and W. H. Harris (2000). Acetabular revision arthroplasty using so-called jumbo cementless components: an average 7-year follow-up study. *J Arthroplasty* 15: 8-15.
- DeBoer, D. K. and M. J. Christie (1998). Reconstruction of the deficient acetabulum with an oblong prosthesis. *J Arthroplasty* 13(6): 674-680.
- Gallo, J., Z. Rozkydal, et al. (2006). Reconstruction of severe acetabular bone defects using Burch-Schneider cage. *Acta Chir Orthop Traumatol Cech* 73(3): 157-63.
- Garcia, J. M., M. Doblare, et al. (2000). Three-dimensional finite element analysis of several internal and external pelvis fixations. *J Biomech Eng* 122(5): 516-22.
- Gerber, A., M. Pisan, et al. (2003). Ganz reinforcement ring for reconstruction of acetabular defects in revision total hip arthroplasty. *J Bone Joint Surg AM* 85(12): 2358-2364.
- Götze, C., C. Sippel, et al. (2003). Grenzen der zementfreien Revisionsarthroplastik. *Z Orthop Grenzgeb* 141: 182-189.
- Haury, J., F. Raeder, et al. (2004). Defekt-adaptiertes, modulares Revisions-Implantat bei Wechseloperationen. *Pressfitpfannen*. H. Effenberger, L. Zichner and J. Richolt. Grieskirchen, MCU: 221-226.
- Heller, M. O., G. Bergmann, et al. (2001). Musculo-skeletal loading conditions at the hip during walking and stair climbing. *J Biomech* 34(7): 883-93.
- Kaku, N., H. Tsumura, et al. (2004). Biomechanical study of load transfer of the pubic ramus due to pelvic inclination after hip joint surgery using a three-dimensional finite element model. *J Orthop Sci* 9(3): 264-9.
- Keller, T. S. (1994). Predicting the compressive mechanical behavior of bone. *J Biomech* 27(9): 1159-68.

- Kluess, D., R. Souffrant, et al. (2009). A convenient approach for finite-element-analyses of orthopaedic implants in bone contact: Modeling and experimental validation. *Comput Methods Programs Biomed* 95: 23-30.
- Kurtz, S., F. Mowat, et al. (2005). Prevalence of primary and revision total hip and knee arthroplasty in the United States from 1990 through 2002. *J Bone Joint Surg Am* 87(7): 1487-97.
- Malchau, H., P. Herberts, et al. (2002). The swedish total hip replacement register. *J Bone Joint Surg Am* 84-A Suppl 2: 2-20.
- Manley, M. T., K. L. Ong, et al. (2006). The potential for bone loss in acetabular structures following THA. *Clin Orthop Relat Res* 453: 246-53.
- Oki, H., M. Ando, et al. (2004). Relation between vertical orientation and stability of acetabular component in the dysplastic hip simulated by nonlinear three-dimensional finite element method. *Artif Organs* 28(11): 1050-4.
- Perka, C. and R. Ludwig (2001). Reconstruction of segmental defects during revision procedures of the acetabulum with the Burch-Schneider anti-protrusio cage. *J Arthroplasty* 16: 568-574.
- Perka, C., F. Schneider, et al. (2002). Revision acetabular arthroplasty using a pedestal cup in patients with previous congenital dislocation of the hip four case reports and review of treatment. *Arch Orthop Traum Su* 122: 237-240.
- Peters, C. L., M. Curtain, et al. (1995). Acetabular revision with the Burch-Schneider antiprotrusio cage and cancellous allograft bone. *J Arthroplasty* 10: 307-312.
- Rasmussen, J., M. Damsgaard, et al. (2003). Designing a general software system for musculoskeletal analysis. IX. International Symposium on Computer Simulation in Biomechanics, Sydney, Australia.
- Rho, J. Y., M. C. Hobatho, et al. (1995). Relations of mechanical properties to density and CT numbers in human bone. *Med Eng Phys* 17(5): 347-55.
- Schoellner, C. (2004). Die Sockelpfanne - eine Problemlösung für massive acetabuläre Pfannendefekte. *Pfannenrevisionseingriffe nach Hüft-TEP*. C. Perka and H. Zippel. Reinbek, Einhorn-Press Verlag: 211-215.
- Schultze, C., D. Kluess, et al. (2007). Finite element analysis of a cemented ceramic femoral component for the real assembly situation in total knee arthroplasty. *Biomed Tech (Berl)* 52(4): 301-7.
- Siebenrock, K. A., A. Trochsler, et al. (2001). Die Hakendachschale in der Revision schwieriger gelockerter Hüftprothesenpfannen. *Orthopäde* 30: 273-279.
- Snyder, S. M. and E. Schneider (1991). Estimation of mechanical properties of cortical bone by computed tomography. *J Orthop Res* 9(3): 422-31.
- Spears, I. R., M. Pfeiderer, et al. (2001). The effect of interfacial parameters on cup-bone relative micromotions. A finite element investigation. *J Biomech* 34(1): 113-20.
- Springer, B. D., T. K. Fehring, et al. (2009). Why revision total hip arthroplasty fails. *Clin Orthop Relat Res* 467(1): 166-173.
- Thompson, M. S., M. D. Northmore-Ball, et al. (2002). Effects of acetabular resurfacing component material and fixation on the strain distribution in the pelvis. *Proc Inst Mech Eng [H]* 216(4): 237-45.
- Wachtl, S. W., M. Jung, et al. (2000). The Burch-Schneider antiprotrusio cage in acetabular revision surgery: a mean follow-up of 12 years. *J Arthroplasty* 15: 959-963.

- Weise, K. and E. Winter (2003). Revision arthroplasty--acetabular aspect: cementless acetabular bone reconstruction. *Int Orthop* 27: 29-36.
- Whaley, A. L., D. J. Berry, et al. (2001). Extra-Large uncemented hemispherical acetabular components for revision total hip arthroplasty. *J Bone Joint Surg AM* 83: 1352-1357.
- Wolff, J. (1892). *Das Gesetz der Transformation des Knochens*. Berlin, Hirschwald.
- Wu, G., S. Siegler, et al. (2002). ISB recommendation on definitions of joint coordinate system of various joints for the reporting of human joint motion--part I: ankle, hip, and spine. International Society of Biomechanics. *J Biomech* 35(4): 543-8.
- Wu, G., F. C. van der Helm, et al. (2005). ISB recommendation on definitions of joint coordinate systems of various joints for the reporting of human joint motion--Part II: shoulder, elbow, wrist and hand. *J Biomech* 38(5): 981-992.
- Yoon, T. R., S. M. Rowe, et al. (2003). Acetabular revision using acetabular roof reinforcement ring with a hook. *J Arthroplasty* 18: 746-750.
- Zacharias, T. (2001). *Präoperative biomechanische Berechnung von Femur-Hüftendoprothese-Systemen zur Ermittlung der individuellen Primärstabilität nach Roboterimplantation*. Aachen, Shaker Verlag.

Orthopaedic Biomechanics: A Practical Approach to Combining Mechanical Testing and Finite Element Analysis

Rad Zdero, Ph.D.^{1,2,3} and Habiba Bougherara, Ph.D.²

¹ *Martin Orthopaedic Biomechanics Lab, St. Michael's Hospital, Toronto, Canada*

² *Dept. of Mechanical and Industrial Engineering, Ryerson University, Toronto, Canada*

³ *Corresponding Author: zderor@smh.ca*

1. Introduction

Clinicians and engineers have long been interested in assessing the mechanical properties of human whole bones and implant devices to address a vast array of orthopaedic pathological conditions and traumatic injury patterns. To this end, experimental methods and computational techniques have been employed over the years, separately and in combination.

Mechanical in vitro tests on human cadaveric longbones and/or longbone-implant constructs have been done for over a century (Roesler, 1987) (Fig.1). Physiologic loading is a complex interplay of anatomical geometry, material properties, muscle activity, and surrounding soft tissue. As such, experimentalists simplify matters by doing isolated axial compression, lateral bending, torsion, 3-point bending, and 4-point bending tests (Dennis et al., 2001; Fulkerson et al., 2006; Martens et al., 1980, 1986; McConnell et al., 2008; Papini et al., 2007). Although quasi-static load regimes are easier to implement practically, impact and cyclic loads simulate more realistic activities of daily living and injury mechanisms (Fulkerson et al., 2006; Mather, 1968). Pre-clinical experiments have assessed extramedullary and intramedullary fracture fixation constructs, joint arthroplasties, and longbone mechanical properties (Dennis et al., 2001; Fulkerson et al., 2006; McConnell et al., 2008; Papini et al., 2007). More complex in vivo and in situ biomechanical trials have used instrumented devices implanted into human subjects, i.e. telemetry (Bergmann et al., 1993; Schneider et al., 2001). Because of the challenges in using cadaveric bone and living subjects, synthetic bone analogs have become an attractive option. Major advantages include no special storage requirements, low cost, commercial availability, no degeneration over time, standardized geometry, and predetermined material properties (Papini et al., 2007). These surrogates have been validated against human bones with good results for axial, torsional, and 4-point bending stiffness, as well as for cortical and cancellous screw pullout strength (Cristofolini and Viceconti, 2000; Cristofolini et al., 1996; Heiner and Brown, 2001; Papini et al., 2007; Zdero et al., 2008b, 2009a).

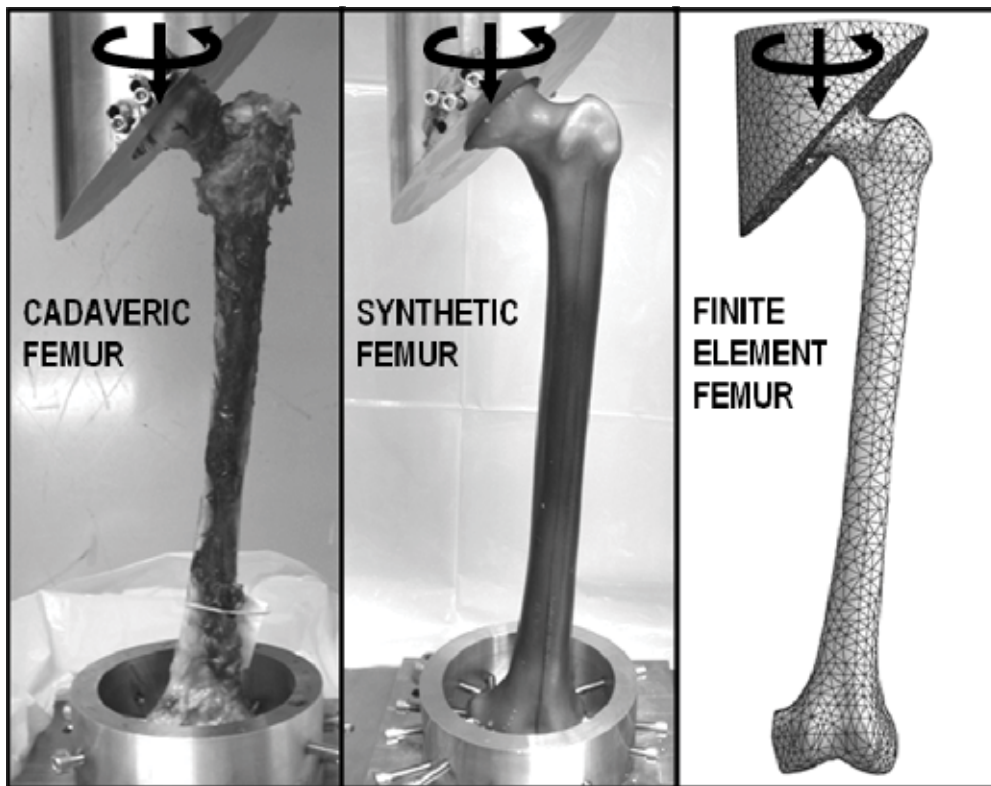


Fig. 1. The assessment of longbone mechanical properties. Experiments on a human cadaveric and a synthetic femur are simulated by a three-dimensional finite element (FE) femur model. Arrows indicate the application of axial compression and torsion.

Synthetic bones have been increasingly used to evaluate the mechanical stability of fracture fixation constructs and joint arthroplasties, as well as being characterized themselves under different loading conditions (Bougherara et al., 2009; Chong et al., 2007; Dennis et al., 2000; Dunlap et al., 2008; Lescheid et al., 2010; Papini et al., 2007; Talbot et al., 2008a, 2008b; Zdero et al., 2008a, 2010a, 2010b).

However, their exact bone quality (i.e. normal, osteopenic, or osteoporotic) has not been conclusively established, and the failure mechanisms of synthetic bone-implant interfaces have not yet been shown to be equivalent to constructs using human host bone. Future advances, therefore, will likely involve characterization and refinement of synthetic bone properties, development of a global open access database of human and synthetic bone material properties and geometries, improvement of experimental instruments and techniques, and implementation of combined loading regimes more representative of actual physiological conditions.

Finite element analysis (FEA) of bone has been done for several decades (Fig.1). The first model of human longbone in the early 1970's assumed two-dimensional geometries and used homogeneous, isotropic, elastic properties (Brekelmans et al., 1972).

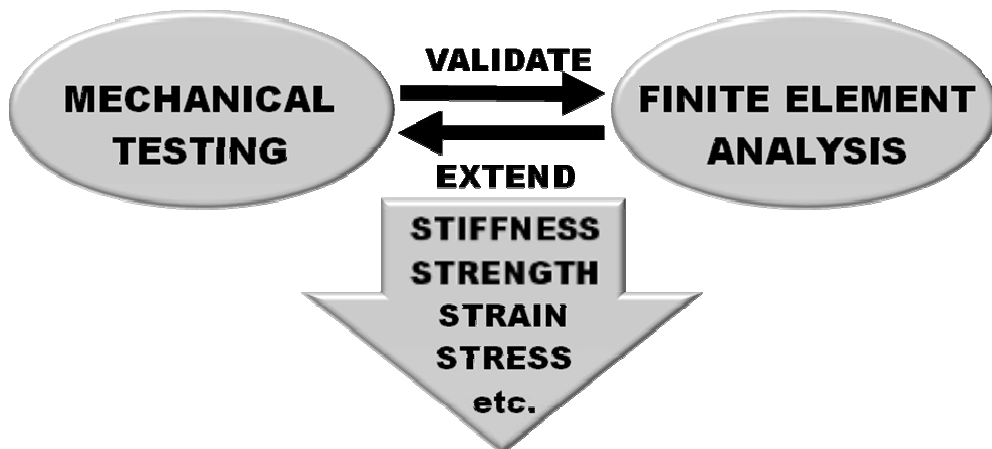


Fig.2. Orthopaedic biomechanics research diagram. The combination of mechanical testing and FEA produces outcome measures of stiffness, strength, strain, and stress.

Later models in the 1970's and early 1980's were three-dimensional, but they often used simplified tube geometries that did not necessarily mimic realistic anatomical structures or complex clinical problems (Cheal et al., 1983, 1984, 1985; Huiskes, 1982; Huiskes and Chao, 1983). Advances in computer hardware, computed tomography (CT), and computer aided design (CAD) now permit the development of more representative geometries and material properties (Bougherara et al., 2009; Cheung et al., 2004; Maldonado et al., 2003; Papini et al., 2007; Sakai et al., 2010; Taylor et al., 1996; Wang et al., 1998; van Rietbergen et al., 1995).

Finite element (FE) modeling and analysis has also been performed on a variety of cobalt chrome, titanium alloy, and polymer-based implant devices, such as longbone fracture fixation plates (Cheal et al., 1983, 1984, 1985; Oh et al., 2010), longbone fracture fixation nails (Bougherara et al., 2009, 2010a; Cheung et al., 2004), knee arthroplasty (Bougherara et al., 2010b), hip arthroplasty (Bougherara and Bureau, 2008; Davis et al., 2009; Huiskes and Stolk, 2005), elbow arthroplasty (Lewis et al., 1996), shoulder arthroplasty (Buchler and Farron, 2004; Orr and Carter, 1985), and so forth. FE models typically treat bone as a continuum, rather than focusing on its microstructural details. This approach is adequate for implant design when relative, rather than absolute, results are used to compare the performance of several devices. However, clinical conditions must be adequately replicated if the absolute quantitative performance of an implant is to be predicted. Many major challenges posed by geometry, material properties, and loading conditions for FE modeling and analysis have been identified and/or addressed in the last several decades. Future advances, therefore, will likely focus on the use of high-resolution CT scans of human bone anatomy, the employment of non-linear material properties for human tissue, the simulation of dynamic physiological loading conditions, the development of a global open access internet database of CAD models and material properties, and the importance of proper verification, validation, and sensitivity analysis (Anderson et al., 2007; Panagiotopoulou, 2009).

Although many researchers continue to legitimately carry out experimentation and FEA separately from one another, the combination of the two has advantages (Fig.2.). Experimentation, on the one hand, is able to validate to what degree an FE model can

successfully replicate more real-world conditions, albeit those of a laboratory setting. FEA, on the other hand, is able to mimic loading conditions and extend data analysis far beyond what might be feasible to perform in a laboratory setting. Two caveats must be kept in mind, however. Firstly, experimental setups should attempt to simulate real-world physiological conditions as much as possible, although their inevitable limitations should be clearly recognized. Secondly, proper verification, validation, and sensitivity analysis of FE models should be performed to ensure that models are working properly. Studies in the biomechanics literature provide tangible examples of how mechanical testing and FEA, given their own drawbacks, can work synergistically to evaluate fracture fixation, joint arthroplasty, and the general mechanical behavior of intact limbs and joints during a variety of activities (Bougherara et al., 2009, 2010a, 2010b; Cheung et al., 2004; Iwamoto et al., 2000; Le Huec et al., 2010; Papini et al., 2007; Zdero et al., 2008b, 2010a). Most of these reports, however, assume that the reader is a somewhat informed and/or experienced experimentalist and FE analyst. As such, many details regarding technique that may be obvious to the author, but not necessarily to the reader, go unreported. Moreover, to the authors' knowledge, there is little in the way of explicit instructional material that addresses the issue of how mechanical testing and FEA can work together.

The aim of this article, therefore, will be to present some practical tools to students, engineers, and clinicians for combining mechanical experiments and FEA to successfully carry out orthopaedic biomechanics research. Specifically, an overview will be given of the purpose, principles, outcome measures, resources required, data assessment, and the various tips, tricks, and troubles when performing mechanical testing and FEA as independent research modalities. This will be followed by a practical case study of a real-world orthopaedic biomechanics problem in which mechanical assessment and computational analysis will be used in combination to generate results.

2. Mechanical Testing

2.1 Purpose

The purpose of in vitro biomechanics testing is to generate valuable pre-clinical information to the researcher and clinician (Fig.3). It allows for the mechanical characterization of intact human joints, longbones, and soft tissues. It assesses the performance of joint prostheses and bone-implant fracture fixation constructs. It allows engineers and surgeons to make improvements to the design of a potentially new implant or surgical technique under controlled laboratory conditions. It helps investigators to decide whether or not to expend the time, energy, resources, and finances on subsequent clinical studies in live animal or human subjects. It prevents the premature introduction of new orthopaedic devices into the market and their implantation into patient populations. In this regard, there has been a recent recognition of the need for an evidence-based approach in bringing new orthopaedic devices to market, which follows a sequential hierarchy of evaluation approaches, starting with basic biomechanical testing (Schemitsch et al., 2010).

2.2 Principles

Several principles should be considered when designing and performing biomechanical tests with respect to specimen factors, testing conditions, and measurement instruments (Fig.3) (Baird, 1962; Holman, 2000).

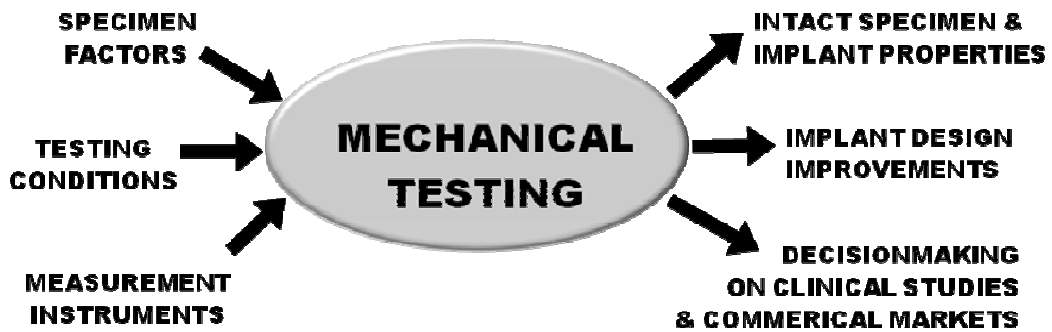


Fig. 3. Overview of mechanical testing shows required inputs and resulting outputs.

Specimen material properties of cadaveric human, cadaveric animal, or synthetic specimens should be similar to in vivo tissue whenever possible. Specimen baseline properties should be established for control specimens or control groups. The number of specimens per test group when statistical comparisons will be made should be sufficient to detect all differences actually present, i.e. to avoid type II error. Test conditions, such as loading modes, loading levels, and specimen fixation devices, should approximate in vivo conditions whenever possible.

Test conditions should be identical from specimen to specimen, unless otherwise required. Test conditions should have precedent in the scientific literature and/or abide by established technical standards, like the ISO (International Standards Organization), ANSI (American National Standards Institute), and ASTM (American Society for Testing and Materials International), as well as conforming to the ethical standards established by institutional and organizational committees if human or animal specimens are used. Measurement instruments should interfere only minimally with the test setup and specimen. Measurement instruments should be calibrated for accuracy and precision and their range and resolution identified. Measurement instruments and methods should be performed and reported so they are reproducible by others investigators.

2.3 Outcome Measures

Some questions should be asked before choosing which mechanical outcome measure(s) will be assessed experimentally. Is it practically feasible to measure? Is it clinically relevant? Is it reproducible by other investigators? Is it sufficient alone or must it be supplemented by other parameters? Although there are a variety of measurements made in biomechanics studies, several are of particular importance. *Stiffness* is the force required to generate a unit of displacement prior to any permanent damage to a specimen, i.e. the slope of the linear portion of the force-vs-displacement curve. *Strength* from a mechanical standpoint is the maximum force needed to cause structural failure of a specimen, i.e. the absolute peak force of the force-vs-displacement curve. Alternately, strength from a clinical standpoint is the force needed to reach a physiologically relevant maximum displacement, i.e. the force corresponding to the pre-determined clinical displacement criterion. *Strain* is the relative deformation per unit length at a location on the test specimen. *Stress* is the amount of relative force per unit area at a location on the test specimen.

2.4 Equipment Required

Several pieces of equipment are vitally important for the biomechanical experimentalist. A mechanical tester such as Instron (Instron Corp., Norwood, MA, USA) or MTS (MTS Systems Corp., Eden Prairie, MN, USA) should be obtained with both axial and torsional load application capabilities. Uniaxial and rosette strain gages accompanied by a dedicated data acquisition system measure specimen surface strains that can be converted to stress values, e.g. strains at several points along the lateral surface of a tibia undergoing axial compression. Pressure sensitive layers like Fuji film (Fuji Prescale USA, Madison, NJ, USA) and K-scan (Tekscan Inc., South Boston, MA, USA), respectively, measure the static and dynamic contact stresses and areas at interfaces, e.g. an intact knee or a total knee prosthesis. linear velocity differential transformers such as the LVDT (Macro Sensors, Pennsauken, NJ, USA) and angular displacement transducers like the ADT (Trans-Tek, Ellington, CT, USA) measure the relative motion between any two bodies, e.g. the relative movement of adjoining bone fragments at a fracture site.

2.5 Data Assessment

Data analysis will be approached depending on the type of study performed. *Statistical studies* are used to determine definitively whether or not outcome measures are different between test groups representing pathological conditions, injury patterns, or normals. A power analysis should be done (preferably before the study) to ensure an adequate number of specimens per group to permit detection of all the statistical differences that actually exist, i.e. avoid type II error. *Non-statistical studies* are usually used to detect patterns or relative changes of an outcome measure on the same specimen under different conditions. However, when comparing two or more specimens, each of which represent a pathology, injury, or normal subject, a definitive statement about statistical differences cannot be made. *Data correlation* using the Pearson coefficient of determination (R^2) is a common approach. As a general guideline, there is low correlation if $0 < R^2 < 0.4$, there is moderate correlation if $0.4 < R^2 < 0.8$, and there is high correlation if $0.8 < R^2 < 1$. It can be employed to determine if two outcome measures are strongly interdependent, e.g. bone density and ultimate failure stress. It is also used to describe the linearity of the data scatter during a measurement process, e.g. if specimens were kept within the linear elastic region during stiffness tests based on the force-vs-displacement curve.

Data display can be done using one of two formats. *Absolute values* allow for comparison of results to known biomechanical threshold values, e.g. material yield stress, or clinical threshold values, e.g. prosthesis failure displacement. The researcher, however, can still highlight certain findings by reporting the percentage change seen for a particularly important outcome measure. *Normalized values* present a direct comparison of results to a control group, e.g. normal subjects, or a control specimen, e.g. intact longbone. In this case, the researcher should still report the absolute values of the control group to which all other data have been normalized. Whether data are given as absolute or normalized values, it is important to present enough data to allow for adequate interstudy comparison.

Data interpretation of results from in vitro biomechanics studies cannot provide or predict the absolute values of in vivo clinical conditions. Rather, mechanical testing gives results that allow for an understanding of the relative performance to one another of the test groups or individual specimens assessed, e.g. the failure strength of implant A versus the failure

strength of implant B. Even so, it is likely that the relative performance can be generalized to actual physiological conditions in vivo.

2.6 Tips, Tricks, and Troubles

Long investigation times can change the mechanical properties of cadaveric bone and bone-implant constructs if the study lasts several weeks or months. For example, a 30% decrease in axial stiffness of femur-implant constructs at an applied load of 500 N can occur over several months during the experimentation period (McConnell et al., 2008). This may happen if the bones are fresh, stored in a freezer, and repeatedly refrozen and re-thawed because all experiments could not feasibly be performed within a shorter period of time. Additionally, implant “migration and settling” within the host bone may also be a factor in this regard as cumulative damage is incurred over time from test-to-test.

Embalming of cadaveric bones and bone-implant constructs could alter the mechanical properties of bone. Several prior investigations, however, have demonstrated that embalming does not significantly affect bone’s mechanical properties (McElhaney et al., 1964; Wainer et al., 1990). Bone mineral density levels measured using dual x-ray absorptiometry (DEXA) reach a steady-state value and remain unchanged, but only if the specimens have been embalmed for several months (Lochmuller et al., 2001). The correlation between bone mass and strength appears to be unaffected by embalming (Edmondston et al., 1994). Given the fact that embalmed specimens, unlike fresh frozen specimens, do not require any special storage requirements beyond the bones being placed in a plastic bag at room temperature, they may be a feasible choice for biomechanical experimentation.

Soft-tissue stripping of muscles, tendons, ligaments, and other tissues allows investigators to isolate cadaveric bone for easier preparation and mechanical testing. This greatly simplifies experiments. However, this ignores the important immediate and longterm role of surrounding tissues in generating forces and moments that influence the stability of joints and bone-implant constructs, as well as bone fracture healing and union. For instance, during the stance phase of normal walking at a speed of 1.2 m/sec, the three tibio-femoral joint force peaks are due primarily to the activity of the hamstrings (3.3 times body weight), the quadriceps (2 times body weight), and the gastrocnemius (2.5 times body weight) (Schipplein and Andriacchi, 1991). Even so, numerous prior investigators have used cadaveric bone stripped of all tissue to successfully assess the relative levels of mechanical stiffness and strength achieved for intact and/or instrumented femurs (Davis et al., 2008; Dennis et al., 2001; Fulkerson et al., 2006; McConnell et al., 2008; Papini et al., 2007).

Loading modes most commonly utilized in testing cadaveric and synthetic longbones and longbone-implant constructs are axial compression, lateral bending, 3- or 4-point bending, and torsion (Bougherara et al., 2009; Cristofolini and Viceconti, 2000; Cristofolini et al., 1996; Davis et al., 2008, 2009; Dennis et al., 2000, 2001; Fulkerson et al., 2006; Martens et al., 1980, 1986; McConnell et al., 2008; Papini et al., 2007; Talbot et al., 2008a, 2008b; Zdero et al., 2008a, 2010a, 2010b). These are meant to simulate the various loading orientations of longbone that might be encountered during common activities of daily living, such as walking, running, jumping, stair climbing/descending, sitting/rising from a chair, etc. Needless to say, there is a large variety of possible conditions that longbones might experience during injuries and accidents which may be difficult to predict and simulate experimentally. As such, a large proportion of experimental biomechanics research attempts

to mimic common activities such as level walking, in which the femur is primarily in a state of axial compression (Duda et al., 1998; Taylor et al., 1996).

Quasi-static tests are used by many investigators to determine the stiffness and strength of longbone or longbone-implant constructs (Bougherara et al., 2009; Cristofolini and Viceconti, 2000; Cristofolini et al., 1996; Davis et al., 2008, 2009; Dennis et al., 2000; Martens et al., 1980, 1986; McConnell et al., 2008; Papini et al., 2007; Talbot et al., 2008a, 2008b; Zdero et al., 2008a, 2010a, 2010b). However, the influence of cyclic fatigue should also ideally be considered when appropriate, since it can cause microdamage and micromovement at the various interfaces created by contact between cement, screws, wires, plates, nails, and arthroplasties with the host bone, thereby potentially leading to decreased longbone-implant stiffness and failure strength (Dennis et al., 2001; Fulkerson et al., 2006; Talbot et al., 2008a). Additionally, some investigators have given attention to simulating impact loading conditions that might occur during falls, which may be particularly relevant for the elderly, or motor vehicle accidents, which may be especially salient for young people with good bone stock (Alho, 1996; Elstrom et al., 2006; Mather, 1968).

Strain gage placement will affect the reliability and reproducibility of results. Gages should ideally be positioned on flat and smooth surfaces, so they will be able to optimally detect the planar strains at that location. Conversely, curved and textured surfaces under load could produce strains over an arc length that would result in an unstable strain gage signal. Moreover, gages should be positioned away from structural discontinuities where there is a steep strain gradient, such as fracture sites and rigid fixation systems, whose proximity could render the gage readings meaningless.

Interspecimen variability between cadaveric tissue with respect to material properties and geometry is a drawback. For instance, axial and torsional stiffnesses for intact human femurs can vary by 3.3 and 3.2 times, respectively (Papini et al., 2007), while for intact human scapholunate ligaments variation can be 28.7 and 9.7 times, respectively (Zdero et al., 2009b). This can pose a challenge in creating a test apparatus or longbone-implant construct with boundary conditions that are able to apply or absorb forces in an identical manner for each specimen. Specimen-to-specimen variability can also lead to a large variation in measured results which are sometimes awkward to accommodate even by normalization relative to control group results. This makes it difficult to draw useful conclusions from a given study. Moreover, researchers may also be faced with the challenge of appropriately comparing their results with previously published investigations.

Synthetic longbones are being increasingly used and have afforded researchers a number of advantages over cadaveric specimens (Cristofolini and Viceconti, 2000; Cristofolini et al., 1996). Synthetic longbones have a standardized geometry that is 20-200 times more uniform than cadaveric bones. There are no special storage requirements. There is no degeneration over time that affects the mechanical properties. Toxicity and financial cost are low. They are commercially available. Moreover, these bones have been shown to replicate human bone axial stiffness, torsional stiffness, 4-point bending stiffness, cortical screw pullout strength, and cancellous screw pullout strength (Cheung et al., 2004; Cristofolini and Viceconti, 2000; Cristofolini et al., 1996; Heiner and Brown, 2001; Papini et al., 2007; Zdero et al., 2008b, 2009a). As such, a number of experimental investigations using these analogs have been reported in the literature to evaluate intact and instrumented longbones (Bougherara et al., 2009; Dennis et al., 2000; Lescheid et al., 2010; Talbot et al., 2008a, 2008b; Zdero et al., 2008a, 2010a, 2010b). However, it has not been shown conclusively that the failure mechanisms of

implants in synthetic host bone are equivalent to those found in cadaveric host bone, therefore, failure strength measurements must be considered cautiously.

Clinical fracture patterns are often complex, resulting in multiple fragments, bone loss, and jagged fracture patterns (OTA, 1996). The interdigitating nature of complex fracture patterns may provide enhanced stability, since interfacial slip at the fracture site between adjoining segments of bone may be reduced. In contrast, experimentalists often simulate simpler injury patterns with smooth bone cuts because they are more easily created in the laboratory setting, thereby permitting duplication of multiple test specimens (Dennis et al., 2000, 2001; Fulkerson et al., 2006; McConnell et al., 2008; Talbot et al., 2008a, 2008b; Zdero et al., 2008a). As such, stiffness and strength may actually be higher clinically than measured experimentally, which may provide a conservative lower estimate.

Biological factors cannot be taken into account by in vitro tests. Biological differences between people will contribute to the degree of fracture healing, limb and joint function, and stability after surgical repair of fractures and implantation of joint prostheses (Claes and Ito, 2005; Huiskes and Van Rietbergen, 2005; Perren, 1975). Biological elements include variations in intrinsic healing capacity of various segments of the longbone, bone quality as influenced by osteoporosis and osteopenia, the organization and structure of osteons, and longbone nutritional needs. This interplay between mechanics and biology substantially influences the stress levels experienced by longbone and, thus, the quality of fracture healing between adjoining bone fragments and bony ongrowth around implant devices.

3. Finite Element Analysis

3.1 Purpose

The purpose of Finite Element Analysis (FEA) in the field of orthopaedic biomechanics is to predict the mechanical behavior of bones, develop and improve the design of implants, and eliminate and/or minimize the time and cost involved for in vivo and in vitro experimentation (Fig.4). FE modeling is also used to simulate the bone remodeling and healing phenomena of conventional and biomimetic devices, such as hip replacements (Bougherara et al., 2010c; Huiskes et al., 1987; Hernandez et al., 2000). More recently, computational modeling combined with biochemistry, bone biology, and thermodynamics was used to develop new therapeutic strategies associated with bone loss and other bone related diseases (Bougherara et al., 2010d; Huiskes et al., 2000; Rouhi et al., 2007).

3.2 Principles

The development of FE models in orthopaedic biomechanics starts with the acquisition of data using medical imaging techniques, such as computed tomography (CT) scans and magnetic resonance imaging (MRI) (Bougherara et al., 2007). These scans should be of sufficient resolution to allow for accurate replication of the geometries of interest. These data are then converted into realistic three-dimensional geometries of the bone tissue of interest using medical imaging software. The software should have a known minimal or zero error in the relative positioning of the image slices that make up these three-dimensional reconstructions. To generate the geometry of prostheses that will be inserted into joints or bones, computer aided design (CAD) software can be employed. The exact positioning of the implant into the bone tissue in the FE model should be verified by x-rays to accurately replicate laboratory experiments or clinical scenarios.

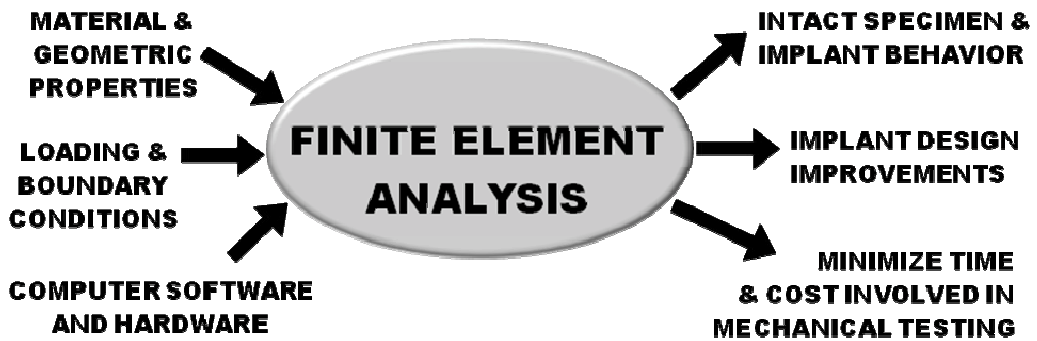


Fig. 4. Overview of FEA shows required inputs and resulting outputs.

At this stage, the assembled geometry is ready to be exported into a software simulation package for analysis. Material properties should be assigned to the bones and implants and physiological load conditions should be applied that are reflective of real-world conditions. FEA should then be performed to obtain a variety of outcome measures, such as stiffness, strength, strain, and/or stress. It should be recognized that the accuracy of the FEA results will rely on the assumptions made to build the FE model. Given the fact that it is very difficult to develop and analyze any model without any assumptions, therefore, it is important to understand to which extent such assumptions may affect the analysis of the results.

Typical assumptions may concern input data, such as material and geometric properties, loading and boundary conditions, subsequent analysis, and the interpretation of results. Each of these has its own set of limitations which will affect the final calculated results.

3.3 Outcome Measures

Typical outcome parameters relevant to clinicians and engineers include bone quality or stiffness, strength, strain, and stress. The combination of FE approaches with numerical algorithms offers potential measurements such as bone density and damage accumulation. In addition, modern FE software packages can offer advanced tools, such as shape and topology optimizations. Any successful FEA must be reproduced by experimental tests. The adequacy of the outcome parameters must be verified by comparison with previous studies and/or clinical observations. Experimentally, strain gages provide a fundamental comparison with FE predictions of average deformations at particular locations on the bone implant surfaces. Thus, FEA is advantageous since it provides a detailed map of strains and stresses of the entire structure, unlike strain gages which provide measurements only at single discrete locations.

3.4 Software and Hardware Required

Three types of software are needed to complete FEA of any structure. Firstly, a computer aided design (CAD) package or a three-dimensional (3D) scanner are needed to create the solid model of the bone or implant. In terms of accuracy, 3D scanners (NextEngine, Inc., Santa Monica, CA, USA) are better for generating CAD models than conventional approaches, which require physical measurements of an object. The 3D scanners

systematically and automatically create highly detailed digital models. Secondly, an FE program package is needed for analysis. Packages such as ANSYS, Inc., ABAQUS, and Altair Engineering combine CAD and FE programs into a single software package, thereby creating a user friendly environment. Thirdly, to generate 3D models of complex geometries, such as bone tissues where density measurements are crucial, medical imaging software (Mimics®, Materialise Group, Leuven, Belgium) processes CT images and transforms them into 3D models. SolidWorks can be used to develop geometries of prosthetic implants. The bone and implant can then be assembled to replicate the experimental setup and exported for analysis to ANSYS Workbench (ANSYS Inc., Canonsburg, PA, USA). Since FE models often need considerable computational time, it is recommended to use a computer known for rapid numerical data processing in order to reduce central processing unit (CPU) time and storage costs.

3.5 Data Assessment

Data analysis in FEA can be either structural or statistical in nature. *Structural analysis* allows the quantification of normal and shear stresses, strains, and strain energy densities, as well as investigating different failure modes of the bone and/or implant under physiological load conditions. *Statistical analysis* enables the prediction of the uncertainty in the modeling variables on the biomechanical behavior of the bone implant system. For example, stochastic approaches such as the Monte Carlo method can be used to evaluate the probability of bone fracture under static loading conditions.

Data correlation of FEA data with experimental data can be done using the Pearson coefficient of determination (R^2). This coefficient is an indicator of the validity of the FE model by determining the level of agreement between FE and experimental results.

Data display of FEA results allows the analyst to observe minimum and maximum stress values, either in nodes or elements, which can be used to measure the risk of bone failure. Conversely, Von Mises stresses (i.e. averaged values of the principal stresses) are used to quantify the load transfer after prosthetic implantation. Other quantities, such as strain energy density values, are used to investigate stress shielding and bone loss. Thus, depending on the parameter investigated, different outcome measures can be displayed.

Data interpretation of FEA results focuses on the assessment of the mechanical performance of implants and/or bones to predict their mechanisms of failure. This can lead to the optimization of prosthetic implants and improved surgical procedures. However, FEA results should be interpreted with caution when used as a basis for clinical treatment.

3.6 Tips, Tricks, and Troubles

Material properties are chosen for bones which sometimes may not be reflective of actual clinical conditions. Often, assumptions of linearity and homogeneity of properties are made to simplify modeling and analysis by simulating commercially available synthetic bones. Although this may initially appear problematic, these analog bones, for instance, can replicate the non-homogeneity of human bone. However, synthetic bones are unable to mimic human bone anisotropy, which affects the bulk mechanical behavior of bones. Nevertheless, a comparison of FE analysis, synthetic femurs, and human cadaveric femurs previously performed suggests that linearity is a fairly good approximation of the behavior of femurs as a whole (Bougherara et al., 2009, 2010a, 2010b; Papini et al., 2007)

Implant positioning of prostheses into host bone can be difficult to accomplish using FE modeling. Malpositioning and the mode of fixation of the implant may affect the load sharing between the implant and host bone tissue (Lotke and Ecker, 1997; Shimmin et al., 2010). To overcome this problem, x-rays of the bone-implant system or computer-aided periprosthetic planning should be used to assist with implant position.

Boundary conditions at the interfacial contact regions may affect the mechanism of load sharing between implants and host bone. In order to achieve perfect osseointegration, i.e. bony ongrowth around the implant, bonded contact at interfaces is usually assumed in FE models. In reality, such a case is unlikely to occur. The incorporation of frictional contact in FE models is important, since it affects the longterm stability of orthopaedic implants (Dammak et al., 1997; Forcione and Shirazi, 1990).

Loading levels considered in most FE models are sometimes much lower than real physiological loads. Most frequently, only static joint forces acting on the bone tissue are considered, while the remaining forces due to the activity of muscles, tendons, ligaments, and other surrounding soft tissues are neglected. This simplifies FEA greatly. However, some researchers have shown that the stress distribution in cancellous bone, since it is prone to stress shielding, is dominated by the vertical component of the joint reaction force (Taylor et al., 1995). In fact, the magnitude of stress shielding and bone resorption may be underestimated by FE models (Simões and Marques, 2005).

Osteolysis or bone resorption, which is due to alterations in the stress distribution of the bone and implant, are often not replicated by FE models. Studies are restricted to analyzing the stress-strain pattern in the bone-implant system. While models of bone adaptation to loading environments exist in published literature (Bougherara et al., 2010c; Huiskes et al., 1987; Taylor et al., 1999; Weinans et al., 1993), no doubt a favorable prediction of bone adaptation will further help assess the potential of implants used in orthopaedics.

Three-dimensional reconstruction of muscles and tendons remains a significant challenge in biomechanics, despite the considerable potential of FEA. It is very difficult to identify the exact locations and intersections of the muscles, as well as their thickness, even from high resolution MRI images (Schutte et al., 2006).

Geometric defects resulting from the reconstruction of bone tissue using CT scans and the subsequent segmentation phase can be optimized using special software, such as Geomatic Studio (Geomatic, Inc., Research Triangle Park, NC, USA), for healing and optimizing complex surfaces. Further topological optimization tools are also available in AnsysWorkbench Suite.

Mesh type, size, and relevance are crucial for the success of FEA simulations. The choice of the mesh type depends on two factors, namely, ease of generation and high quality or accuracy. Hexahedral meshes are quite difficult to generate (Schonning et al., 2009), but offer more accuracy than tetrahedral meshes, particularly for biomechanical applications involving contact analysis. However, tetrahedral meshes can be used if validated with experimental results (Papini et al., 2007). Furthermore, it is important to optimize the mesh sensitivity using mesh refinement tools, in order to obtain a high quality mesh, keeping in mind the computation cost, such as CPU time, data storage, software required, hardware required, human cost, etc.

4. Case Study: The T2 Intramedullary Nail for Fixing Femoral Shaft Fractures

4.1 Background

Intramedullary femoral nails are the most common method of fixing fractures of the femoral diaphysis (Pankovich and Davenport, 2006). The T2 Femoral Nailing System (Stryker, Mahwah, NJ, USA) provides surgeons with a multi-functional nail that can be oriented in antegrade (inserted into the proximal femur end) or retrograde (inserted into the distal femur end) and can be fixed to the femur with standard or advanced locking. Advanced (or dynamic) locking uses oblong holes to permit up to 10 mm of motion in either the distal or proximal end of the femur. Advanced locking allows for load transfer through the fractured ends of the femur and has both locked and unlocked modes by which the movement of the nail within the femur is restrained. Presently, results are given of a retrograde nail (RA-on) and an antegrade nail (AA-on) with advanced locking activated. Outcome measures included axial stiffnesses, strain distributions along the medial surface of the femur, and three-dimensional stress maps of the femur-nail construct. This case study is a brief summary of a more comprehensive prior examination by the current authors of the T2 nailing system implanted into synthetic femurs (Bougherara et al., 2009).

4.2 Mechanical Testing

Two large left Third Generation Composite Femurs (Model 3306, Pacific Research Labs, Vashon, WA, USA) were obtained and instrumented with T2 Femoral Nailing Systems (Stryker, Mahwah, NJ, USA) (Fig.5a). Implantation was done using the manufacturer's protocol by an orthopaedic surgeon. One femur had a nail inserted retrograde with advanced locking activated (RA-on). The other femur had a nail inserted antegrade with advanced locking activated (AA-on). Femurs were not osteotomized, thereby modeling an anatomically-reduced mid-shaft fracture which has begun to heal 6 weeks after surgical repair (Schneider et al., 1990, 2001).

Femurs were instrumented with five general purpose uniaxial linear pattern strain gages (Model CEA-06-125UW-350, Vishay Measurements Group, Raleigh, NC, USA). Gages were spaced 40 mm apart along the medial surface, starting at 20 mm proximal to the lesser trochanter. Wire leads were soldered to each gage, secured to the femur using electrical tape, and attached to an 8-channel Cronos-PL data acquisition system (IMC Mess-Systeme GmbH, Berlin, Germany). This system was attached to a laptop computer for data collection and signal analysis using FAMOS V5.0 software (IMC Mess-Systeme GmbH, Berlin, Germany).

Femurs were secured distally using nine pointed screws inserted into the cortex of the condyles through holes in a stainless steel ring-like jig. The jig was mounted rigidly onto the baseplate of the test machine. All experiments were done on a mechanical tester (MTS 793, MTS Systems Corporation, Eden Prairie, MN, USA).

A vertical compressive load to the femoral head was applied using displacement control (max displacement = 0.5 mm, rate = 10 mm/min). The load applicator was a flat plate. No slippage of the femoral head at the applicator-head interface was visually observed. The slope of the load-vs-displacement curve was defined as axial stiffness. Based on stiffness, a displacement value was then calculated and used as the maximum value for a second test run to reach 580 N, as done previously (Cheung et al., 2004). It was during this second test run that final strain gage values were recorded and averaged over a 30 second period.

4.3 FE Modeling and Analysis

For the T2 nail, SolidWorks 2007 CAD software was used to create a solid model of the same nail used experimentally (Fig.5b and 5c). From a basic cross-section, the “sweep” function was used to extrude the nail geometry. Holes were cut and chamfers were added to simulate the nail. The channels running down the exterior surface of the nail were omitted, and a circular cross-section was maintained along the length of the nail. A second CAD model was created to aid in the ANSYS DesignModeler assembly prior to FEA. This “dummy” model did not have holes cut and did not have a canal running along its length.

For the synthetic femur, CT scans were performed every 0.5 mm along the length of the same synthetic femurs used experimentally (Fig.5b and 5c). Using MIMICS Medical Imaging Software, cross-sectional geometries of the femur were exported into SolidWorks 2007 Suite and used to fit splines representing the boundaries of the cortical and cancellous bones. The splines were imported into SolidWorks and joined together using a “loft” function, resulting in separate cancellous and cortical CAD models. The intramedullary canal was created by cutting away from the cancellous bone.

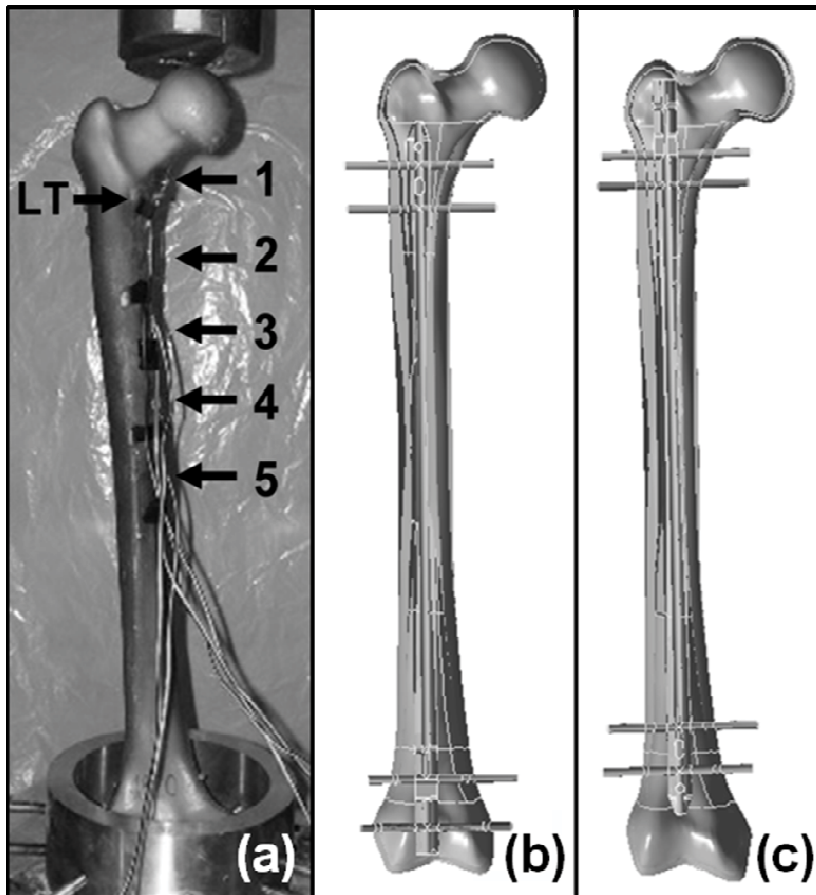


Fig. 5. T2 nail configurations. (a) Experimental setup of femur-nail instrumented with strain gages 1 to 5. (b) CAD model for RA-on nail inserted from the bottom of the femur. (c) CAD model for AA-on nail inserted from the top of the femur. LT is the lesser trochanter.

To assemble the CAD models, the cancellous bone, cortical bone, and T2 models were imported into SolidWorks 2007. The T2 nail was positioned into the intramedullary canal while attempting to minimize intrusion into the cortical bone. The exact position of the nail was verified by x-rays of the experimental specimens. The use of CAD-modeled screws was avoided by simulating partially and fully threaded screws with pins 5.0 mm in diameter. A flat plate load applicator was added to simulate the loading on the femoral head. ANSYS DesignModeler was used to import the SolidWorks 2007 models into ANSYS. A “dummy” nail was imported and a slice operation introduced to remove any cortical and cancellous bone overlapped by the T2 nail. The actual CAD model of the T2 nail was then imported and fitted into the space left by the slice operation. Any unwanted residual elements were suppressed, and the geometry was exported for FEA to be carried out.

Regarding mesh creation, ANSYS Workbench 10.0 was used. For the RA-on nail configuration, the number of nodes and elements was 51,108 and 24,530, respectively. For the AA-on nail configuration, the number of nodes and elements was 48,520 and 23,198, respectively. Body elements used were 10-Node Quadratic Tetrahedron for the cortical bone, cancellous bone, and T2 nail and pins, whereas a 20-Node Quadratic Hexahedron was used for the flat plate load indenter. Contact elements used were Quadratic Triangular for the cortical bone, cancellous bone, and T2 nail and pins, whereas a Quadratic Quadrilateral was used for the load applicator. Mesh relevance for both configurations was > 80%.

Concerning material properties, the T2 nail was set at values for titanium ($E=110$ GPa, $\nu=0.3$). Synthetic femurs were assumed to be isotropic and linearly elastic with material properties for cortical ($E=10$ GPa, $\nu=0.3$) and cancellous ($E=206$ MPa, $\nu=0.3$) based on a previous study on the same bone (Cheung et al., 2004). The value for Young's Modulus of cortical bone ($E=10$ GPa) was an average of compressive (7.6 GPa) and tensile (12.4 GPa) values (Cheung et al., 2004).

FEA was done in the Simulation window of ANSYS Workbench 10.0 suite. The displacement of the distal end of the femur was restrained. This was an approximation of the nine-point jig used in experimental trials. The loading simulated that of the experimental apparatus with a flat plate load applicator applying 580 N at the apex of the femoral head. Contact regions between the T2 nail and synthetic femur were set to frictional (coefficient = 0.08) (Grant et al., 2007). For advanced locking activated, the pins set in dynamic positions were allowed to have frictionless contact with the nail, while maintaining bonded contact to the femur. All other pin and bone contacts were set to bonded. The contact region between the indenter and femoral head was set to bonded, assuming that there would be no slipping.

4.4 Mechanical Testing Results

Axial stiffness results are shown (Fig.6). The AA-on configuration was 1.7 times stiffer than the RA-on configuration. The average linearity coefficient for axial load versus axial displacement data was $R^2 > 0.99$. Thus, specimens remained within the linear elastic region and incurred no permanent damage during tests. Compressive microstrains along the medial surface of the femur ranged from 294 to 968 (Fig.7). As anticipated, the highest strain values occurred for proximal Location 1 (968, RA-on) or Location 2 (820, AA-on) at the base of the femoral neck, since the bone there is composed of a soft cancellous region with only a thin cortical shell. This allows for substantial bending during axial compression.

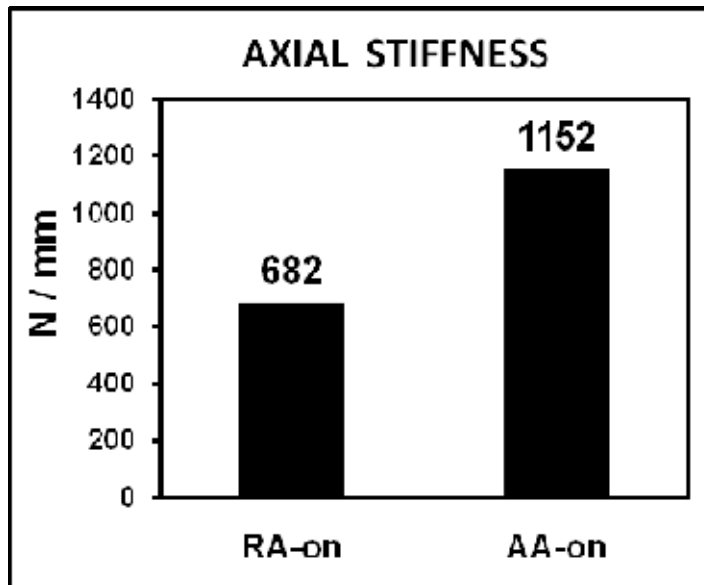


Fig. 6. Experimental stiffness results for RA-on and AA-on configurations. The AA-on femur-nail construct was 1.7 times stiffer than the RA-on femur-nail construct.

The lowest strain values occurred for distal Location 5 (294, RA-on; 302, AA-on) at about the midpoint of the femoral shaft, since the bone there is composed solely of rigid cortical material. This prevents substantial bending during axial compression. A highly linear relationship existed for strain gage values (microstrain) versus strain gage location (mm) along the medial length of the femurs, yielding linear correlation coefficients of $R^2 = 0.88$ (RA-on) and $R^2 = 0.91$ (AA-on). This illustrated the uniformity of material properties and strain along the medial surface of the femur bone.

4.5 FEA Results

FE strain readings were taken at the same strain gage locations as experimentally. Elastic (Von Mises) strain readings were used and represent compressive strains occurring at each location, ranging from 363 to 799 (Fig.7). A linear relationship was noted for the FE results between gage value (microstrain) and location (mm) along the femur, yielding $R^2 = 0.89$ (RA-on) and $R^2 = 0.83$ (AA-on). FE strains agreed reasonably well with experimental strains, showing that the actual line of best fit ($R^2 = 0.89$, slope = 0.69) was similar to the theoretical line of perfect correlation ($R^2 = 1$, slope = 1) (Fig.7). Three-dimensional stress maps demonstrated lower stresses for the RA-on compared to the AA-on configuration (Fig.8). Peak stresses were 7.7 MPa (RA-on) and 8.2 MPa (AA-on) at the locking screws.

4.6 Discussion

Overall there was reasonable agreement between experimental strain data obtained from mechanical tests and FE analysis, thereby validating the FE model. The resulting FE three-dimensional stress maps may serve as preliminary estimates for this fracture repair method once healing begins, albeit at low axial loads.

Compared to previous studies, the present study yielded favorable results despite some differences in methodology. Current axial stiffnesses were 682 N/mm (RA-on) and 1152 N/mm (AA-on) with compressive medial microstrain ranges of 294-968 (mechanical tests) and 363-799 (FEA). Kraemer et al. (1996) loaded intact human cadaveric femurs equipped with the Russell-Taylor nail (axial stiffness = 850 N/mm; microstrain = 975) and a long intramedullary rod hip screw (axial stiffness = 750 N/mm; microstrain = 1125).

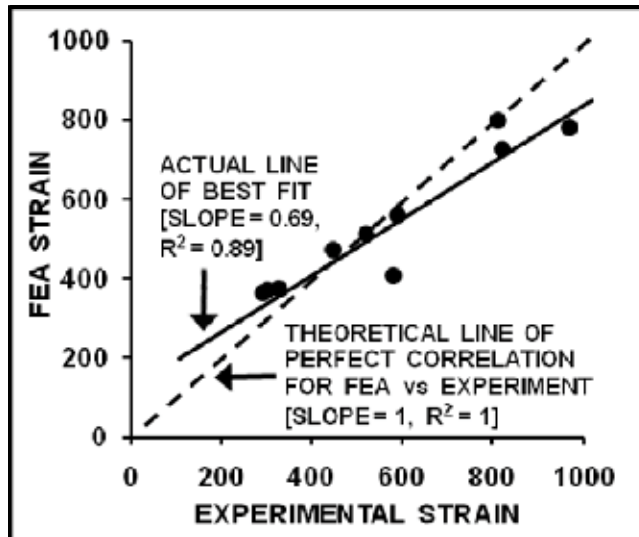


Fig. 7. Correlation of surface strain obtained from experiment and FEA. The actual line of best fit showed a high linear correlation, R^2 , but its slope was only moderately close to that of the theoretical line of perfect correlation. Values are in microstrain units ($\mu\text{m}/\text{mm}$).

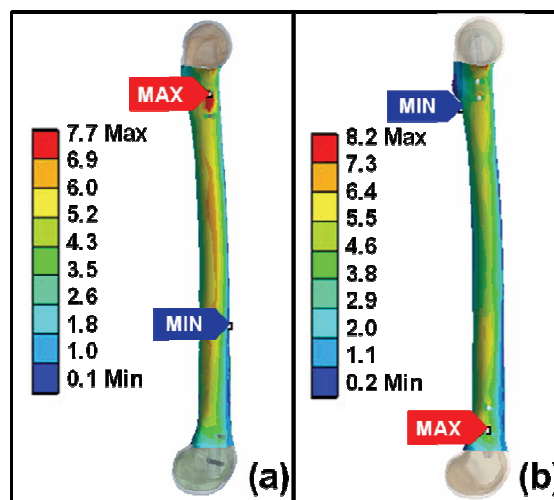


Fig. 8. FEA stress distributions. (a) RA-on configuration stress map. (b) AA-on configuration stress map. Stresses shown are equivalent Von Mises stresses (MPa).

Cheung et al. (2004) tested the same synthetic femur used currently and instrumented it with a Synthes reamed retrograde intramedullary nail, yielding microstrains that ranged from approximately 800 (compressive) to 200 (tensile).

Clinically, the majority of axial load is taken by the femur in a femur-nail construct during walking gait. Cheung et al. (2004) showed that an intact femur instrumented with a fully locked intramedullary nail absorbs about 70 to 74% of total axial force during the stance phase and 91% during the swing phase of gait. Although load sharing was not monitored in the present investigation, similar proportions would be expected because the synthetic femurs used were identical to Cheung et al. and the axial forces applied were almost identical (600 N versus 580 N presently).

The T2 nail offers optimal configurations that may be useful for young healthy and older patients. A previous study using the same test setup showed that axial stiffness was 1290 ± 30 N/mm for these same synthetic femurs without nails, thereby modeling young healthy bone, and 757 ± 264 N/mm for human cadaveric femurs without nails with an average age of 75.4 years (Papini et al., 2007). By matching these intact femur stiffnesses with current data for the instrumented femurs, the AA-on configuration may be optimal for young healthy patient bones, whereas the RA-on arrangement may be more suitable for older patients. However, this suggestion would need to be demonstrated at higher clinical loads.

Stress maps showed higher peak stresses for AA-on compared to RA-on around locking screws. Screw locations could potentially be sites of clinical failure once weight bearing is permitted (Pankovich and Davenport, 2006). The relative similarity of stress distribution along the diaphysis of the two configurations suggests that neither would necessarily be more susceptible to failure at mid-shaft. However, the application of higher clinical forces would need to be performed experimentally to confirm this supposition.

The weaknesses of this study should be recognized. Firstly, low static axial loads were applied experimentally and modeled using FEA. The femur, in reality, is exposed to higher dynamic forces and moments during activities of daily living (Bergmann et al., 1993).

Secondly, the FE model assumed linear isotropic bone properties. In reality, nonlinearity, anisotropy, and viscoelasticity might affect bulk mechanical behavior of femurs. Even so, comparison of FEA, synthetic femurs, and human cadaveric femurs previously performed (Papini et al., 2007) suggests linear behavior is a good approximation.

Thirdly, synthetic, rather than human, femurs were used. However, these synthetic surrogates have been successfully used in investigations dealing with femoral mid-diaphyseal fracture repair constructs (Talbot et al., 2008a, 2008b; Zdero et al., 2008a). They have demonstrated good agreement in axial and torsional stiffness when compared to human cadaveric femurs (Heiner and Brown, 2001, 2003; Papini et al., 2007).

Fourthly, although experimental and FE strains showed reasonable agreement, a clockwise tilt of the actual line of best fit indicates exaggeration of experimental strains. Experimentally, this may have been due to femoral head movement medially under the flat load applicator plate, which would increase the strain values, whereas the FE model restricted femoral head movement to only the vertical axial direction.

Finally, load to failure, i.e. strength, was not currently assessed. This would be an important parameter to measure to fully understand the potential benefits and limitations of the T2 nail. Even so, the investigation was able to identify locations of high stress around locking screws, which could be potential sites of failure at clinical loads.

In conclusion, the RA-on nail may be more suitable for the elderly, whereas the AA-on nail may be more optimal for the young. The results simulated post-operative stability at low static axial loads once fracture union begins to occur at the mid-shaft fracture site.

5. Conclusion

In orthopaedic biomechanics, an assessment of the mechanical properties of human bones and implants can be done using mechanical experimentation combined with FEA. A practical overview of the methodology was presently given. The outcome measures discussed were stiffness, strength, strain, and stress. A real-world case study was also examined. The authors hope that students, engineers, and clinicians will now be better prepared to engage in more effective orthopaedic biomechanics research.

6. References

- Alho, A. (1996). Concurrent ipsilateral fractures of the hip and shaft of the femur: a meta-analysis of 659 cases. *Acta Orthop Scand*, Vol.67, No.1, 19-28.
- Anderson, A.E., Ellis, B.J., and Weiss, J.A. (2007). Verification, validation and sensitivity studies in computational biomechanics. *Comput Methods Biomech Biomed Engin*. Vol.10, No.3, 171-84.
- Baird, D.C. (1962). *Experimentation: An Introduction to Measurement Theory and Experiment Design*, Prentice-Hall, Englewood Cliffs, USA.
- Bergmann, G., Graichen, F., and Rohlmann, A. (1993). Hip joint loading during walking and running measured in two patients. *J Biomech*, Vol.26, No.8, 969-990.
- Bougherara, H., and Bureau, M.N. (2008). Biomimetic composite-metal hip resurfacing implant. *Research Letters in Materials Science*, June 8, No. 368985.
- Bougherara, H., Bureau, M.N., Campbell, M., Vadean, A., and Yahia, L. (2007). Design of a biomimetic polymer-composite hip prosthesis. *J Biomed Mater Res A*, Vol. 82, No.1, 27-40.
- Bougherara, H., Zdero, R., Miric, M., Shah, S., Hardisty, M., Zalzal, P., and Schemitsch, E.H. (2009). The biomechanics of the T2 femoral nailing system: a comparison of synthetic femurs with finite element analysis. *Proc Instn Mech Engrs (Part H): J Engineering in Medicine*, Vol.223, No.H3, 303-314.
- Bougherara, H., Miric, M., Zdero, R., Shah, S., Hardisty, M., Zalzal, P., and Schemitsch, E.H. (2010a). A novel biomimetic intramedullary nailing implant for the treatment of femur fracture. *International Journal of Engineering*, in press.
- Bougherara, H., Zdero, R., Mahboob, Z., Dubov, A., Suraj, S., and Schemitsch, E.H. (2010b). The biomechanics of a validated finite element model of stress shielding in a novel hybrid total knee replacement. *Proc Instn Mech Engrs (Part H): J Engineering in Medicine*, in press.
- Bougherara, H., Bureau, M.N., and Yahia, L. (2010c). Bone remodeling in a new biomimetic polymer-composite hip stem. *J Biomed Mater Res Part A*, Vol.92, No.1, 164-174.
- Bougherara, H., Klika, V., Maršík, F., Mařík, I.A., and Yahia, L. (2010d). New predictive model for monitoring bone remodeling. *J Biomed Mater Res A*, in press.
- Brekelmans, W.A.M., Poort, H.W., and Sloof, T.J. (1972). A new method to analyse the mechanical behaviour of skeletal parts. *Acta Orthop Scand*, Vol.43, No.5, 301-317.

- Buchler, P., and Farron, A. (2004). Benefits of an anatomical reconstruction of the humeral head during shoulder arthroplasty: a finite element analysis. *Clin Biomech (Bristol, Avon)*, Vol.19, No.1, 16-23.
- Cheal, E.J., Hayes, W.C., White, A.A., III, and Perren, S.M. (1983). Stress analysis of a simplified compression plate fixation system for fractured bones. *Comput Struct*, Vol.17, No.5-6, 845-855.
- Cheal, E.J., Hayes, W.C., White, A.A., III, and Perren, S.M. (1984). Three-dimensional finite element analysis of a simplified compression plate fixation system. *J Biomech Eng*, Vol.106, No.4, 295-301.
- Cheal, E.J., Hayes, W.C., White, A.A., III, and Perren, S.M. (1985). Stress analysis of compression plate fixation and its effects on long bone remodelling. *J Biomech*, Vol.18, No.2, 141-150.
- Cheung, G., Zalzal, P., Bhandari, M., Spelt, J.K., and Papini, M. (2004). Finite element analysis of a femoral retrograde intramedullary nail subject to gait loading. *Med Eng Phys*, Vol.26, No.2, 93-108.
- Chong, A.C., Friis, E.A., Ballard, G.P., Czuwala, P.J., and Cooke, F.W. (2007). Fatigue performance of composite analogue femur constructs under high activity loading. *Ann Biomed Eng*, Vol.35, No.7, 1196-1205.
- Claes, L.E., and Ito, K. (2005). Biomechanics of fracture fixation and fracture healing, In: *Basic Orthopaedic Biomechanics and Mechano-Biology: Third Edition*, V.C. Mow and R. Huiskes (Eds.), 563-584, Lippincott Williams and Wilkins, Philadelphia, USA.
- Cristofolini, L., and Viceconti, M. (2000). Mechanical validation of whole bone composite tibia models. *J Biomech*, Vol.33, No.3, 279-288.
- Cristofolini, L., Viceconti, M., Cappello, A., and Toni, A. (1996). Mechanical validation of whole bone composite femur models. *J Biomech*, Vol.29, No.4, 525-535.
- Dammak, M., Shirazi-Adl, A., Schwartz, J., and Gustavson, L. (1997). Friction properties at the bone-metal interface: Comparison of four different porous metal surfaces. *J Biomed Mater Res*, Vol. 35, No.3, 329-336.
- Davis, E.T., Olsen, M., Zdero, R., Waddell, J.P., and Schemitsch, E.H. (2008). Femoral neck fracture following hip resurfacing: the effect of alignment of the femoral component. *J Bone Joint Surg Br*, Vol.90, No.11, 1522-1527.
- Davis, E.T., Olsen, M., Zdero, R., Papini, M., Waddell, J.P., and Schemitsch, E.H. (2009). A biomechanical and finite element analysis of femoral neck notching during hip resurfacing. *J Biomech Eng*, Vol.131, No.4, 041002-1-8.
- Dennis, M.G., Simon, J.A., Kummer, F.J., Koval, K.J., and Di Cesare, P.E. (2000). Fixation of periprosthetic femoral shaft fractures occurring at the tip of the stem: a biomechanical study of 5 techniques. *J Arthroplasty*, Vol.15, No.4, 523-528.
- Dennis, M.G., Simon, J.A., Kummer, F.J., Koval, K.J., and Di Cesare, P.E. (2001). Fixation of periprosthetic femoral shaft fractures: a biomechanical comparison of two techniques. *J Orthop Trauma*, Vol.15, No.3, 177-80.
- Duda, G.N., Heller, M., Albinger, J., Schulz, O., Schneider, E., and Claes, L. (1998). Influence of muscle forces on femoral strain distribution. *J Biomech*, Vol.31, No.9, 841-846.
- Dunlap, J.T., Chong, A.C.M., Lucas, G.L., and Cooke, F.W. (2008). Structural properties of a novel design of composite analogue humeri models. *Ann Biomed Eng*, Vol.36, No.11, 1922-1926.

- Edmondston, S.J., Singer, K.P., Day, R.E., Breidahl, P.D., and Price, R.I. (1994). In-vitro relationships between vertebral body density, size, and compressive strength in the elderly thoracolumbar spine. *Clin Biomech (Bristol, Avon)*, Vol.9, No.3, 180-186.
- Elstrom, J.A., Virkus, W.W., and Pankovich, A.M. (Eds.) (2006). *Handbook of Fractures: Third Edition*, McGraw-Hill, New York, USA.
- Forcione, A. and Shirazi-Adl A. (1990). Biomechanical analysis of a porous-surfaced knee implant: A finite element contact problem with nonlinear friction. In: *Mechanical Engineering Forum*, Canadian Society of Mechanical Engineers, University of Toronto, Toronto, Canada, Vol.2, 19-24.
- Fulkerson, E., Koval, K., Preston, C.F., Iesaka, K., Kummer, F.J., and Egol, K.A. (2006). Fixation of periprosthetic femoral shaft fractures associated with cemented femoral stems: a biomechanical comparison of locked plating and conventional cable plates. *J Orthop Trauma*. Vol.20, No.2, 89-93.
- Grant, J.A., Bishop, N.E., Götzen, N., Sprecher, C., Honl, M., and Morlock, M.M. (2007). Artificial composite bone as a model of human trabecular bone: The implant-bone interface. *J Biomech*, Vol.40, No.5, pp.1158-1164.
- Heiner, A.D., and Brown, T.D. (2001). Structural properties of a new design of composite replicate femurs and tibias. *J Biomech*, Vol.34, No.6, 773-781.
- Heiner, A.D., and Brown, T.D. (2003). Structural properties of an improved redesign of composite replicate femurs and tibias. *Transactions 29th Society for Biomaterials*, Vol.26, Reno, NV, USA, April 30-May 3, p.702.
- Hernandez, C., Beaupré, G.S., and Carter, D.R. (2000). A model of mechanobiologic and metabolic influences on bone adaptation. *J Rehabil Res Dev*, Vol.37, No.2, 235- 244.
- Holman, J. (2000). *Experimental methods for engineers: Seventh edition*, McGraw-Hill, New York, USA.
- Huiskes, R. (1982). On the modeling of long bones in structural analyses. *J Biomech*, Vol.15, No.1, 65-69.
- Huiskes, R., and Chao, E.Y. (1983). A survey of finite element analysis in orthopedic biomechanics: the first decade. *J Biomech*, Vol.16, No.6, 385-409.
- Huiskes, R., and Stolk, J. (2005). Biomechanics and preclinical testing of artificial joints: the hip, In: *Basic Orthopaedic Biomechanics and Mechano-Biology: Third Edition*, V.C. Mow and R. Huiskes (Eds.), 585-656, Lippincott Williams and Wilkins, Philadelphia, USA.
- Huiskes, R., and Van Rietbergen, B. (2005). Biomechanics of Bone, In: *Basic Orthopaedic Biomechanics and Mechano-Biology: Third Edition*, V.C. Mow and R. Huiskes (Eds.), 123-179, Lippincott Williams and Wilkins, Philadelphia, USA.
- Huiskes, R., Weinans, H., Grootenboer, H.J., Dalstra, M., Fudala, B., Slooff, T.J. (1987). Adaptive bone-remodeling theory applied to prosthetic-design analysis. *J Biomech*, Vol.20, No.11-12, 1135-1150.
- Huiskes, R., Ruimerman, R., Harry van Lenthe, G., and Janssen, J.D. (2000). Effects of mechanical forces on maintenance and adaptation of form in trabecular bone. *Nature*, Vol.405, No.6787, 704-706.
- Iwamoto, M., Miki, K., Mohammad, M., Nayef, A., Yang, K.H., Begeman, P.C., and King, A.I. (2000). Development of a finite element model of the human shoulder. *Stapp Car Crash J*, Vol.44, 281-297.

- Kraemer, W.J., Hearn, T.C., Powell, J.N., and Mahomed, N. (1996). Fixation of segmental subtrochanteric fractures. A biomechanical study. *Clin Orthop Relat Res*, Vol.332, pp.71-79.
- Le Huec, J.C., Lafage, V., Bonnet, X., Lavaste, F., Josse, L., Liu, M., and Skalli, W. (2010). Validated finite element analysis of the maverick total disc prosthesis. *J Spinal Disord Tech*, Jan 11. [Epub ahead of print]
- Lescheid, J., Zdero, R., Shah, S., Kuzyk, P.R.T., and Schemitsch, E.H. (2010). The biomechanics of locked plating for repairing proximal humerus fractures with or without medial cortical support. *J Trauma*, Jan 28 [Epub ahead of print].
- Lewis, G., Clark, M.C., Harber, M.S., and Vaughan, M. (1996). The elbow joint and its total arthroplasty. Part II. Finite element study. *Biomed Mater Eng*, Vol.6, No.5, 367-377.
- Lochmüller, E.M., Krefting, N., Burklein, D., and Eckstein, F. (2001). Effect of fixation, soft-tissues, and scan projection on bone mineral measurements with dual x-ray absorptiometry (DXA). *Calcif Tissue Int*, Vol.68, No.3, 140-145.
- Lotke, P.A., and Ecker M.L. (1977). Influence of positioning of prosthesis in total knee replacement. *J Bone Joint Surg Am*, Vol.59, No.1, 77-79.
- Maldonado, Z.M., Seebeck, J., Heller, M.O., Brandt, D., Hepp, P., Lill, H., and Duda, G.N. (2003). Straining of the intact and fractured proximal humerus under physiological-like loading. *J Biomech*, Vol.36, No.12, 1865-1873.
- Martens, M., Van Audekercke, R., De Meester, P., and Mulier, J.C. (1980). The mechanical characteristics of the longbones of the lower extremity in torsional loading. *J Biomech*, Vol.13, No.8, 667-676.
- Martens, M., Van Audekercke, R., De Meester, P., and Mulier, J.C. (1986). Mechanical behaviour of femoral bones in bending loading. *J Biomech*, Vol.19, No.6, 443-454.
- Mather, B.S. (1968). Observations on the effects of static and impact loading on the human femur. *J Biomech*, Vol.1, No.4, 331-335.
- McConnell, A., Zdero, R., Syed, K., Peskun, C., and Schemitsch, E.H. (2008). The biomechanics of ipsilateral intertrochanteric and femoral shaft fractures: a comparison of 5 fracture fixation techniques. *J Orthop Trauma*, Vol.22, No.8, 517-524.
- McElhaney, J., Fogle, J., Byers, E., and Weaver, G. (1964). Effect of embalming on the mechanical properties of beef bone. *J Appl Physiology*, Vol.19, Nov, 1234-1236.
- Oh, J.K., Sahu, D., Ahn, Y.H., Lee, S.J., Tsutsumi, S., Hwang, J.H., Jung, D.Y., Perren, S.M., and Oh, C.W. (2010). Effect of fracture gap on stability of compression plate fixation: a finite element study. *J Orthop Res*, Vol.28, No.4, 462-467.
- Orr, T.E., and Carter, D.R. (1985). Stress analyses of joint arthroplasty in the proximal humerus. *J Orthop Res*, Vol.3, No.3, 360-371.
- OTA (1996). Orthopaedic Trauma Association Committee for Coding and Classification. Fracture and dislocation compendium. *J Orthop Trauma*. Vol.10, Supp.1, 51-55.
- Panagiotopoulou, O. (2009). Finite element analysis (FEA): applying an engineering method to functional morphology in anthropology and human biology. *Ann Hum Biol*, Vol.36, No.5, 609-623.
- Pankovich, A.M., and Davenport, A.M. (2006). Fractures of the femoral shaft, In: *Handbook of Fractures, 3rd Edition*, J.A. Elstrom, W.W. Virkus, and A.M. Pankovich (Eds.), 293-303, McGraw-Hill, Toronto, Canada.

- Papini, M., Zdero, R., Schemitsch, E.H., and Zalzal, P. (2007). The biomechanics of human femurs in axial and torsional loading: comparison of finite element analysis, human cadaveric femurs, and synthetic femurs. *J Biomech Eng*, Vol.129, No.1, 12-19.
- Perren, S.M. (1975). Physical and biological aspects of fracture healing with special reference to internal fixation. *Clin Orthop Rel Res*. Vol.138, 175-194.
- Roesler, H. (1987). The history of some fundamental concepts in bone biomechanics. *J Biomech*, Vol.20, No.11-12, 1025-1034.
- Rouhi, G., Epstein, M., Sudak, L., and Herzog, W. (2007). Modeling bone resorption using mixture theory with chemical reactions, *J Mech Mater Struct*, Vol.2, No.6, 1141-1155.
- Sakai, K., Kiriya, Y., Kimura, H., Nakamichi, N., Nakamura, T., Ikegami, H., Matsumoto, H., Toyama, Y., and Nagura, T. (2010). Computer simulation of humeral shaft fracture in throwing. *J Shoulder Elbow Surg*, Vol.19, No.1, 86-90.
- Schemitsch, E.H., Bhandari, M., Boden, S., Bourne, R., Bozic, K.J., Jacobs, J., and Zdero, R. (2010). The evidence-based approach in bringing new orthopaedic devices to market. *J Bone Joint Surg Am*, Vol.92, No.4, 1030-1037.
- Schipplein, O.D., and Andriacchi, T.P. (1991). Interaction between active and passive knee stabilizers during level walking. *J Orthop Res*, Vol.9, No.1, 113-119.
- Schonning, A., Oommen, B., Ionescu, I., and Conway, T. (2009). Hexahedral mesh development of free-formed geometry: The human femur exemplified. *Computer-Aided Design*, Vol.41, No.8, 566-572.
- Schutte, S., van den Bedem, S.P.W., van Keulen, F., van der Helm, F.C.T, and Simonsz, H.J. (2006). A finite-element analysis model of orbital biomechanics. *Vision Research*, Vol. 46, No.11, 1724-1731.
- Schneider, E., Michel, M.C., Genge, M., and Perren, S.M. (1990). Loads acting on an intramedullary femoral nail, In: Bergmann, G., Graichen, F., Rohlmann, A., eds, *Implantable Telemetry in Orthopaedics*, Berlin: Freie Universitat Berlin, pp.221-227.
- Schneider, E., Michel, M.C., Genge, M., Zuber, K., Ganz, R., and Perren, S.M. (2001). Loads acting in an intramedullary nail during fracture healing in the human femur. *J Biomech*, Vol.34, No.7, 849-857.
- Shimmin, A.J., Walter, W.L., and Esposito, C. (2010). The influence of the size of the component on the outcome of resurfacing arthroplasty of the hip: a review of the literature. *J Bone Joint Surg Br*, Vol.92, No.4, 469-76.
- Simões, J.A., and Marques, A.T. (2005). Design of a composite hip femoral prosthesis. *Materials and Design*, Vol.26, No.5, 391-401.
- Talbot, M., Zdero, R., and Schemitsch, E.H. (2008a). Cyclic loading of periprosthetic fracture fixation constructs. *J Trauma*, Vol.64, No.5, 1308-1312.
- Talbot, M., Zdero, R., Garneau, D., Cole, P.A., and Schemitsch, E.H. (2008b). Fixation of longbone segmental defects: a biomechanical study. *Injury*, Vol.39, No.2, 181-186.
- Taylor, M., Tanner, K.E., Freeman, M.A.R., and Yettram, A.L. (1995). Cancellous bone stresses surrounding the femoral component of a hip prosthesis: an elastic-plastic finite element analysis, *Med Eng Phys*, Vol.17, No.7, 544-550.
- Taylor, M.E., Tanner, K.E., Freeman, M.A.R., and Yettram, A.L. (1996). Stress and strain distribution within the intact femur: compression or bending? *Med Eng Phys*, Vol.18, No.2, 122-131.

- Taylor, M., Verdonschot, N., Huiskes, R., and Zioupos, P. (1999). A combined finite element method and continuum damage mechanics approach to simulate the in vitro fatigue behavior of human cortical bone. *J Mater Sci: Mater Med*, Vol.10, No.12, 841-846.
- van Rietbergen, B., Weinans, H., Huiskes, R., and Odgaard, O. (1995). A new method to determine the trabecular bone elastic properties and loading using micromechanical finite element models. *J Biomech*, Vol.28, No.1, 69-81.
- Wainer, R.A., Wright, P.H., Gilbert, J.A., and Taylor, D.F. (1990). Biomechanics of ender rods, compression screw, and Zickel nail in the fixation of stable subtrochanteric femur osteotomies. *J Orthop Trauma*, Vol.4, No.1, 58-63.
- Wang, C.J., Yettram, A.L., Yao, M.S., and Proctor, P. (1998). Finite element analysis of a gamma nail within a fractured femur. *Med Eng Phys*, Vol.20, No.9, 677-683.
- Weinans, H., Huiskes, R., Van Rietbergen, B., Sumner, D.R., Turner, T.M., and Galante, J.O. (1993). Adaptive bone remodeling around bonded noncemented total hip arthroplasty: A comparison between animal experiments and computer simulation. *J Orthop Res*, Vol. 11, No.4, 500-513.
- Zdero, R., Walker, R., Waddell, J.P., and Schemitsch, E.H. (2008a). Biomechanical evaluation of periprosthetic femoral fracture fixation. *J Bone Joint Surg Am*, Vol.90, No.5, 1068-1077.
- Zdero, R., Olsen, M., Bougherara, H., and Schemitsch, E.H. (2008b). Cancellous bone screw purchase: a comparison of synthetic femurs, human femurs, and finite element analysis. *Proc Instn Mech Engrs (Part H): J Engineering in Medicine*, Vol.222, No.H8, 1175-1183.
- Zdero, R., Elfallah, K., Olsen, M., and Schemitsch, E.H. (2009a). Cortical screw purchase in synthetic and human femurs. *J Biomech Eng*, Vol.131, No.9, 094503-1-7.
- Zdero, R., Olsen, M., Elfatori, S., Skriniskas, T., Nourhosseini, H., Whyne, C., Schemitsch, E.H., and Von Schroeder, H. (2009b). Linear and torsional mechanical characteristics of intact and reconstructed scapholunate ligaments. *J Biomech Eng*, Vol.131, No.4, 041009-1-7.
- Zdero, R., Bougherara, H., Dubov, A., Shah, S., Zalzal, P., Mahfud, A., and Schemitsch, E.H. (2010a). The effect of cortex thickness on intact femur biomechanics: a comparison of finite element analysis with synthetic femurs. *Proc Instn Mech Engrs (Part H): J Engineering in Medicine*, in press.
- Zdero, R., Shah, S., Mosli, M., and Schemitsch, E.H. (2010b). The effect of load application rate on the biomechanics of synthetic femurs. *Proc Instn Mech Engrs (Part H): J Engineering in Medicine*, Vol.224, No.H4, 599-605.

Finite element modeling for a morphometric and mechanical characterization of trabecular bone from high resolution magnetic resonance imaging

Angel Alberich-Bayarri¹, Luis Martí-Bonmatí¹,
M. Ángeles Pérez², Juan José Lerma³ and David Moratal⁴

¹*Department of Radiology, Hospital Quirón Valencia, Valencia, Spain.*

²*Group of Structural Mechanics and Materials Modeling, Instituto de Investigación en Ingeniería de Aragón, Universidad de Zaragoza, Zaragoza, Spain.*

³*Department of Rheumatology, Hospital Quirón Valencia, Valencia, Spain.*

⁴*Center for Biomaterials and Tissue Engineering, Universitat Politècnica de València, Valencia, Spain.*

1. Introduction

Trabecular bone disorders such as osteoporosis are currently becoming a worldwide concern, affecting nearly 200 million of people (Reginster & Burlet, 2006). This syndrome is characterized by a diminution of the bone density and also deterioration in the trabecular bone architecture, both factors inducing an increased bone fragility and higher risk to suffering skeletal fractures. The current diagnosis of osteoporosis disease is based on bone mineral density (BMD) measurements, which are obtained using Dual emission X-ray Absorptiometry (DXA) techniques and do not provide any information about the trabecular bone properties at structural level. Although it is accepted that low BMD increases fracture risk, not all patients who are osteoporotic will sustain a fracture. Low BMD as a risk factor for fracture can be compared to high cholesterol as a risk factor for heart disease, being BMD a rather poor predictor of fracture risk (Wehrli et al., 2007). In fact, a large analysis showed that, on the average, BMD explains about 60% of bone strength (Wehrli et al., 2002).

The majority of bone fractures associated with osteoporosis take place in vertebrae, hip or wrist. In the European Union, in the year 2000, the number of osteoporotic fractures was estimated at 3.79 million, of which 24% were hip fractures (Kanis & Johnell, 2004). Of all fractures due to osteoporosis, hip fractures are the ones that are most disabling. In 1990, about 1.7 million new hip fractures occurred worldwide, and this figure is expected to rise to 2.6 million by 2025 (Gullberg et al., 1997). Furthermore, women who have sustained a hip fracture have a 10% to 20% higher mortality than would be expected for their age (Cummings & Melton, 2002).

Fractures are the consequence of mechanical overloading. The evaluation of elasticity and fluency limits of different materials and structures is commonly assessed by mathematics and engineering techniques. It seems reasonable to consider that bone fracture risk should also be analyzed by exact and precise engineering methods. Nevertheless, although bone mechanical properties can be directly extracted from *ex vivo* experiments, the mechanical analysis of bone *in vivo* has always been limited to indirect measurements since there is no physical interaction with the sample. These indirect measurements have evolved through time, being initially age and sex two good indicators of bone health. Nowadays, the most extended measure of bone quality is the BMD measured using DXA techniques. In fact, the DXA technique has been shown to be accurate, precise and reproducible in the BMD quantification. However, other imaging-based techniques that work at microstructural level are of high interest in the evaluation of bone microarchitecture quality. Therefore, the investigation in new methods to characterize trabecular bone microarchitecture alterations produced by the disease is crucial. Such advances may not only be useful for a better diagnosis of the disease, but also to evaluate the efficacy of the existing therapies.

The advances of digital radiology and modern medical imaging techniques like computed tomography (CT), high-resolution peripheral quantitative computed tomography (HR-pQCT) and magnetic resonance imaging (MRI) have allowed the research of *in vivo* characterization of bone microarchitecture (Ito et al., 2005; Boutroy et al., 2005). Although a high sensitivity to bone tissue structure is appreciated in X-Ray based techniques, like CT or HR-pQCT which can achieve very high spatial resolutions, the high inherent contrast between bone and bone marrow and its non-ionizing radiations places MRI as an appropriate technique for the *in vivo* characterization of cancellous bone.

Recent high field MRI scanners permit the acquisition of images with a high spatial resolution, in combination with satisfactory signal-to-noise ratio (SNR) and contrast-to-noise ratio (CNR). In fact, the higher spatial resolutions achievable with high field 3 Tesla MR scanners allow the development of reliable three-dimensional (3D) reconstructions of tissues with the possibility of applying engineering-based methodology to simulate structures and behaviours under certain conditions. The finite element (FE) method is the most extended mathematical tool for complex modelling and simulations and it has been widely applied to civil, spatial and many different engineering disciplines. This method can also be used in biomedical engineering for the simulation of organs and tissues under different scenarios (Prendergast, 1997). In the case of bone tissue, the combination of high quality image acquisitions with proper image post-processing and computational based mechanical simulations using FE with reduced element size, also called micro-FE (μ FE), will improve osteoporosis characterization, reinforce the diagnosis process, and perform an accurate evaluation of the response to different therapies (Ito et al., 2005; Boutroy et al., 2005; Alberich-Bayarri et al., 2008).

In the present chapter, the process for trabecular bone mechanical simulation and analysis from high spatial resolution MRI acquisitions is detailed. Results of the FE application to healthy and osteoporotic populations are also shown.

2. Image acquisition and processing

Image acquisition

Due to the high spatial resolutions required for the assessment of trabecular bone tissue, SNR tends to be significantly reduced in MRI acquisitions. There exists a number of imaging problems inherent to the use of high spatial resolution MRI where technical parameters have to be considered. The best MRI configuration for proper acquisitions of the trabecular bone region is based on 3D gradient echo (GRE) pulse sequences with reduced echo-time (TE), repetition time (TR) and flip angle (α) in order to obtain a high contrast between bone and surrounding bone marrow. A high SNR is achieved with the shortest TE and the use of phased array surface coils that can cover the entire region under analysis. In terms of acquisition quality, an appropriate area for the trabecular bone examination is the wrist region, which can be easily surrounded by the receiving coil and the sample - receptor distance is minimized. Metaphyses of the ulna and radius in the wrist region are rich in trabecular bone tissue. Acquired images obtained at metaphyseal level of the radius have a high detail of the bone trabeculae, as it can be appreciated in figure 1.

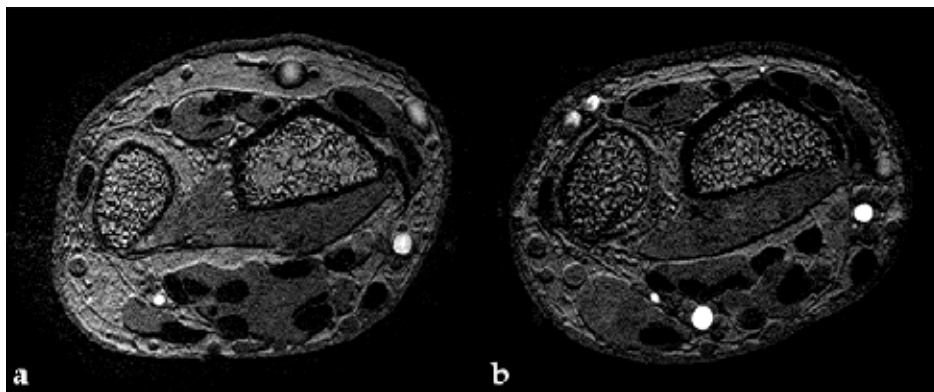


Fig. 1. Images acquired with a 3 Tesla MRI system (Achieva, Philips Healthcare, Best, The Netherlands). In a), wrist slice corresponding to the acquisition performed in a left-handed 62 years old female patient with osteoporosis. In b), image corresponding to a series acquired in a left-handed 64 years old female healthy volunteer.

In the examples shown in figure 1, acquired images have a very high spatial resolution, with an isotropic voxel size of $180 \times 180 \times 180 \mu\text{m}^3$, which is even higher than the achievable with modern multi-detector CT scanners. However, the use of surface coils in the acquisitions produce slight modulations of the signal intensities across the images, also known as coil shading phenomena. Furthermore, although a high spatial resolution of $180 \mu\text{m}$ was used, achieved voxel size was slightly larger than typical thickness of the trabeculae, which is about $100\text{-}150 \mu\text{m}$. As it can be concluded, although the images have a very good quality, some pre-processing must be applied before 3D reconstruction, meshing and simulation of the trabecular bone structures.

Segmentation

Segmentation of the trabecular bone from MR images is performed by placing a rectangular region of interest (ROI) in the first slice corresponding to the most proximal position and

thereafter propagated to the rest of slices. Segmented areas must be verified to exclusively contain marrow and trabecular bone (Figure 2-a).

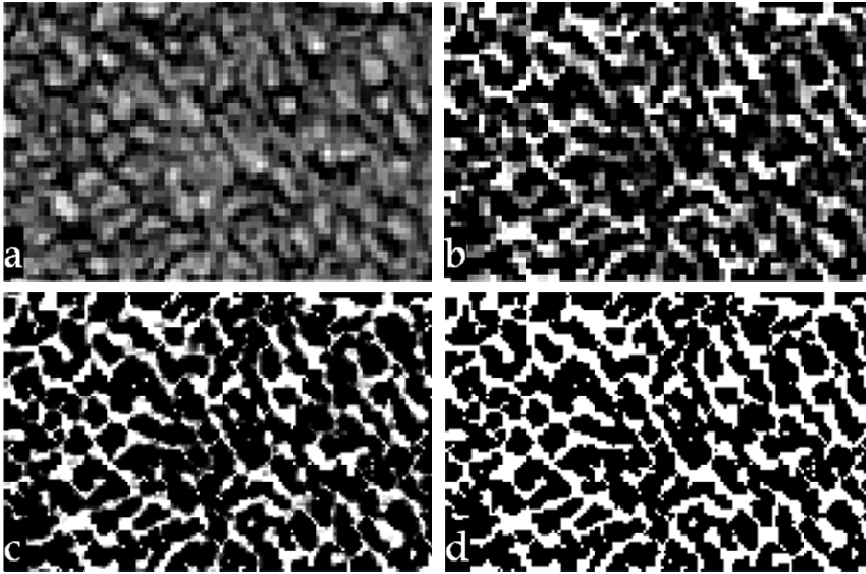


Fig. 2. Representative images resulting from the application of the different pre-processing algorithms. Example of the isolated region exclusively containing trabecular bone and marrow (a). Image resulting from the application of the coil shading correction through the local intensities determination (b). Interpolated image after the execution of the sub-voxel processing algorithm for an apparently increased spatial resolution (c). Binarized image obtained at the end of the pre-processing chain where only complete bone or marrow voxels exist (d).

Coil heterogeneities correction

Slight modulations of the signal intensities across the acquisition volume, also known as coil shading phenomena, are corrected with nearest-neighbour statistics by the application of an implemented 3D local thresholding algorithm (LTA), as a generalization from its 2D version (Vasilic & Wehrli, 2005). Marrow intensity values in the neighbourhood of each voxel are determined and bone voxels intensities are scaled using calculated local intensities. Concretely, the method is based on the calculation of the average Laplacian values $\langle \tilde{L} \rangle_{\vec{r}}(I)$

in a sphere region $S(\vec{r})$ with a radius $R=15$ pixels, being I the voxel intensity and \vec{r} the center of the sphere, which is displaced through all voxels of the volume. When the calculated Laplacian equals zero $\langle \tilde{L} \rangle_{\vec{r}}(I_t(\vec{r}))=0$, the corresponding marrow intensity $I_t(\vec{r})$

can be determined locally. After the marrow intensity is obtained, it is used as a threshold and voxels can be directly classified into pure marrow voxels or scaled by its local threshold value and labeled as partially occupied by bone voxels (Figure 2-b).

Sub-voxel processing

The extreme conditions in terms of low SNR and partial volume effects due to larger voxel size than typical thickness of the trabeculae, which is about 100-150 μm , forced the

implementation of a method to increase the reconstructed spatial resolution. A subvoxel-processing algorithm is applied to minimize partial volume effects and therefore improving the cancellous bone structural quantification from MR images (Hwang & Wehrli, 2002). The method consists on a two-pass algorithm where each voxel is initially divided into eight subvoxels which are assigned a level of intensity conditioned by the corresponding level of their voxel and near subvoxels, and also under the assumption that the amount of bone intensities must be conserved. In the first pass of the algorithm, each subvoxel is assigned an intensity value depending on the intensities of the adjacent voxels and the local sum of intensities. The second pass of the algorithm consists on the refinement of the previously calculated subvoxel intensities considering the intensities of the neighbouring subvoxels and the total amount of intensity conservation. Finally, an increased apparent isotropic spatial resolution of 90 μm is achieved and partial volume effects are minimized (Figure 2-c).

Binarization

Resulting images are binarized into exclusively bone or marrow voxels (Figure 2-d). Histogram shape-based thresholding is applied through the Otsu's method implemented in 3D. The method consists in the minimization of the intra-class variance of volume intensities (Otsu N., 1979), which has been shown to be equivalent to maximizing the between-class variance. Thus, the optimum separation threshold is calculated for the entire volume using:

$$t^* = \arg \max_t \frac{\sigma_B^2}{\sigma_T^2} \quad (1)$$

where σ_T^2 is the total variance and σ_B^2 is the between-class variance. The optimum binarization threshold is t^* . Finally, voxels can be classified as bone or marrow depending on their intensity value.

3. Volumetric reconstruction and meshing

Once the region under analysis has been binarized, a logical 3D matrix is obtained, with bone voxels represented by 1's and marrow voxels by 0's. For visualization, the volume is smoothed by a three-dimensional routine and the marching cubes algorithm (Lorensen & Cline, 1987) is then applied in order to obtain a volumetric reconstruction of the structure, as can be seen in figure 3.

Depending on the model to be simulated, different element types can be used to construct the mesh. Two-dimensional (2D) surface based meshes are usually compound of triangles or quadrilaterals. Considering a trabecular bone structure, a 3D volume mesh needs to be generated for a volumetric simulation and evaluation of trabeculae mechanical resistance. Volume meshes are usually compound of solid elements such as tetrahedrons and hexahedrons. Isotropic hexahedrons with 8 nodes are usually referred as brick elements.

To develop an efficient meshing of the trabecular bone structure, brick elements are the best option in terms of computational burden. The 3D reconstruction is conserved and a more efficient stress analysis can be implemented (Alberich-Bayarri et al., 2007).

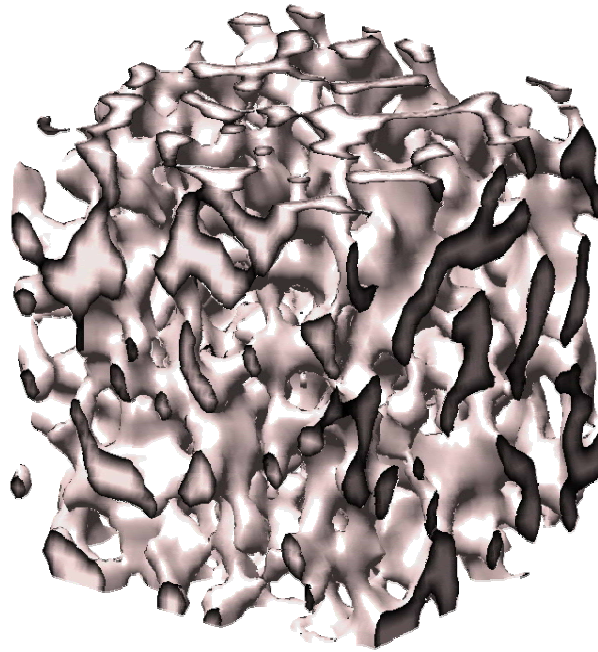


Fig. 3. Smoothed 3D reconstruction of the trabecular bone obtained from the application of volume rendering algorithms to the binary volume of data.

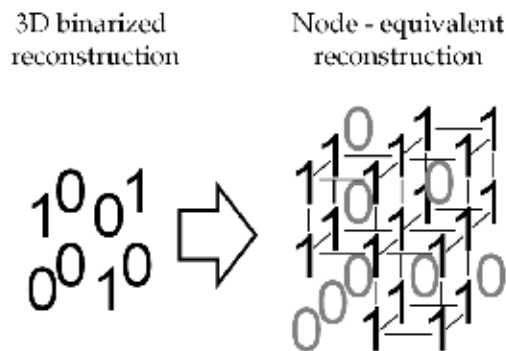


Fig. 4. Example showing the algorithm basis for conversion from the 3D binarized matrix into a 3D matrix containing the node information of the structure. Node coordinates and element connectivity extraction from this process is straightforward.

An algorithm is needed to convert 3D reconstruction geometry information into FE structural data where node coordinates and element interconnections are detailed. Although a very intuitive algorithm would consist in the voxel-by-voxel sequential analysis of structural coordinates and corresponding connectivity, an optimized algorithm implemented by the authors is used for the mesh generation process.

The idea of this algorithm is a direct detection of nodes of the structure working completely with matrices and without sequentiality. Considering the 3D binarized reconstruction matrix of size $m \times n \times p$, a new matrix nodal equivalent matrix of size $(m+1) \times (n+1) \times (p+1)$ is created and filled with zeros. Then, for each bone voxel found in the 3D binarized reconstruction, eight ones representing the nodes of the element are situated on the corresponding position of the node equivalent reconstruction (see Fig. 4). The extraction of the node coordinates from the node equivalent matrix is direct. The relationship between the original bone voxel and the corresponding calculated nodes is also conserved. The implemented meshing algorithm can build a mesh with 1048706 nodes and 285148 elements in 9.78 seconds using an Intel® Core™ 2 Quad CPU at 2.83 GHz and 8 Gb of RAM memory. At the end of the process, the lists containing the nodes with the corresponding coordinates and the elements with the corresponding node connectivity are completed and the FE mesh is fully defined.

Once the mesh is strictly defined, it must be assembled in a file to ease the importation from specific FE commercial software where the simulation is performed. Different FE tools like ANSYS (Ansys Inc., Southpointe, PA, USA) or ABAQUS (Simulia, Providence, RI, USA) are used for the FE model definition and posterior simulation.

4. FE model definition

The generated mesh must be prepared for the simulation after it has been imported in the FE software, that is, a model must be fully generated considering the bulk material properties, boundary conditions and loads application.

Initially, the bulk material properties for each element must be defined. In this case, elements composition is supposed to be formed by compact bone, with linear, elastic and isotropic behaviour. Numerically, the elastic properties of the material forming the elements consist of a Young's modulus given by $E_b=10\text{GPa}$ and a Poisson's ratio of $\nu=0.3$ (Fung, 1993; Newitt et al., 2002).

Experimentally, resistance and elasticity properties of materials and other structures are evaluated by a compression essay using specific equipment in laboratory. Stress-strain relationship is analyzed and Young's modulus of the whole structure can be easily calculated from the linear slope of the stress-strain curve. Analogously, in the FE simulation, null displacement is imposed on nodes from one side while a total strain (ϵ) of 10% of the edge length is specified on nodes from the opposite side for the compression simulation. An iconography of the scenario to be modelled can be appreciated in figure 5.

The modelling of a trabecular bone uniaxial compression essay requires deciding the loading direction. In our case, in the radius bone, the principal orientation of the trabecular bone corresponds to the longitudinal dimension (Gullberg et al., 1997). Since an axial acquisition is performed for the images generation, the longitudinal direction corresponds to the MR scanner slice z direction. In this sense, a crucial advantage of 3 Tesla systems comparing with 1.5 Tesla is the capability of performing high resolution acquisitions not only in plane but also in slice direction (isotropic voxel size). In figure 6, a FE mesh of a trabecular bone structure loaded to the commercial package ANSYS 10.0 is shown. The boundary conditions and the strain application can be also appreciated in the figure.

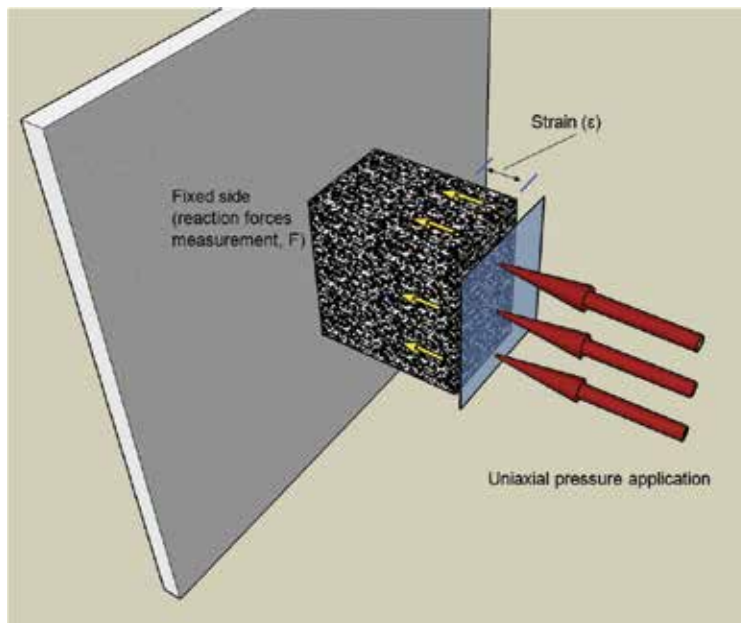


Fig. 5. Representation of the simulation applied using the FE method. A compression test is applied to the trabecular bone sample with a side fixed to a null displacement and the opposite side suffering an imposed deformation.

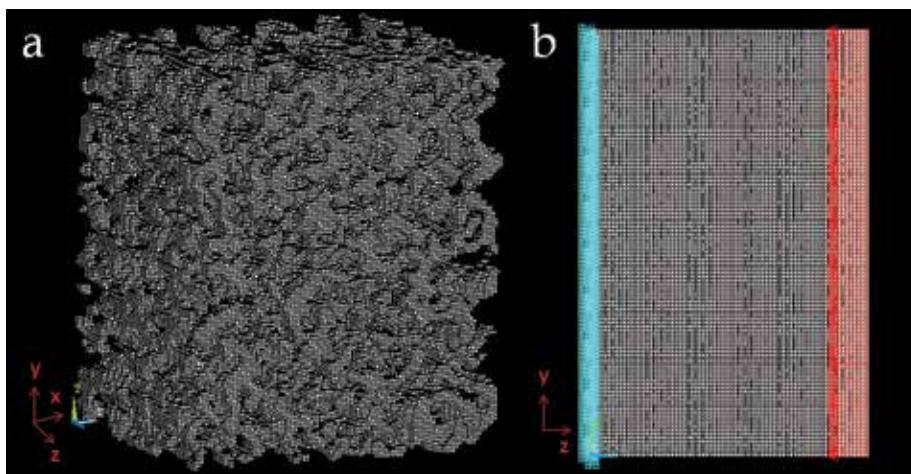


Fig. 6. In a), a FE mesh of trabecular bone imported in the commercial software ANSYS 10.0 is shown. The coordinates system is also shown, with the z direction corresponding to the longitudinal direction of the bone. In b), lateral view of the cancellous bone structure with boundary conditions definition, that is, null displacement on the left side (blue) and 10% strain in the right side (red).

Once the model is completely defined, including forces application, material properties and boundary conditions, simulation process is ready to begin.

5. Simulation and results calculation

The calculation of the mechanical results from the simulation is a very demanding task in terms of computational cost. The number of equations of our system is considerably high, since there are 3 degrees of freedom per node, and each element has 8 nodes. As an example, a mesh with 800000 nodes defines a system of 2400000 equations.

In order to solve the large systems of equations, different strategies can be used, depending on the kind of solver (sparse, conjugate gradients, minimal residual). In the present case, for the characterization of the trabecular bone response, systems are solved using a standard sparse solver.

Although the different techniques for handling the matrices and approaching to the final solution, the mathematical problem to be solved is summarized as follows (Zienkiewicz et al., 2006):

The stiffness matrix of each element K^e can be calculated by equation (2),

$$K^e = \int_{V^e} B^T D B d(vol) \quad (2)$$

where D is the elasticity matrix, which exclusively depends on Young's modulus and Poisson's ratio (see equation (3)) and B is a matrix containing spatial information.

$$D = \frac{E}{1-\nu^2} \begin{bmatrix} 1 & \nu & 0 \\ \nu & 1 & 0 \\ 0 & 0 & (1-\nu)/2 \end{bmatrix} \quad (3)$$

The global stiffness matrix of the structure under analysis can be assembled by equation (4), taking into consideration the stiffness matrix of each element.

$$K_{ij} = \sum K_{ij}^e \quad (4)$$

Once the stiffness matrix is assembled, the objective of the FE problem is to solve the structural equation (5), which relates the global stiffness matrix with the nodal displacements (u) and forces (f):

$$K \cdot u = f \quad (5)$$

After nodal displacements and forces are calculated, the calculation of stresses and strains is straightforward.

A very interesting parameter that can be extracted from the FE results is the apparent Young's modulus of the whole structure by the application of the homogenization theory (Hollister et al., 1991; Hollister & Kikuchi, 1992) and equation (6).

$$E_{app} = \frac{1}{\varepsilon A} \sum_n F \quad (6)$$

Where F corresponds to the nodal reaction forces measured in the fixed side, A is the area of the fixed side, and ε is the imposed strain.

Apparent Young's modulus (E_{app}) is used as an estimation of cancellous bone resistance to compression.

6. Validation with sheep models

The mechanical trabecular bone biomarker parameters extracted from high resolution 3T MR images can be validated by a comparison with 64-MDCT and micro-CT derived parameters in the cancellous bone of sheep tibiae.

Five fresh legs were extracted from adult sheep cadavers in order to perform the trabecular bone analysis of the tibia. The samples were prepared after skin removal and cleaning for the MR and 64-MDCT acquisitions. The μ CT examinations required the preparation of small trabecular bone samples that were extracted from the metaphysis of the sheep tibiae. Results of the Von Mises stress maps can be appreciated in an example in figure 7.

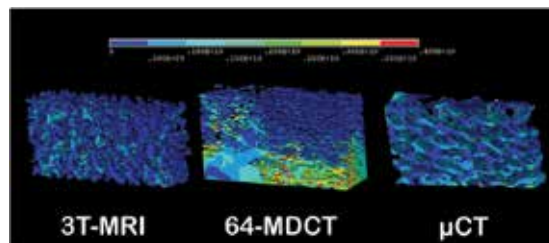


Fig. 7. Parametric reconstructions of the Von Mises nodal stress results for the compression simulation using three different modalities (3T-MRI, MDCT and μ CT).

The values obtained for the analyzed parameters in the 3 acquisition modalities can be appreciated in Table 1. Setting micro-CT results as the reference, MR calculated elastic modulus were slightly smaller while CT derived elastic modulus were clearly much higher. This bone overestimation in MDCT acquisitions is probably due to the bone hyper estimation effect of the point spread function (PSF) at high spatial resolutions.

	Ex (MPa)	Ey (MPa)	Ez (MPa)
3T-MRI	79.910	320.160	827.390
MicroCT	105.920	485.911	995.390
MDCT-64	1122.509	1430.551	2577.300

Table 1. Results for the apparent Young's modulus obtained in a trabecular bone sample from the sheep legs.

The results obtained for the elastic modulus after compression FE test of the MR-derived mesh are more proximal to the calculated for micro-CT than the obtained from 64-MDCT. Compared to micro-CT, a bone overestimation is observed in the 64-MDCT calculated parameters, as previously mentioned. Similar stress distributions can be appreciated in

Figure 1 between 3T-MR and micro-CT while 64-MDCT derived stress map shows a high and non-uniform stress arrangement.

7. Application to healthy subjects and patients with osteoporosis

The Young's modulus parameter was initially evaluated in a healthy population in order to define a rank of normality for the parameter and evaluate differences between sex or age.

The Young's modulus results on a healthy population of 40 subjects showed significant differences between males and females ($p=0.012$) with lower values in females (143.02 ± 24.83 MPa vs. 241.58 ± 28.07 MPa, mean \pm standard error of the mean (SEM), female and male respectively). Age showed no relationship with the Young's modulus parameter ($r^2=0.033$) (Alberich-Bayarri et al., 2008).

An interesting point is that the volume of analyzed bone had an influence on the elasticity results. There was a statistically significant difference ($p<0.001$, ANOVA test) between Eapp results obtained for the large volume of interest (VOI) and those obtained for a small VOI (189.84 ± 20.02 vs. 81.56 ± 16.86 , mean \pm standard deviation). Furthermore, as it can be appreciated in figure 8, differences between Eapp results for the two different volumes analyzed become greater when the values of the elastic parameter grow. The reason for small VOI analysis was the minimization of the computational burden in compression simulation. However, it was shown that the goodness of the results diminished as the restricted volume dimensions decreased (Alberich-Bayarri et al., 2008). Authors detected that the mechanical analysis results applied to the large sample volumes of trabecular bone better characterized its properties, with significant differences between male and female subjects for Young's modulus. As expected, Eapp results were greater in men (Keaveny & Yeh, 2002).

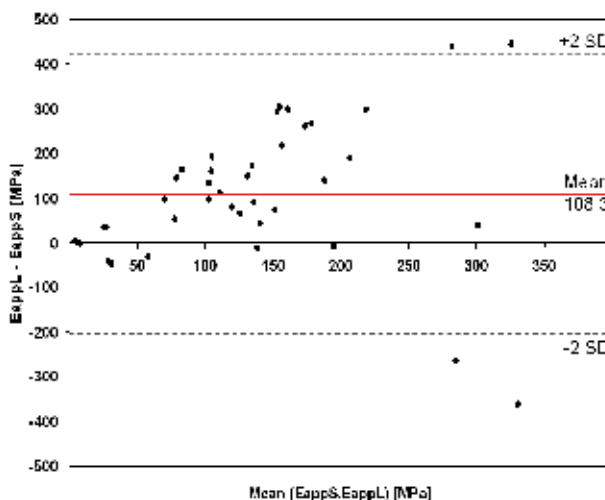


Fig. 8. Bland-Altman plot comparing the results of the Eapp for two different VOI analyzed, the large volume (L) and the small volume (S). (adapted from Alberich-Bayarri et al.).

Authors' results for the mean elastic modulus were lower than the value obtained for the uniaxial compression (Newitt D.C. et al., 2002) (189.84 ± 126.62 MPa vs. 2050 ± 590 MPa, respectively). However, these differences are probably due to different acquisition

configurations. In this sense, author's results are close to those obtained by (Dagan et al., 2002) with an Eapp of 150 MPa. In conclusion, acquisition technique and processing tools clearly influence the Eapp values.

Also, the same methodology was applied to a population of 20 female osteoporosis patients in order to evaluate Young's modulus sensitivity to the disease. Mean Young's modulus computed for the osteoporotic population and a paired reference subjects is shown in table 2.

Parameter	Healthy (n=21)	Patients (n=20)
Eapp (MPa)	142 ± 25	43.1 ± 24.8

Table 2. Young's modulus results in a population of healthy volunteers and a group of patients with Osteoporosis.

Resultant Von Mises stress parametric maps were also qualitatively different between healthy and osteoporotic subjects. While healthy volunteers present a uniform structural response to compression, with a distributed stresses, in the case of patients with osteoporosis the parametric reconstructions showed a high heterogeneity with a specific zone of high stress and resulting in a less resistant structure.

8. Conclusions and future challenges

Although further evaluation is needed about the usefulness of these methods in osteoporosis, it seems clear that computational generated 3D models of the cancellous bone from high resolution 3 Tesla MRI can be used to characterize bone *in vivo*, analyzing different mechanical conditions of the cancellous microstructure.

The FE analysis could be performed in a more complex philosophy, if bone anisotropy is considered for bulk material properties definition (Hellmich et al., 2008), and also if not only the part with linear behaviour of the stress-strain curve is considered, but also the curve region showing plasticity or non-recoverable deformation. These studies should add more information to the mechanisms involved in bone fracture, like buckling phenomena, just before the breaking point arrives.

Patients are classified nowadays using the World Health Organization (WHO) criteria, depending on the amount of bone loss, in osteopenic or osteoporotic patient. These are the clinical references used for new biomarkers evaluation. However, multivariate studies may provide different groups or classification patterns for the patient populations.

Results of the mechanical simulations among a large population should provide knowledge for the establishment of new biomarkers of disease. New clinical trials and studies should help to analyze the sensitivity of these parameters to the treatment.

Acknowledgements

The support of the Spanish Ministry of Science and Innovation through project TEC2009-14128 and of the Generalitat Valenciana through projects GV/2009/126 (grups d'investigació emergents) and ACOMP/2010/022 (ajudes complementàries) is acknowledged. The funding from Instituto de la Mediana y Pequeña Industria Valenciana (IMPIVA) of the Generalitat Valenciana (IMIDTP/2009/334) is also acknowledged.

9. References

- Alberich-Bayarri, A.; Moratal, D.; Martí-Bonmatí, L.; Salmerón-Sánchez, M.; Vallés-Lluch, A.; Nieto-Charques, L. & Rieta JJ. (2007). Volume mesh generation and finite element analysis of trabecular bone magnetic resonance images. *Proceedings of: Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 1603-1606, 1557-170X, Lyon, France, August 2007, *IEEE Engineering in Medicine and Biology Society*.
- Alberich-Bayarri, A.; Marti-Bonmati, L.; Sanz-Requena, R.; Belloch, E. & Moratal, D. (2008). In vivo trabecular bone morphologic and mechanical relationship using high-resolution 3-T MRI. *American Journal of Roentgenology*, 191, 3, (Sep 2008) 721-726, 0361-803X.
- Boutroy, S.; Bouxsein, M.L.; Munoz, F. & Delmas, P.D. (2005). In vivo assessment of trabecular bone microarchitecture by high-resolution peripheral quantitative computed tomography. *The Journal of clinical endocrinology and metabolism*, 90, 12 (Dec 2005), 6508-6515, 0021-972X.
- Cummings, S.R. & Melton III, J.R. (2002). Epidemiology and outcomes of osteoporotic fractures. *Lancet*, 359, 9319 (May 2002), 1761-1767, 0140-6736.
- Fung, Y.C. (1993). Biomechanics. *Mechanical properties of living tissues*, pp. 500-518, 2nd ed, Springer, New York, NY.
- Dagan, D.; Be'ery, M. & Gefen, A. (2004). Single-trabecula building block for large-scale finite element models of cancellous bone. *Medical & biological engineering & computing*, 42, 4 (Jul 2004), 549-556, 0140-0118.
- Geraets, W.G.; Van der Stelt, P.F.; Lips P.; Elders P.J.; Van Ginkel F.C. & Burger E.H. (1997). Orientation of the trabecular pattern of the distal radius around the menopause. *Journal of Biomechanics*. 30, 4 (Apr 1997), 363-370, 0021-9290.
- Gullberg, B.; Johnell, O. & Kanis, J.A. (1997). World-wide projections for hip fracture. *Osteoporosis International*, 7, 5, 407-413, 0937-941X.
- Hellmich, C.; Kober, C. & Erdmann, B. (2008). Micromechanics-based conversion of CT data into anisotropic elasticity tensors, applied to FE simulations of a mandible. *Annals of biomedical engineering*, 36, 1 (Jan 2008), 108-122, 0090-6964.
- Hollister, S.J.; Fyhrie, D.P.; Jepsen, K.J. & Goldstein, S.A. (1991). Application of homogenization theory to the study of trabecular bone mechanics. *Journal of biomechanics*, 24, 9, 825-839, 0021-9290.
- Hollister, S.J. & Kikuchi, N. (1992). A comparison of homogenization theory and standard mechanics analyses for periodic porous composites. *Computational Mechanics*, 10, 2 (Mar 1992), 73-95, 0178-7675.
- Hwang, S.N. & Wehrli, F.W. (2002). Subvoxel processing: a method for reducing partial volume blurring with application to in vivo MR images of trabecular bone. *Magnetic Resonance in Medicine*, 47, 5 (May 2002), 948-957, 0740-3194.
- Ito, M.; Ikeda, K.; Nishiguchi, M.; Shindo, H.; Uetani, M.; Hosoi, T. & Orimo, H. (2005). Multi-detector row CT imaging of vertebral microstructure for evaluation of fracture risk. *Journal of bone and mineral research : the official journal of the American Society for Bone and Mineral Research*, 20, 10 (Oct 2005), 1828-1836, 0884-0431.
- Kanis, J.A. & Johnell, O. (2005). Requirements for DXA for the management of osteoporosis in Europe. *Osteoporosis International*, 16, 3 (Mar 2005), 229-238, 0937-941X.

- Keaveny, T.M. & Yeh, O.C. (2002). Architecture and trabecular bone - toward an improved understanding of the biomechanical effects of age, sex and osteoporosis. *Journal of musculoskeletal & neuronal interactions*, 2, 3 (Mar 2002), 205-208, 1108-7161.
- Lorensen, W.E. & Cline, H.E. (1987). Marching Cubes: A high resolution 3D surface construction algorithm. *Computers & Graphics*, 21, 4 (July 1987), 163-169, 0097-8493.
- National Institutes of Health. (2000). Osteoporosis Prevention, Diagnosis, and Therapy. *Consensus Development Conference Statement*. 17, 1 (March 27-29, 2000), 1-45.
- Newitt, D.C.; Majumdar, S.; van Rietbergen, B.; von Ingersleben, G.; Harris, S.T.; Genant, H.K.; Chesnut, C.; Garnero, P. & MacDonald, B. (2002). In vivo assessment of architecture and micro-finite element analysis derived indices of mechanical properties of trabecular bone in the radius. *Osteoporosis International*, 13, 1 (Jan 2002), 6-17, 0937-941X.
- Otsu, N. (1979). A threshold selection method from gray-level histogram. *IEEE Transactions on Systems, Man, and Cybernetics*, 9, 1, 82-86, 1083-4427.
- Prendergast, P.J. (1997). Finite element models in tissue mechanics and orthopaedic implant design. *Clinical biomechanics (Bristol, Avon)*, 12, 6 (Sep 1997), 343-366, 0268-0033.
- Reginster, J.Y. & Burlet, N. (2006). Osteoporosis: a still increasing prevalence. *Bone*, 38, 2 Suppl 1 (Feb 2006), S4-9, 8756-3282.
- Vasilic, B. & Wehrli, F.W. (2005). A novel local thresholding algorithm for trabecular bone volume fraction mapping in the limited spatial resolution regime of in-vivo MRI. *IEEE transactions on medical imaging*, 24, 12 (Dec 2005), 1574-1585, 0278-0062.
- Wehrli, F.W.; Saha, P.; Gomberg, B.; Song, H.K.; Snyder, P.J.; Benito, M.; Wright, A. & Weening, R. (2002). Role of magnetic resonance for assessing structure and function of trabecular bone. *Topics in magnetic resonance imaging : TMRI*, 13, 5 (Oct 2002), 335-355, 0899-3459.
- Wehrli, F.W. (2007). Structural and functional assessment of trabecular and cortical bone by micro magnetic resonance imaging. *Journal of magnetic resonance imaging : JMRI*, 25, 2 (Feb 2007), 390-409, 1053-1807.
- Zienkiewicz, O.C.; Taylor, R.L. & Zhu, J.Z. (2006). *The finite element method, its basis & fundamentals*. 6th edition. Elsevier, Oxford, UK. 2006. 0750663200.

Finite Element Modeling of the Human Lumbar Spine

Marta Kurutz

*Budapest University of Technology and Economics
Hungary*

1. Introduction

Finite element (FE) numerical simulation is an effective tool for analyzing phenomena that cannot be clarified by experimental methods, like most of the biomechanical processes, for example the age-related spinal degeneration processes. Moreover, numerical simulation techniques have the potential to reduce costs and to save time during the development of new effective spinal treatment methods or implants. Consequently, there is a need to obtain more and more realistic and correct numerical models for the very complicated structure, the human spine.

In this chapter the FE modeling aspects of the most frequented spinal part, the lumbar spine is presented. After giving a short overview of the anatomy of the lumbar spine, biomechanical effects, loads, internal forces and movements are detailed. Then the three steps of FE modeling procedure, the geometric, material and element/mesh type modeling is discussed, followed by the validation of the complete FE model. Finally, an example for FE numerical simulation closes the chapter.

2. Structural anatomy of the lumbar spine

The lumbar spine is the section of spinal column between the thorax and the sacrum. It consists of five *vertebrae* named L1 to L5 with their posterior elements and articular facet joints, of *intervertebral discs*, *ligaments* and the surrounding *muscles*. In ideal case, the axis of the lumbar spine is a plane curve closed to the section line of the two vertical anatomic planes. The sagittal plane is the vertical plane of symmetry of the body; the frontal or coronal plane is the vertical lateral plane perpendicular to the sagittal one, parallel to the shoulders. The horizontal anatomic plane meets them in the intersection point located in the lumbar spine. The clinical anatomy of the lumbar spine can be studied in the books of Bogduk and Twomey (1987), White and Panjabi (1990), Dvir (2000), Benzel (2001), Adams et al. (2002) and Bogduk (2005).

2.1 Vertebral body, posterior elements, articular facet joints

The lumbar vertebrae are roughly cylindrical with a lateral diameter (width) of 40-50 mm and sagittal diameter (depth) of 30-35 mm. The lumbar vertebrae are thicker anteriorly than

posteriorly resulting in anteriorly convex curvature of the spine known as the lumbar lordosis. Thus, the dorsal body heights are 25-27 mm, the ventral body heights are 26-29 mm.

The *vertebral body* consists of an outer shell of high strength *cortical bone* reinforced internally by the *cancellous bone* as a network of vertical and horizontal narrow bone struts called *trabeculae*. The superior and inferior surface of the vertebral body is covered by the bony endplates of thin cortical bone perforated by many small holes which allow the passage of metabolites from bone to the central regions of the avascular discs.

Towards the upper end of the posterior surface of the vertebral body is a pair of stout pillars of bone called the *pedicles* supporting the *posterior elements*. From each pedicle a skew plate of bone called the *lamina* goes towards the midline where they fuse. The pedicles and the laminae together form the *neural arch*, which together with the posterior surface of the vertebral body encloses a channel, the *vertebral foramen*. The peak-like junction of the two laminae in the midline forms the *spinous process*, while at the lateral junction of each lamina and pedicle starts a long flattened bar of bone called the *transverse process*. At the root of each transverse process, two extensions of bone rise from the lamina: upwards the superior, downwards the inferior *articular process*.

The superior articular processes of a vertebra meet the inferior articular process of the adjacent vertebra above, forming the two symmetrically located synovial joints called *articular facet joints* with an average area of 1.6 cm² each. The articular surfaces are approximately vertical in the upper lumbar spine, but are more oblique at the lower part of it.

2.2 Intervertebral discs

The *intervertebral discs* separate the adjacent vertebrae following the geometric measures of the relating vertebrae. They are roughly cylindrical with a lateral diameter (width) of 40-45 mm and sagittal diameter (depth) of 35-40 mm. The ratio of disc height to height of the vertebral body is about 1:3 in the lumbar region, so the height of the lumbar disc is about 10 mm. During daily activity, when the discs are under compressive load, water is squeezed out of them, so they lose their height. After bed rest and sleep in the night, water flows back into the discs, so the height of them is restored again.

The structure of intervertebral disc tissue is anisotropic. The disc consists of three components: the gelatinous center, the *nucleus pulposus*, surrounded by the concentrically arranged fibrous layers or lamellae of the *annulus fibrosus*, and the superior and inferior cartilaginous endplates.

The nucleus pulposus forms 25-50% of the sagittal cross-area of the disc. It is located more posterior than central in the lumbar spine. The nucleus is a hydrated gel, a semi-fluid mass, an incompressible sphere that exerts pressure in all directions. Although there are significant differences in their structure, there is no clear boundary between the nucleus and the annulus.

The annulus fibrosus consists of 15-25 concentric laminated layers of collagen lamellae tightly connected to each other in a circumferential form around the periphery of the disc. Each lamella consists of ground substance and collagen fibers. Within each lamella the collagen fibers are arranged in parallel, running at an average direction of 30° to the disc's horizontal plane. In adjacent lamellae they run in opposite directions and are therefore oriented at 120° to each other. The outermost fibers called *ligamentous* portion of annulus fibrosus are attached directly to the external rim of vertebral bone, while the internal fibers called *capsular* portion of annulus fibrosus insert into the cartilaginous endplates, forming a continuous

envelope around the nucleus pulposus. Outside the lamellae are mutually connected by the anterior and posterior longitudinal ligaments.

The *endplates* separate the nucleus and annulus from the vertebral bodies. The plates have a mean thickness of 0.6 mm. The cartilaginous endplates cover the superior and inferior surface of the disc, the top and bottom of the nucleus and annulus, binding the disc to its respective vertebral bodies. *Cartilaginous endplates* cover almost the entire surface of the adjacent vertebral bodies, the *bony endplates*, only a narrow rim of bone around the perimeter of the vertebral body is uncovered by cartilage.

2.3 Ligaments

In the lumbar spine seven types of ligaments are distinguished, five of them connect the several parts of posterior elements of vertebrae and two of them connect the vertebral bodies itself.

The *ligamentum flavum* (LF), the most elastic ligament of the spine connects the lower and upper ends of the internal surfaces of the adjacent laminae, closing the gap between the consecutive laminae. The *intertransverse ligaments* (ITL) connect the transverse processes by thin sheets of collagen fibers. The *interspinous ligaments* (ISL) connect the opposing edges of spinous processes by collagen fibers, while the *supraspinous ligaments* (SSL) connect the peaks of adjacent spinous processes by tendinous fibers. The *capsular ligaments* (CL) connect the circumferences of the joining articular facet joints, being perpendicular to the surface of the joints.

The *anterior longitudinal ligament* (ALL) covers the anterior surfaces of the vertebral bodies and discs, attached strongly to the vertebral bone and weakly to the discs. Consequently, it is wider at the bone and narrower at the discs. The *posterior longitudinal ligaments* (PLL) covers the posterior aspects of the vertebral bodies and discs, attached strongly to the discs and weakly to the bone. Consequently, it is wider at the discs and narrower at the bone.

2.4 Muscles

The muscles of the lumbar spine can be distinguished by their location around the spine: *postvertebral* and *prevertebral muscles*. The postvertebral muscles can be divided to *deep*, *intermediate* and *superficial* categories. The prevertebral muscles are the *abdominal muscles*.

The *postvertebral deep muscles* consist of short muscles that connect the adjacent spinous and transverse processes and laminae. The *intermediate muscles* are more diffused, arising from the transverse processes of each vertebra and attaching to the spinous process of the vertebra above. The *superficial postvertebral muscles* collectively are called the *erector spinae*. There are four *abdominal muscles*, three of them encircle the abdominal region, and the fourth is located anteriorly at the midline.

2.5 Functional spinal units or motion segments

A motion segment or functional spinal unit (FSU) is the smallest part of the spine that represents all the main biomechanical features and characteristics of the whole spine. Thus the entire spinal column can be considered as a series of connecting motion segments. The motion segment is a three-dimensional (3D) structure of six degree of statical/kinematical freedom, that is, mathematically, it is 3D in the geometric and 6D in the function space. The FSU consists of the two adjacent vertebrae with its posterior elements and facet joints, and

the intervertebral disc between them, moreover the seven surrounding ligaments, without muscles.

3. Biomechanics of the lumbar spine

The spine is a typical mechanical structure. The biomechanics of the spine concerns the mechanical behaviour of the spine as a living load bearing structure subjected to physiologic and other loads. The spinal column is the main load bearing structure of the human musculoskeletal system. It has three fundamental biomechanical functions: (1) to guarantee the load transfer along the spinal column without instability, (2) to allow sufficient physiologic mobility and flexibility, and (3) to protect the delicate spinal cord from damaging forces and motions. The lumbar spine has a distinguished role in these functions: it has the greatest loads, forces and moments, and the greatest mobility at the same time. As a consequence, the lumbar spine is a very complicated compound mechanical structure.

3.1 Loads acting on the lumbar spine

The loads acting on the spine can be divided into two classes: *physiologic* and *traumatic* loads. The physiologic loads due to the common, normal activity of the spine have further classes: *short-term loads* (in flexion, extension), *long-term loads* (in sitting, standing), *repeated or cyclic loads* (in gait, walk), *dynamic loads* (in running, jumping). The traumatic loads generally occurs suddenly, accidentally with great amplitude (in impact, whiplash). The spinal loads based on biomechanical studies are summarized by Dolan and Adams (2001).

The loads can be classified concerning the *origin* of them (gravity, muscles, etc.). Each part of the body is subjected to *gravity load*, proportionally to its mass. The compressive gravity load increases downwards, towards the support of the body. Standing upright, the weight of the upper body loads vertically the lower lumbar spine. This load can be multiplied in acceleration, during a fall, or other affect with acceleration or deceleration.

Muscle loads depend on the muscle activity. The muscles are *active tissues*, they can contract, their ability of contraction is governed by the nervous system. The back and abdomen muscles stabilize the spine in upright standing position, moreover, they prevent the spine from extreme movements. At the same time, since the strongest muscles run parallel and closed to the long axis of the spine, their contraction subjects it to high compressive forces. According to Nachemson (1981) and Sato (1999), during relaxed standing or sitting, the compressive load from the muscles can double the concerning load of body weight. In bending forwards with lifting weights, the back muscles generate very high tensile forces to equilibrate the effects of the vertically acting upper body load and the lifted weight together (Adams et al. (2002)). Due to this extra large tensile force in the back muscles, the lumbar spine is subjected to a high compression.

The *ligaments* are *passive tissues*, they cannot contract, but can sustain high tensile forces being stretched. Stretched passive ligaments store elastic energy that can release and unload the muscles (Dolan et al., 1994).

The *intra-abdominal pressure* decreases generally the spinal compression due to the abdominal muscle activity. Holding the breath increases the intra-abdominal pressure that increases the spinal stability. Wide abdominal belts help to reduce the spinal compressive forces during lifting. It is proved that a belt can reduce sudden unexpected loading events as well.

Ergonomic loads afflict mostly the lumbar spine. By lifting and holding weights the lumbar spine is subjected to high compressive load, depending on the horizontal distance of the load from the lumbar spine. Long-term vibration and cyclical effects may increase the compression in the lumbar spine leading to structural changes and fatigue effects in the tissue of discs and vertebrae.

It has been proved experimentally, that *traumatic overload* of the spine may cause damage in the discs and facet joints. When a disc is loaded beyond its load-bearing capacity, structural damage may occur as a consequence of high-level short-term loads that exceed load tolerance. However, submaximal long-term and repeated loads may also cause disc failure. Although muscles can save the spine from excessive injurious loads and movements, this protection works only if the neural system has time enough to activate the muscles. This time is very short, however, in most cases it is not enough to avoid the injuries. By static loading the protection of muscles generally does work, still injury happens, mainly by coupled torsion and bending in the morning, by sudden lifting, by tired muscles, by vibrational circumstances. By dynamical effects, by car incidents, by whiplash-like loading effects, the arrival of neural information to the muscles is too late. The unexpected dynamic effect of a simple stumble may cause 30-70% larger compression to the lumbar spine. The defensive reaction of the cervical and dorsal muscles can save the life of a person in a car impact.

3.2 Internal forces arising in the lumbar spine

The main internal force acting on the lumbar spine is the *compressive normal force* acting perpendicularly to the middle plane of the discs, causing high compression in the discs. It is accompanied by mainly sagittal and less lateral *shear forces* acting in the middle plane of the discs, causing the slope of the discs to each other. The *moment components* causing the forwards/backwards bending (flexion/extension) and the lateral bending of the spine are the sagittal and lateral *bending moments*, respectively; and the component that causes the spine to rotate about its long axis is the *torque or torsional moment*. The *tensile force* is also a normal force acting perpendicularly to the middle plane of the discs and causing the elongation of it. Although from physiologic loads there is no pure tensile force acting on the spine, since it acts generally to a part of the discs only as a side effect of other internal forces, however, the aim of traction therapies is even to apply pure tensional force to the lumbar spine.

The *compressive force* arisen in the lumbar spine from the body weight depends on the weight of the trunk, head and arms together. Standing upright this vertical weight load is approximately 55-60% of the total body, that is, about 400 N for the standard body weight of 700 N. Taking into account that the lumbosacral disc has approximately 30° inclination to the horizontal plane, this force can be decomposed to a 350N compressive normal force and a 200N sagittal shear force. These forces can be doubled by the effect of muscle forces, or by the effect of dynamic loads.

The compressive force in the lumbar spine depends strongly on the posture of the body. In laying posture about 150-250 N, by standing erect about 500-800 N, by sitting erect about 700-1000 N can arise in the lumbar spine (Adams et al. (2002)). During forward bending and lifting weight, the magnitude of the compressive forces increases.

Measuring intradiscal pressure in vivo was made originally by Nachemson (1964, 1981), by sticking a pressure-sensitive needle into the L3-4 disc of volunteers. This method has been improved by better technology by Sato (1999), Wilke et al. (1999) and Adams et al. (1996, 2002).

The lumbar *shear force* has never been measured in vivo, thus it can be calculated only, depending on the compressive force and the inclination angle of the actual disc. The shear force is higher in the lower lumbar part due to the higher inclination.

The *bending moments* that play an important role in damaging the discs can be estimated also mainly by mechanical models combined with experimental flexion movement measurements. Adams and Dolan (1991) obtained about 10 Nm bending moment when lifting 100 N weight with bent knees, and 19 Nm with straight knees. Evidently, the lumbar bending moment depends on several variable factors. The back muscles protect the spine from excessive bending during moderate lifting, however, in the case of repeated loading the protective reflex can be eliminated, and the recovery of muscles needs considerable time.

The lumbar spine is subjected to direct *torsion* in some sport or ergonomic activity, accompanied generally by lateral bending. Very little is known about the torsional stresses in the spine in vivo. It is possible to measure the torsional rotation movements in vivo that can be compared with the torque-rotation properties of spine obtained by cadaveric experiments. The upper bound of torque causing damage may be 15-30 Nm in vivo, the safety torque without damage can be 6-12 Nm.

As for the *tensile force*, there are only a few results of measured human spine elongations in pure centric tension, since tensile deformations are analyzed associated with flexion and extension (Bader and Bouten (2000), White and Panjabi (1990)). However, by traction therapies, the lumbar spine is subjected to pure centric tension (Kurutz et al. 2003, Kurutz 2006a, 2006b).

3.3 Mobility of the lumbar spine

The range of spinal movements can be measured both in vivo and in vitro. The spinal motions are described in the space coordinate system, related to the anatomic planes. The spinal movement has six components: three deflections and three rotations. The physiologic movements are the flexion and extension in the sagittal plane, the lateral bending in the frontal plane and the rotation around the long axis of the spine. The spinal motions are generally characterized by three parameters: (1) the *neutral zone* in which the spine shows no resistance, (2) the *elastic zone* in which the spinal resistance works, and (3) the *range of motion*, the sum of the two latter zones. The reference position is the erect standing. The range of motion of a general lumbar segment is about $12-16^\circ$ for flexion/extension, increasing from L1-2 to L4-5; about 6° for lateral bending, and about 2° for axial torsion. The neutral zone is about $2-3^\circ$ for all physiologic motions. The displacement of a motion segment in the neutral zone is about 0.3 mm for tension and compression, and about 0.8 mm for shear. The range of motion is about 0.1-1.9 mm for tension, compression and shear.

The mobility of the spine depends on several factors. It depends first of all on the state of the intervertebral discs: the geometry, the stiffness, the fluid content, the degeneration and aging of it. The smaller its diameter and the larger its height is, the larger its mobility, its any range of motion is. Because of the greater proportional height of the discs, the lumbar region of the spine has greater mobility than the thoracic spine. The range of motion is influenced also by the state of ligaments, the articular facet joint and the posterior bony elements. Viscoelastic properties of discs and ligaments also have an effect on the mobility.

3.4 Degeneration of the lumbar spine

Degeneration means a specific injurious change in the composition, structure and function of the spine. Degeneration is a general terminology, it may be divided basically into two main classes: (1) the long-term *age-related degenerations* and (2) the shorter-term *environmental degenerations* caused by mechanical, chemical, electromagnetic or any other environmental effects. Among the latter category, there is a large class of sudden short-term degenerations caused by unexpected mechanical effects, like a sudden overload or wrong movements, car accidents or other traumatic loads, leading to mechanical damage of the spine.

Age-related degenerations may lead to annular tears or disc prolapse and herniation or osteoporotic failures in the trabecular bone. The forms of environmental degeneration depend on the given special causing effects. Compression overload may cause vertebral failures, endplate fractures cancellous bone collapse, and internal disc disruption, annulus buckling, due to some alarming events, heavy weight lifting in a rapid manner, accidents involving falls, collisions. Shear overload may lead to articular joint damage and fracture, spondylolysis, caused by a heavy back pack or an excessively lordotic posture. Torsion overload can cause impact of articular joint again and anterior annulus tears, caused by a fall or some sports. Forward bending may lead to intervertebral ligament sprains, caused by putting on a shoe, or bending forward suddenly after a long period of flexed posture like car driving, during repetitive bending and lifting, Backwards bending may yield impact of neural arch and posterior disc bulging, caused by an overhead manual work or some sports. Lateral bending may cause ligament sprains and articular joint impaction, caused by special movements during daily activity. However, highest risk of mechanical damage and the most dangerous injuries occur during the combination of the listed mechanical effects that are generally in multiple combinations with each other. Bending and compression combined with torsion has the highest risk to spinal damage. Unexpected, sudden load effects, vibration, repetitive load and fatigue may increase the risk for damage (Adams et al., 2002).

Any age-related or environmental change in the structure of discs or vertebrae leads to the change in the load-bearing capacity of the lumbar segment. Osteoporosis of vertebrae, or decreasing water-binding capacity of nucleus, or calcification of the fibro-cartilaginous endplates, or a disc prolapse of sudden loading or a wrong movement, equally leads to metabolic disorders and biomechanical defects, to limited ability of the required mechanical function and loading tolerance of segments (Adams et al., 2002).

4. Biomechanics of the elements of the lumbar functional spinal unit

The three-dimensional FSU has six force and six motion components that occur in combination with each other. These general force and motion systems depend highly on the mechanical properties, stiffness or flexibility, or load bearing capacity of each structural component of the motion segment.

4.1 Biomechanics of the vertebral body and the articular facet joints

Lumbar vertebral bodies resist most of the compressive force acting down along the long axis of the spine. Most of this load must be resisted by the dense network of trabeculae, and less by the cortical shell. Namely, in the load transfer of axial compression, the nucleus of the disc pressurizes the cartilaginous endplates to bulge inward the cancellous core of vertebrae, when the trabecular bone columns start to buckle due to the excessive load, conse-

quently, first the trabecular bone fails during compression. During this process, radial stresses occur in the endplates, causing cracks in it, to allow the nucleus to bulge also into the vertebral body. Thus, the state of the cancellous bone is the main factor of failure tolerance of vertebrae (McGill, 2000). Moreover, the cancellous bone of vertebrae acts as shock absorber of the spine in accidental injurious effects.

The load bearing capacity of vertebrae depends mainly on the geometry, mass, bone mineral density (BMD) and the bone architecture of the vertebral cancellous bone, which are in correlation with aging, sex and degeneration. *Mosekilde* (2000) demonstrated that age is the major determinant of vertebral bone strength, mass, and micro-architecture. Some papers consider the effect of aging and sex on the compressive strength characteristics of vertebrae. There is a significant decrease of vertical compressive strength and load bearing capacity of vertebral trabecular bone with the development of osteoporosis during aging, occurring in different life periods for men and women (Duan et al. 2001, Keaveny and Yeh 2002). Analysis of regional inhomogeneity inside the vertebral body showed that the weakest part is the central region of the vertebral body (Gong et al. 2005, Briggs et al. 2004, Banse et al. 2001).

The posterior elements of vertebrae (pedicles, laminae, spinous and transverse processes and facet joints) have also important role in the load bearing capacity and mobility of segments. Facet joints work as typical contact structures governed by unilateral conditions, limiting the spinal movements, extension, lateral bending and axial torsion. Failure of the posterior elements, together with facet damage, leading to spondylolisthesis, is generally caused by anteroposterior shear forces. The articular facet joints stabilize the lumbar spine in compression, and prevent excessive bending and translation between adjacent vertebrae. In this manner they are able to protect the disc. Lumbar facet joints are able to resist forces acting perpendicular to the vertical articular surfaces, approximately in the plane of the disc. Thus, they limit the range of axial rotation, with great contact stresses occurring in the joint surfaces.

4.2 Biomechanics of the intervertebral disc

The intervertebral discs provide the compressive force transfer between the two adjacent vertebrae, at the same time, they allow the intervertebral mobility and flexibility. The arrangement of the collagen fibers in the annulus fibrosus is optimal for absorbing the stresses generated by the hydrostatic compression state of the nucleus pulposus in axial loading of the disc, moreover, they play an important role in restricting axial rotation of the spine.

Axial compressive stiffness is higher in the outer and posterior regions than in the inner and anterior regions. Tensile stiffness is higher in the anterior and posterior part than in the lateral and inner regions. Thus, the inner annulus near the nucleus seems to be the weakest area of annulus, and the outer posterior part the strongest region.

In sustained loading the spine shows viscoelastic features. In quasi-static compression the disc creep is 5-7 times higher than the creep in the bony structures of the segment. Thus, the main factor of segment viscosity is the disc, mainly the disc annulus. The creep of the disc depends on the fluid content of it. A 3 hours long 1200 N compressive loading yields a 10% decrease in disc height and 5-13% increase in the sagittal diameter (Adams and Hutton, 1983, McNally and Adams, 1992) mainly due to fluid loss, similarly to the diurnal variation, namely the effect of overnight bed rest with fluid recovery.

The load bearing capacity of segments is mainly influenced by the degeneration state of the disc. Degeneration means an injurious change in the function and structure of the disc,

caused by aging or by environmental effects, like mechanical overloading (Adams et al., 2000). Degeneration of FSU starts generally in the intervertebral discs. Changes to any tissue property of the disc markedly alter the mechanics of load transfer and stability of the whole segment (Ferguson and Steffen, 2003). The first age-related changes of disc occur within the nucleus.

Long-term age-related degeneration of the disc is manifested in the loss of hydration, a drying and stiffening procedure in the texture of mainly the nucleus (McNally and Adams 1992; Adams et al. 2002; Cassinelli and Kang, 2000). The functional consequences of aging are that the nucleus becomes dry, fibrous and stiff. The volume of nucleus and the region of hydrostatic pressure of it decrease, consequently, the compressive load-bearing of the disc passes to the annulus. However, the annulus becomes weaker with aging, so the overloading of it can lead to the inward buckling of the internal annulus, or to circumferential or radial tears, fiber break in the annulus, disc prolapse or herniation, or to large radial bulging of the external annulus, reduction of the disc height, or moreover, to endplate damages (Natarajan et al., 2004). The main cause of all these problems is that while the healthy disc has a hydrostatic nucleus, it becomes fibrous during aging, being no longer as a pressurized fluid.

Short term sudden degeneration or damage may yield the sudden loss of hydrostatic compression in nucleus, accompanied or due to some other failures mentioned above. Several recent studies concluded that light degeneration of young discs led to instability of lumbar spine, while the stability restored with further aging (Adams et al., 2002).

4.3 Biomechanics of the ligaments and muscles

The ligaments are passive tissues working only against tension. The primary action of the spinal ligaments lying posterior to the centre of sagittal plane rotation is to protect the spine, by preventing excessive lumbar flexion. However, during this protection the ligaments may compress the discs by 100% or more. Indeed, the effectiveness of a ligament, its contribution to the integrity of the spine depends mainly on the moment arm through which it acts.

The most elastic ligament, the ligamentum flavum being under pretension throughout all levels of flexion prevents any forms of buckling of spine. The interspinous and supraspinous ligaments may protect against excessive flexion. The capsular ligaments of facet joints restrict joint flexion and distraction of the facet surfaces of axial torsion.

The failure strength of the lumbar ligaments are about 450 N for the anterior longitudinal, 330 N for the posterior longitudinal, 220 N for the capsular, 120 N for the interspinous ligaments and 280 N for the ligamentum flavum (Benzel, 2001).

The muscles and the neuromuscular controls are required (1) to provide dynamic stability of the spine in the given activity and posture, and (2) to provide mobility during physiologic activity, moreover (3) to protect the spine during trauma in the post-injury phase. Two mechanical characteristics are necessary to provide these physiologic functions: (1) the muscles must generate forces isometrically and by length change, and (2) they must increase the stiffness of the spinal system.

4.4 Biomechanics of the functional spinal unit

The mechanical behaviour of the FSU depends on the physical properties of its components, mainly on behaviour of the intervertebral disc, ligaments and articular facet joints.

The average *load tolerance* of lumbar segments under quasi-static loading is about 5000 N for compression, 2800 N for tension, 150 N for shear and 20 Nm for axial rotation (Bader and Bouten, 2000).

Flexibility of the FSU is the ability of the structure to deform under the applied load. Inversely, the *stiffness* is the ability to resist by force to a deformation. The stiffness of the spinal segments increases from the cervical to the lumbar regions for all loading cases. In lumbar region the stiffness is about 2000-2500 N/mm for compression, 800-1000N/mm for tension, 200-400 N/mm for lateral and 120-200 N/mm for anterior/posterior shear. The rotational stiffness is about 1.4-2.2 Nm/degree for flexion, 2.0-2.8 Nm/degree for extension, 1.8-2.0 Nm/degree for lateral bending and 5 Nm/degree for axial torsion (White and Panjabi, 1990, Bader and Bouten, 2000). The stiffness of the lumbar spine depends on the age and degeneration. In advanced degeneration the stiffness is higher. The stiffness is influenced by the viscous properties of the segments and the load history as well.

5. Geometrical modeling of lumbar functional spinal unit

Geometrical modeling of the FSU needs precise geometrical data of the real object; it must follow the anatomy of the segment. Beside the topology, additional data such as volume density, surface texture, etc. are needed. Different methods of acquisition of geometrical data can be used, like scanners, computer tomography, or magnetic resonance imaging methods.

In FE modeling the vertebral body, its cortical shell, cancellous core, posterior bony elements and the bony endplates are generally distinguished. For the thickness of the vertebral cortical wall Lu et al. (1996) used 1.5 mm, and for the thickness of the cartilaginous endplates considered 1 mm. Baroud et al. (2003) applied 1 mm cortex and 0.5 mm endplate thickness. For the thickness of the cartilage layer of facet joint Schmidt et al., (2009) considered 0.2 mm.

In FE modeling the intervertebral disc, its nucleus, annulus ground substance, annulus fibers and the cartilaginous endplates are generally distinguished. For the volumetric relation between annulus and nucleus, ratio 3:7 is generally used for the lumbar part L1-S1 (Goto et al., 2002, Moramarco et al., 2010). Chen et al. (2001) considered 30-50% of the total disc area in cross section as the nucleus and the rest of the region as the annulus. Lu et al. (1996) applied 38% nucleus area ratio to the total disc cross-sectional area, based on the measurements of a normal disc. Baroud et al. (2003) assumed the nucleus to occupy the 43% of the total disc volume. The diameter length of the disc from anterior to posterior end is about 36 mm, the lateral length is about 44 mm. For the orientation of annulus fibers Rohlmann et al. (2007) considered to the mid cross-sectional area of the disc under alternating direction of about 30° and 150°.

6. Material modeling of lumbar functional spinal unit

Since FSU is a highly heterogeneous compound structure, the material modeling must be related to the components of it. First the material models of the healthy components are considered.

6.1 Material models of the vertebral body and the articular facet joints

The high strength *vertebral cortical shell* is generally considered linear elastic isotropic or transversely isotropic, orthotropic material, seen in Table 1. Vertebral *cancellous bone* is modeled generally by linear elastic isotropic or transversely isotropic or orthotropic material, seen in Table 2.

Vertebral cortical bone				
Material model	E [MPa]	ν	G [MPa]	References
linear elastic, isotropic	5 000	0.3		Rohlmann et al. 2006b, Zander et al. 2006,
linear elastic, isotropic	10 000	0.3		Argoubi, Shirazi-Adl 1996, Kumaresan et al. 1999, Rohlmann et al. 2006a, 2006c, 2007,
linear elastic, isotropic	11 300	0.2		Little et al. 2008,
linear elastic, isotropic	12 000	0.3		Shirazi-Adl et al. 1984, Cassidy et al. 1989, Lavaste et al. 1992, Goel et al. 1995a, 2002, Lu et al. 1996b, Smit et al. 1997, Wang et al. 2000, Chen et al. 2001, 2008, Baroud et al. 2003, Zhong et al. 2006, Denoziere and Ku, 2006, Williams et al. 2007, Ruberté et al. 2009, Zhang et al. 2009, Kurutz and Oroszváry 2010
linear elastic transversely isotropic	11 300 11 300 22 000	0.48 0.20 0.20	3 800 5 400 5 400	Lu et al. 1996a, Schmidt et al. 2009,
linear elastic, transversely isotropic	8 000 8 000 12 000	0.40 0.23 0.35	2 857 3 200 3 200	Noailly et al. 2005, 2007, Malandrino et al. 2009,
poroelastic	10 000	0.3		Ferguson et al. 2003,

Table 1. Material properties for the FE models of lumbar vertebral cortical bone

Vertebral cancellous bone				
Material model	E [MPa]	ν	G [MPa]	References
linear elastic, isotropic	10	0.2		Shirazi-Adl et al. 1984, Cassidy et al. 1989, Lu et al. 1996b, Smit et al. 1997, Goel et al. 1995a, 2002, Wang et al. 2000, Chen et al. 2001, 2008, Denoziere and Ku 2006, Zhong et al. 2006, Ruberté et al. 2009,
linear elastic, isotropic	50 81 140 100 100 150 500	0.2 0.2 0.2 0.29 0.3 0.3 0.2		Rohlmann et al. 2006a, Baroud et al., 2003, Little et al., 2008, Zhang et al. 2009. Lavaste et al., 1992, Kurutz and Oroszváry, 2010, Rohlmann et al. 2006b, Zander et al., 2006,
poroelastic	100	0.2		Argoubi and Shirazi-Adl 1996, Williams et al. 2007
linear elastic transversely isotropic	200 140	0.45 0.315		Rohlmann et al. 2006c, 2007,
linear elastic transversely	140 140	0.45 0.176	48 77	Noailly et al. 2005, 2007, Malandrino et al. 2009,

isotropic	250	0.315	77	Lu et al. 1996a, Schmidt et al. 2009,
linear elastic	140	0.45	48.3	
transversely	140	0.32	48.3	
isotropic	200	0.32	48.3	

Table 2. Material properties for the FE models of lumbar vertebral cancellous bone

The high strength *bony endplate* of vertebrae and the lower strength *cartilaginous endplate* of disc can hardly be distinguished in FSU. The authors generally give information about it when specifying material properties. Table 3 shows the generally applied material moduli of the endplates.

Endplate			
Material model	E [MPa]	ν	References
<i>bony</i> , linear elastic, isotropic	12 000	0.3	Baroud et al. 2003,
	1000	0.3	Noailly et al. 2005, 2007,
	1000	0.4	Chen et al. 2001, 2008,
	500	0.4	Lavaste et al. 1992,
	500	0.3	Zhang et al. 2009,
	100	0.4	Kurutz and Oroszváry 2010,
<i>cartilaginous</i> , linear elastic, isotropic	24	0.4	Shirazi-Adl et al. 1986, Goel et al. 1995a, 2002, Noailly et al. 2005, 2007, Wang et al. 2000, Zhong et al. 2006, Ruberté et al. 2009, Lu, et al. 1996, Schmidt et al. 2009
<i>cartilaginous</i> , poroelastic	5	0.1	Argoubi and Shirazi-Adl 1996,
	5	0.17	Malandrino et al. 2009,
	20	0.4	Williams et al. 2007,
<i>bony</i> , outer:	12 000	0.3	Denoziere and Ku 2006
intermediate:	6 000	0.3	
central:	2 000	0.3	

Table 3. Material properties for the FE models of lumbar vertebral and disc endplates

The *posterior bony elements* are considered linear elastic isotropic material, generally by the same Young's modulus $E=2500$ MPa and Poisson's coefficient $\nu=0.25$ or $\nu=0.2$.

The *articular facet joints* are considered as unilateral frictionless connections transmitting only compressive forces with an initial gap of generally 0.5 mm (Rohlmann et al. 2006a, Chen et al., 2008, Zhang et al., 2009); or 0.6 mm of a nonlinear frictionless contact problem (Schmidt et al., 2009); or assuming soft contact with exponentially increasing contact force with decreasing contact gap (Sharma et al., 1995, Rohlmann et al. 2007). Zhong et al. (2006) used 1 mm initial gap for the surface-to surface frictional contact with friction coefficient 0.1. Little et al. (2008) used a finite sliding frictionless tangential relationship with softened contact in the normal direction that means exponentially increasing contact stresses with initial gap of 0.1 mm.

6.2 Material models of the intervertebral disc

The *intervertebral disc* is the most critical component of the spine in both its mobility and load bearing ability, therefore its FE modeling has a great importance.

6.2.1 Material models of the nucleus pulposus

Nucleus pulposus is the most important element in the compressive stiffness of the disc: the hydrostatic compression in it guarantees the stability of the whole disc and segment. The healthy young nucleus is generally modeled as an incompressible fluid-like material, seen in Table 4.

Nucleus pulposus			
Material model	E [MPa]	ν	References
Fluid-like solid, linear elastic, isotropic	1	0.499	Shirazi-Adl et al. 1984, 1986, Goel et al. 1995a, Chen et al. 2001, Zhong et al. 2006, Denoziere and Ku 2006, Zhang et al. 2009, Ruberté et al. 2009, Kurutz and Oroszváry 2010,
	4	0.499	Shirazi-Adl et al. 1984, Lavaste et al. 1992, Goel et al. 1995b, Fagan et al. 2002.
	10	0.4	Chen et al., 2008,
Incompressible fluid			Lu et al. 1996, Little et al. 2008, Zander et al. 2006, Rohlmann et al. 2006a, 2006b, 2006c
Quasi incompressible			Rohlmann et al. 2007
Hyperelastic, neo-Hookean			Moramarco et al. 2010,
Mooney-Rivlin incompressible			Smit et al. 1997, Noailly et al. 2007, Baroud et al. 2003, Schmidt et al. 2007, 2009,
Poroeleastic	varied	0.17	Malandrino et al. 2009,
	1	0.45	Williams et al. 2007,
	1.5	0.1	Argoubi and Shirazi-Adl 1996,
	1.5	0.17	Ferguson et al. 2004,
Viscoelastic solid	2	0.49	Wang et al. 2000,
Osmoviscoelastic	0.15	0.17	Schroeder et al. 2006

Table 4. Material properties for the FE models of lumbar disc nucleus

6.2.2 Material models of the annulus fibrosus

Annulus fibrosus represents a typical composite-like material with a ground substance and fiber reinforcements of many layers, seen in Table 5.

Annulus fibrosus					
Material model	Ground substance		Fibers		References
	E [MPa]	ν	E [MPa]	ν	
Linear elastic, isotropic matrix, tension only elastic fibers	4	0.4	500	-	Lu et al. 1996,
	4	0.45	500	0.3	Fagan et al. 2002)
	4	0.45	400/500/300	0.3	Kurutz and Oroszváry 2010,
	4.2	0.45	450	-	Zhong et al. 2006,
	4.2	0.45	450	0.3	Shirazi-Adl et al. 1984, 1986,
	4.2	0.45	175	-	Chen et al. 2001,
	4.2	0.45	360/420/485/550	0.3-	Denoziere and Ku, 2006,
	2	0.45	500	-	Lavaste et al. 1992,

	2 8 10 4.2 4.2	0.45 0.45 0.4 0.45 0.45	500 360/420/485/500 360/385/420/440/ 495/550 nonlinear 500	- 0.3	Goel et al. 1995b, Baroud et al. 2003, Chen et al. 2008, Goto et al. 2002, Zhang et al. 2009,
Hyperelastic matrix, nonlinear outwards stiffen- ing fibers	3.15	0.45			Schmidt et al. 2007, 2009, Rohlmann et al. 2006a, 2006b, 2006c, 2007, Ruberté et al. 2009, Noailly et al. 2007, Moramarco et al., 2010,
Hyperelastic matrix, linear elastic tension only fibers			500	0.3	Little et al. 2008,
Viscoelastic ma- trix, nonlinear elastic fibers	8	0.45			Wang et al. 2000,
Poroelastic ma- trix, nonlinear elastic fibers	2.5 2.5	0.4 0.1			Williams et al. 2007, Argoubi and Shirazi-Adl 1996,
Poroelastic ma- trix, linear elastic fibers	2.5	0.17	60	0.33	Ferguson et al. 2004,

Table 5. Material properties for the FE models of lumbar disc annulus

6.3 Material models of the ligaments

Numerical modeling of *ligaments*, as typical exponentially stiffening soft tissues is not a simple task. Generally, the seven ligaments are incorporated to the FE models as tension only elements. In contrast to its strong nonlinear behaviour (White and Panjabi, 1990), most of the reported FEM studies have adopted linear elastic models (Lavaste et al., 1992, a et al., 1995, Zhong et al, 2006, Chen et al, 2008), but bilinear models (Pintar et al., 1992, Goel et al, 1995a, Chen et al., 2001, Goto et al., 2002, Denoziere and Ku, 2006, Moramarco et al, 2010), moreover, trilinear approaches are also used (Pintar et al, 1992, Ruberté et al, 2009), seen in table 6.

Lumbar Ligaments								
Liga- ments	E_1 [MPa]	ε_1 [%]	E_2 [MPa]	ε_2 [%]	E_3 [MPa]	ε_3 [%]	CS area [mm ²]	References
ALL	20						63.7	Goel et al. 1995a, Zhong et al. 2006
PLL	20						20	
LF	19.5						40	
ITL	58.7						3.6	
CL	32.9						60	
ISL	11.6						40	
SSL	15						30	
ALL								
PLL	70						20	

LF	50						60	Chen et al. 2008,
ITL	50						10	
CL	20						40	
ISL	28						35.5	
SSL	28						35.5	
ALL	7.8	12	20				63.7	Goel et al. 1993, 1995a, Denoziere and Ku 2006,
PLL	10	11	50				20	
LF	15	6.2	19				40	
ITL	10	18	59				1.8	
CL	7.5	25	33				30	
ISL	8	20	15				30	
SSL	10	14	12				40	
ALL	7.8	12	20				63.7	Chen et al. 2001,
PLL	10	11	20				20	
LF	15	6.2	19.5				40	
ITL	10	18	58.7				1.8	
CL	7.5	25	32.9				30	
ISL	10	14	11.6				30	
SSL	8	20	15				40	
ALL	7.8	12	20				32.4	Pintar et al. 1992, Goel et al, 1995a, Goto et al. 2002, Moramarco et al, 2010,
PLL	1	11	2				5.2	
LF	1.5	6.2	1.9				84.2	
ITL	10	18	59				1.8	
CL							43.8	
ISL							35.1	
SSL	3	20	5				25.2	
ALL	12.6	8	15.6				32.5	Pintar et al. 1992, Ruberté et al, 2009,
PLL	27.1	7	40	25	31.6	38	5	
LF	24	8	40	20	36	25	91.6	
ITL	125	8	313				2	
CL	7.5	25	12.7				51.2	
ISL	4.15	20	11.4				34	
SSL	4.15	20	11.4				34	

Table 6. Material properties for lumbar ligament FE models

Here the transition strains $\varepsilon_1, \varepsilon_2$ and ε_3 separate the concerning Young's moduli of the polygonal stress-strain function. The generally used - very divergently applied - cross sectional areas of each ligament are also illustrated in Table 6. Several further nonlinear FEM models of ligaments can be found in the literature (Shirazi-Adl, 1986a, 1986b, Wang et al., 2000, Zander et al., 2006, Noailly et al., 2007, Rohlmann et al., 2006a, 2006b, 2006c, 2007, Williams et al., 2007, Schmidt et al., 2009). Based on several FE studies Eberlein et al., (2004) summarize the ligament models applied in FEM analyses and suggests a membrane model with a new constitutive equation as the special case of the equation obtained for the annulus fibrosus.

7. Material and geometrical modeling of the degenerated lumbar functional spinal unit

Aging type degeneration starts generally in the nucleus. A healthy young fluid-like nucleus is in a hydrostatic compression state. During aging, the nucleus loses its incompressibility and becomes even stiffer and stiffer, changing from fluid to solid material. This kind of nucleus degeneration can be modeled by decreasing Poisson's ratio with increasing Young's modulus (Kurowski and Kubo, 1986; Kim et al., 1991). This behavior is generally accompanied by the stiffening process of the disc as a whole and by the volume reduction of the nucleus and volume extension of the annulus, furthermore, height reduction of the disc. Moreover, at the same time, annulus tears or internal annulus buckling, or break of the annular fibers, damage and crack or rupture of endplates, osteoporotic defects of vertebral cancellous bone can happen. Consequently, modeling age-related degeneration of FSU is a compound task; it must be done in its progress, relating to a lifelong process.

In contrast to the age-related degeneration, the nucleus may lose its incompressibility without any stiffening and volume change process. due to a *sudden unexpected traumatic load* effect. In this case the nucleus may quasi burst out and the hydrostatic compression may suddenly stop in it. This kind of nucleus degeneration can be modeled by suddenly decreasing Poisson's ratio with unchanged Young's modulus of nucleus (Kurutz and Oroszváry, 2010). This behaviour is generally caused or accompanied by the tear or buckling of the internal annulus, break of the annular fibers, fracture of endplates, or collapse of vertebral cancellous bone, depending on the age in which the accidental event happens. Namely, accidental failures can happen in a young disc, as well, or in any age and aging degeneration phases. These effects can be modeled by sudden damage of tissues of the concerning components of the segment. In contrast to the long term aging degeneration, these kinds of damage instability occurs suddenly, generally due to a mechanical overloading (Acaraglou et al., 1995).

Rohlmann et al. (2006c) have developed a FE model of a lumbar motion segment of different grades of age-related disc degenerations to simulate the effect of degeneration on the biomechanical behaviour of the segment. They introduced three grades of disc degeneration: mild, moderate and severe degenerations. Compared to the healthy disc, the three grades have 20, 40 and 60% less disc height, respectively. Parallel to the disc height reduction, the length of the annulus fibers was also reduced, compensated by offsetting their nonlinear stiffness curves. The facet orientation was also changed with disc height reduction. The compressibility of nucleus was increased with the loss of fluid-like behaviour from 0.0005 to 0.15 mm²/N, by using linear interpolation for the different grades of degeneration (mild: 0.0503, moderate: 0.0995 mm²/N). It was assumed that the disc degeneration has no effect on the material properties of the annulus fibrosus. By analyzing a healthy and a slightly degenerated lumbar spine Rohlmann et al. (2007) assumed compressible nucleus with the compressibility of 0.0005 mm²/N for healthy, and 0.0503 mm²/N for mildly degenerated case.

Schmidt et al. (2007) verified the hypothesis that with increasing disc degeneration, the internal pressure and strains of the disc decrease, therefore, the risk of disc prolapse decreases. They assumed mildly, moderately and severely degenerated disc with 16.5, 49.5 and 82.5% reduced height, respectively, and by supposing increasing osteophytes formations with progressing degeneration (1.5, 4.5 and 7.5 mm, respectively). Parallel to the disc height reduction, the facet orientation was also changed, compared to the parallel position of the surfaces (0°), the angle between them increased (0.8, 1.9 and 2.9°). Moreover, it was

supposed that the endplate flattens with progressing disc degeneration. By assuming 0% for the healthy endplate curvature and 100% for the planar endplate, the degeneration percents followed the percents of the disc height reduction. Since the disc bulging increase with the progress of degeneration, and however, there were no data available in the literature, it was assumed the same ratio for bulging increase as for disc height decrease. In the length of the annulus fibers, the results of Rohlmann et al (2006c) were used. The Young's modulus of the nucleus was increased during degeneration from the value of the healthy nucleus to the value of the annulus ground substance by supposing that the nucleus and the annulus become structurally similar with increasing disc degeneration, and that the disc degeneration has no effect on the material properties of the annulus fibrosus.

Ruberté et al. (2009) simulated aging degeneration of L4-5 lumbar segment by introducing three grades of degeneration: healthy, mild and moderate phases; by decreasing disc height (12.0, 10.2 and 8.0 mm) and nucleus area (388, 269 and 101 mm²) with increasing annulus area (731, 850 and 1022 mm²), and by modifying the material properties of annulus ground substance (Mooney-Rivlin C_1/C_2 : 0.2/0.05, 0.4/0.1 and 0.9/0.23) and nucleus (Young's modulus/Poisson's coefficient: 1.0/0.49, 1.26/0.45 and 1.66/0.4). The nucleus area was reduced by following the stress profilometry results of Adams et al. (1996). Material properties were taken from the literature and from the results of Umehara et al. (1996), Iatridis et al. (1997) and Elliott and Setton (2001). Umehara et al. (1996) measured experimentally the distribution of the compressive elastic moduli in the lumbar intervertebral disc in term of degeneration. The distribution of elastic moduli in normal discs was symmetric about the midsagittal plane, degenerated discs showed irregular distributions of elastic moduli. The elastic moduli of the degenerated nucleus were higher than those in normal discs.

Zhang et al. (2009) applied healthy, and two degenerated grades in modeling L4-5 motion segment. For grade 1 the elastic modulus of the disc nucleus was two times the elastic modulus of the annulus in the intact model and the Poisson's ratio was adopted to be the same as that of the annulus, and the disc height was reduced by 20%, following Kim et al. (1991), Iatridis et al. (1997) and Kumaresan et al. (2001). For grade 2, in addition, the elastic modulus of the annulus was doubled, and the annulus fiber volume was reduced by 25%, and the disc height was reduced by 40%, by considering Rohlmann et al. (2006c).

Kurutz and Oroszváry (2010) introduced five grades of age-related degeneration from healthy (1) to fully degenerated (5) cases, modeled the loss of hydrostatic state in nucleus by decreasing Poisson's ratio ($\nu=0.499, 0.45, 0.40, 0.35, 0.30$, respectively), accompanied by nucleus stiffening modeled by increasing Young's modulus ($E=1, 3, 9, 27, 81$ MPa, respectively). Simultaneously, in the annulus matrix a gradual increase ($E=4.0, 4.5, 5.0, 5.5, 6.0$ MPa), while in the vertebral cancellous bone ($E=150, 125, 100, 75, 50$ MPa) and endplates ($E=10, 80, 60, 40, 20$ MPa) a gradual decrease of Young's modulus were considered with aging. The age-related change of tensile moduli of annulus ground substance and nucleus equally ($E=0.4, 1.0, 1.6, 2.2, 2.8$ MPa) were modeled by using a parameter identification method based on the in vivo measured lumbar disc elongations by Kurutz et al. (2003, 2006a, 2006b).

Most of the geometric models of the degenerated disc apply the reduction of the disc height. Lu et al. (1996) by using FE simulation concluded that variation in disc height had a significant effect on the axial displacement, the posterolateral disc bulge and the tensile stress in the peripheral annulus fibers, but the influence on the intradiscal pressure and the longitudinal stress distribution at the endplate-vertebra interface was minimal.

Permeability is a key factor in poroelastic FE analysis, representing the ability of interstitial fluid flow within the tissues. Experimental studies have shown a large variation of this parameter. Gu et al. (1999) and Johannessen and Elliott (2005) correlated anisotropic nucleus and annulus permeabilities with disc degeneration.

Gu et al. (1999) concluded that the fluid transport within a disc is crucial to its viscoelastic behavior, fluid pressure redistribution, and cell nutrition, in viscoelastic behaviors of healthy and degenerate discs as well as the biomechanical etiology of disc failure. The hydraulic permeability of human non-degenerated annulus fibrosus is direction-dependent, namely, anisotropic, with the greatest permeability in the radial direction. With disc degeneration, the radial permeability of annulus decreases, mainly because of the decreased water content, and the axial and circumferential permeability coefficients increase, mainly because of the structural change, leading to more isotropic permeability behavior for more degenerated discs.

Johannessen and Elliott (2005) by measuring the biphasic compressive material properties of normal and degenerate human nucleus pulposus tissue in confined compression concluded that swelling is the primary load-bearing mechanism in both nondegenerate and degenerate nucleus pulposus. Degeneration produced significant decreases in swelling stress, while permeability increased with degeneration.

Malandrino et al. (2009) aimed to study the poromechanical responses of the L3-4 disc FE model under compression, flexion and rotation, by varying the Young's moduli and permeabilities of the tissues of the main components of disc. This may represent a typical degeneration process. They considered the Young's modulus for the healthy nucleus as 1 MPa, for the healthy annulus 2.56 MPa; and for the fully degenerated case for the nucleus 1.66 MPa and for the annulus 12.29 MPa (Iatridis et al., 1998, Natarajan et al., 2006). Four hydraulic permeabilities were chosen, for annulus, nucleus, cartilaginous endplate and for the trabecular bone. This allowed studying the significance of fluid-solid interaction in the disc.

8. Loads in finite element simulation

Loads on lumbar spinal motion segments in FE modeling depend on the aims of the analysis. The segment is generally supported rigidly along the inferior endplate of the lower vertebra, thus, the loads are generally applied on the superior endplate of the upper vertebra.

The loads can be applied as *static or dynamic loads*. Constant static loads or *incrementally changing quasi-static loads* are generally applied in lumbar spine analyses. The basic loading types are the *force or displacement type loads*, in a *load or displacement controlled device*, in a *load history analysis*.

Chen et al. (2001) in analyzing the adjacent segment syndrome of a rigid fixation used equally 10 Nm for flexion, extension, lateral bending and axial torsion under 150 N preload.

Goto et al (2002) investigated the numerical analysis of lumbar vertebrae of L4-5 segment by using intradiscal pressure in nucleus to establish the model, and by applying incremental loading device. Compressive loading was performed to 294 N in ten steps, then flexion and extension loads of 15 Nm was applied in 15 steps. The intradiscal pressure was set as 1.32 MPa for flexed and for standing position; 0.6 MPa for extended position, and zero pressure was assumed for degenerated disc model.

Baroud et al. (2003) applied displacement control in load history FE analysis of vertebroplasty of the L4-5 segment, by applying quasi-static compression load of 2.8 mm in steps of 0.2

mm. Ferguson et al. (2004) analyzed the fluid flow within the disc by simulating the diurnal loading cycle consisting of an 8 hours resting period, followed by a 16 hours constant compressive load equivalent to 0.5 MPa average mechanical stress. The tendency of the disc to swell due to an osmotic potential was simulated by the addition of a 0.2 MPa pressure at the bony endplate and the annulus. The loading cycle was discretized into fixed time steps.

Denoziere and Ku (2006) applied for physiologic loads a pre-compression of 720 N simulating the intervertebral pressure of standing position and additional compressive forces were applied to the suitable areas of the endplates to simulate severe motions: 2000 N for lifting a load with straight legs in full flexion; 1000 N for full extension; 1300 N for full lateral bending; 11.45 Nm axial torque for axial rotation, based on the work of Nachemson (1966), White and Panjabi (1990) and Wilke et al. (1999). Zhong et al. (2006) analyzed a new cage by topology optimization by applying the maximum possible load without causing spinal injury, thus, 10 Nm for flexion, extension and torsion and lateral bending with 150 N compressive preload.

By analyzing rigid and dynamic fixation Rohlmann et al. (2007), vertebro- and kyphoplasty Rohlmann et al. (2006a) and dynamic implant Zander et al. (2006) applied follower load. An upper body weight of 260 N and a compressive follower load of 200 N were considered standing, flexion and extension. The follower load simulated the stabilizing effect of the local muscle forces. The lever arms to the disc center of erector spinae and rectus abdominis were 40 and 153 mm, respectively, and the direction of the muscle forces was quasi parallel to the lumbar spine. For axial rotation, a 10° axial pre-rotation was applied, and a follower load of 500 N representing the upper body weight and the muscle forces together was considered. Spinal loads were assumed to be 25% higher for walking than for standing. The method to estimate the muscle forces for standing and different inclination of the spine in the sagittal plane has been described in details by Zander et al. (2001), Wilke et al. (2003) and Rohlmann et al. (2006b).

Schmidt et al. (2007) in a degeneration analysis used unconstrained moment load 7.5 Nm with changing loading directions between each pair of main anatomical planes to simulate the combinations of the anatomical loadings, for example rotation with lateral bending, and so on. All these load cases were additionally combined with an axial compressive preload of 500 N. Interbody fusion and fixation techniques were compared by Chen et al. (2008) by applying a compressive preload of 150 N together with four different kinds of 10 Nm moments simulating the physiologic loading cases. Zhang et al. (2009) evaluated the load transfer of a dynamic stabilization device under compression, by applying axial compressive force of 2000 N for validating the model and 1000 N to investigate the load transmission characteristics of different implants.

Schmidt et al. (2009) analyzed flexible lumbar stabilization system by applying pure unconstrained moments in the three anatomical main planes, simulating flexion, extension lateral bending and axial rotation. The loads were increased incrementally in 10 load steps from zero to the predetermined maximum values of 7.5 and 20 Nm. For poroelastic disc analysis for physiologic loading Malandrino et al. (2009) used 1000 N compression force load and 7.5 Nm moment load, similarly to Noailly et al. (2007).

By analyzing the effect of degenerative disc disease on adjacent segments Ruberté et al. (2009) applied different moment loads of flexion, extension, axial rotation and lateral bending for healthy (8, 6, 4 and 6 Nm, respectively) for healthy model, and these moments were

modified for mild and moderate degenerations. A compressive follower preload of 800 N was also applied.

Poroelastic creep analysis of a lumbar segment was investigated by Argoubi and Shirazi-Adl (1996). The creep response of the segment was studied for a period of 2 hours under a constant axial force of 400, 1200 and 2000 N. Viscoelastic analysis of segment was investigated by Wang et al. (2000) for combined compression and sagittal flexion. A 600 N axial compressive load with a 60 N anterior shear load for a duration of 30 seconds simulated the preload of physiologic neutral posture. The hybrid loading mechanism was used to simulate the lowering task by specifying the final net compression of 2000 N, the anterior shear force of 200 N, and the sagittal flexion of 10° . The three different durations of 0.3, 1 and 3 seconds represented the fast, medium and slow movements, respectively.

Dynamic loading of the L4-5 segment with poroelastic disc was analyzed by Williams et al. (2007) by applying short-term creep and cyclic loading. In simulating the short-term creep and standing recovery, a 400 N compressive preload was applied on the superior surface of L4, followed by an additional 400 N for 20 minutes, after which the load was reduced to 400 N for 10 minutes for recovery. In simulating the short-term cyclic load and standing recovery, after the 400 N compressive preload, a peak-to-peak compressive force of 400 N and a peak-to-peak flexion moment of 5 Nm at a rate of 12 lifts per minute was applied for 20 minutes, then the disc was allowed to relax for 10 minutes with just the preload present.

Moramarco et al. (2010) validated the FE model of a lumbosacral segment by applying four simulation by different loading cases. First, an incremental 4 Nm pure flexion moment was applied. Then an axial compressive pre-loading of 100 N was applied first with a flexion moment than with an extension moment, both of 10 Nm, applied incrementally. Finally, 10 Nm lateral moment was considered.

Kurutz and Oroszváry (2010) analyzed by FE simulation the stretching effect of a special underwater traction treatment when the patients are suspended cervically in vertical position in the water, supported on a cervical collar alone, loaded by extra lead weights on the ankles. There were two parts of the applied traction load: the removal of the compressive preload of body weight and muscle forces in water, named indirect traction load; and the direct traction load consisting of the tensile force of buoyancy with the applied extra loads. Based on mechanical calculations, for the standard body weight of 700 N, and the applied extra lead weights 40 N, the indirect and direct traction loads yields 840 N and 50 N, respectively.

9. Validation of the finite element models

By using FE models in a numerical simulation, the results should be trustworthy. But how do we know that the results or predictions can be believed with confidence? FE model validation can answer these questions. Correlation between FE results and experimental results can lead to use the FE model predictions with confidence. Recently there are many technologies for evaluating and improving the accuracy and validity of linear and nonlinear FE models.

In numerical simulations of biomechanics, for FE prediction accuracy assessment, a gold standard can be the *experimental validation of numerical results*. This enables the analyst to improve the quality and reliability of the FE model and the modeling methodology. If there is very poor agreement between the analytical and experimental data, by using certain nu-

merical techniques for model updating allow the user to create improved models which represent reality much better than the original ones.

For example, analyzing the effect of disc height Lu et al. (1996) validated the FE model by direct comparison of the model predictions with experimental results of axial displacement, axial compressive stress and posterolateral disc bulge obtained on cadaveric motion segments. In viscoelastic analysis Wang et al. (2000) validated the segment model by the experimental data of constant compressive strain rate loading, creep loading and cyclic relaxation loading. In analyzing the adjacent segment syndrome Chen et al. (2001) validated the L1-5 multisegment model by comparing the kinematics data of the model with in vivo experiments under the same loading condition. In fluid flow analysis within the disc Ferguson et al. (2004) validated the FE model against in vitro creep/swelling data for isolated discs. Denozziere and Ku (2006) validated the segment model by comparing the average mobility of the healthy model in flexion-extension, lateral bending and axial rotation with various experimental reports on vertebral motions. In short-term creep and cyclic analysis Williams et al. (2007) validated the model by in vivo creep and recovery disc height variations. For interbody fusion and fixation analysis Chen et al. (2008) validated the intact FE model by comparing the flexion-extension angles with experimental data. Moramarco et al. (2010) validated the FE model of a lumbosacral segment by in vitro experiments of axial displacements and posterior disc bulge. Kurutz and Oroszváry (2010) validated the lumbar segment model for both compression and tension and for both healthy and degenerated disc. Distribution of vertical compressive stresses of healthy and degenerated discs in the mid-sagittal horizontal section of the disc was compared with the experimental results of Adams et al. (1996, 2002), obtained by stress profilometry. In axial tension, the calculated disc elongations were compared with the in vivo measured elongations of Kurutz (2006a) for healthy and degenerated segments.

10. Types of elements applied to the segmental structures

The cancellous core and the posterior bony elements of vertebrae can be modeled as *3D solid continuum elements*, as isoparametric 8-node hexahedral (brick) elements, or as 20- or 27-noded brick elements, moreover, as 10-noded tetrahedral elements. The cortical shell and the endplates can be modeled as *thin shell elements*, like 4-node shell elements. Quasi-rigid *beam elements* can connect the posterior vertebra with the medial transverse processes (pedicles) and from the medial transverse processes to the medial spinous process (lamina). Beam elements can also be used to represent the transverse and spinous processes. The bony surface of the facet joints can be represented by shell elements where beam elements link these facets to the lamina, simulating the inferior and superior articular processes (Little et al. 2008). The facet joints can be modeled as 3D 8-node surface-to-surface *contact elements* (Zhong et al., 2006).

The disc annulus ground substance is generally modeled as 3D continuum elements. The collagen fibers can be modeled as *truss elements* or as *reinforced bar (rebar) type elements* embedded in 3D solid elements (Lodygowski et al. 2005). The nucleus pulposus can be modeled as *hydrostatic fluid volume elements*.

The anterior and posterior longitudinal ligaments can be modeled as thin shell elements, or, the ligaments can be modeled as 2-node axial elements, that is, *tension only linear or nonlinear truss or cable or spring elements*.

11. Example: FE numerical simulation of age-related degeneration of lumbar segments

Spinal aging degeneration processes are the most typical example of those phenomena that can not be clarified in their progress by experimental methods, but exclusively by numerical simulation. 3D FE simulation of age-related degeneration processes of lumbar segments L3-S1 was investigated in axial compression. Aging degeneration of the segment was modeled by the material properties of its components, validated both for compression and tension, by comparing the numerical results with experimental data. Five grades of aging degeneration were distinguished from the healthy to fully degenerated case.

A 3D geometrical model of a typical lumbar FSU was created, obtained by using Pro/Engineer code. The geometrical data of the FSU were obtained by the measures of a typical lumbar segment. Cortical and cancellous bones of vertebrae were separately modeled, including posterior bony elements, too. The thickness of vertebral cortical walls and endplates were 0.35 and 0.5 mm, respectively. The height of the disc was 10 mm. Annulus fibrosus consisted of ground substance and elastic fibers. Annulus matrix was divided to internal and external ring; with three layers of annulus fibers of 0.1 mm² cross section. The geometry and orientation of facet joints were chosen according to Panjabi et al., (1993).

The FE mesh was generated by ANSYS Workbench, the connections between several geometrical components were integrated to the FE model by ANSYS Classic. The FE model consisted of solid, shell and bar elements. Annulus matrix, nucleus, cancellous bone, articular joints and different types of attachments were modeled by Solid_186/187 elements with quadratic displacement behavior. Cortical shells and endplates were modeled by Shell_181 elements with four nodes at each element. All ligaments were modeled by Shell_41 elements, with tension-only material. Annulus fibers were mapped into Link_10 bar elements with bilinear stiffness matrix resulting in a uniaxial tension-only behaviour.

By means of a systematic numerical analysis of the separated effect of the two main mechanical components of aging degeneration, it was proved that at the beginning period of the aging process, the effect of the loss of hydrostatic stress state of the nucleus had the dominant effect, while in further aging, the stiffening of nucleus dominated. This fact leads to the largest deformability and the smallest compressive stiffness, consequently, to the risk of segmental instability at mildly degenerated case in young age, while the stiffness and stability increased with further aging and degeneration. Exclusively by FE numerical analysis the observation and question of many international papers, why low back pain problems insult so frequently the young adults, could be answered.

For the introduced numerical model, the mean axial compressive stiffness of the nucleus, internal and external annulus was about 700, 1200, 500 N/mm for healthy; 500, 1000, 400 N/mm for mildly; 800, 1200, 600 N/mm for medium; 2100, 2000, 1100 N/mm for severely and 5500, 4500, 3000 N/mm for fully degenerated discs, respectively. The stiffness of the whole disc was about 2400, 1900, 2600, 5200 and 13000 N/mm for the five degeneration grades. While vertical intradiscal stresses showed significant change during aging degeneration, between 0.6-1.6 MPa, the horizontal stresses remained quasi constant and small, between 0.2-0.5 MPa for 1000 N compression. In the numerical modeling of hydrostatic state of a healthy nucleus, smaller than 1 MPa Young's modulus of nucleus must be considered to cut down the nuclear stress divergence below 10%. For healthy nucleus, $E=0.1$ MPa seemed to be acceptable.

FE simulations of degeneration processes of lumbar segments may help clinicians to understand the initiation and progression of disc degeneration and to treat lumbar discopathy problems even more effectively.

12. References

- Acaroglu, E.R., Iatridis, J.C., Setton, L.A., Foster, R.J., Mow, V.C., Weidenbaum, M. (1995). Degeneration and aging affect the tensile behaviour of human lumbar annulus fibrosus, *Spine*, 20(24), 2690-2701.
- Argoubi, M., Shirazi-Adl, A. (1996). Poroelastic creep response analysis of a lumbar motion segment in compression, *Journal of Biomechanics*, 29(10), 1331-1339.
- Adams, M.A., Dolan, P. (1991). A technique for quantifying the bending moment acting on the lumbar spine in vivo, *Journal of Biomechanics*, 24(2), 117-126.
- Adams, M.A., Freeman, B.J., Morrison, H.P., Nelson, I.W., Dolan, P. (2000). Mechanical initiation of intervertebral disc degeneration, *Spine*, 25(13), 1625-1636.
- Adams, M.A., Bogduk, N., Burton, K., Dolan, P. (2002). *The Biomechanics of Back Pain*, Churchill Livingstone, Edinburgh, London, New York.
- Adams, M.A., Hutton, W.C. (1983). The effect of posture on the fluid content of lumbar intervertebral discs, *Spine*, 8(6), 665-671.
- Adams, M.A., McNally, D.S., Dolan, P. (1996). Stress distributions inside intervertebral discs. The effects of age and degeneration. *J. Bone Joint Surg. Br.* 78(6), 965-972.
- Bader, D.L., Bouten, C. (2000). Biomechanics of soft tissues. In: Dvir, Z. (Ed.), *Clinical Biomechanics*, Churchill Livingstone, New York, Edinburgh, London, Philadelphia, pp. 35-64.
- Banse, X., Devogelaer, J.P., Munting, E., Delloye, C., Cornu, O., Grynpas, M. (2001). Inhomogeneity of human vertebral cancellous bone: systematic density and structure patterns inside the vertebral body. *Bone*, 28(5), 563-571.
- Baroud, G., Nemes, J., Heini, P., Steffen, T. (2003). Load shift of the intervertebral disc after a vertebroplasty: a finite element study, *European Spine Journal*, 12(4), 421-426.
- Benzel, E.C.: *Biomechanics of Spine Stabilization*, (2001). American Association of Neurological Surgeons, Rolling Meadows, Illinois.
- Bogduk, N. (2005), *Clinical Anatomy of the Lumbar Spine and Sacrum*, Elsevier, Churchill Livingstone, Edinburgh, London, New York, Oxford, Philadelphia, St Luis, Sidney, Toronto.
- Bogduk, N., Twomey, L.T. (1987). *Clinical Anatomy of the Lumbar Spine*, Churchill Livingstone, New York.
- Briggs, A.M., Greig, A.M., Wark, J.D., Fazzalari, N.L., Bennell, K.L. (2004). A review of anatomical and mechanical factors affecting vertebral integrity. *International Journal of Medical Sciences*, 1, 170-180.
- Cassidy, J.J., Hiltner, A., Baer, E. (1989). Hierarchical structure of the intervertebral disc, *Connective Tissue Research*, 23(1), 75-88.
- Cassinelli, E., Kang, J.D. (2000). Current understanding of lumbar disc degeneration, *Operative techniques in orthopaedics*, 10(4), 254-262.
- Chen, C.S., Cheng, C.K., Liu, C.L., Lo, W.H. (2001). Stress analysis of the disc adjacent to interbody fusion in lumbar spine, *Medical Engineering and Physics*, 23(7), 483-491.

- Chen, S.H., Tai, C.L., Lin, C.Y., Hsieh, P.H., Chen, W.P. (2008). Biomechanical comparison of a new stand-alone anterior lumbar interbody fusion cage with established fixation techniques – a three-dimensional finite element analysis, *BMC Musculoskeletal Disorders*, 9(88), 1-10.
- Denoziere, G., Ku, D.N. (2006). Biomechanical comparison between fusion of two vertebrae and implantation of an artificial intervertebral disc, *Journal of Biomechanics*, 39(4), 766-775.
- Dolan, P., Mannion, A.F., Adams, M.A. (1994). Passive tissues help the back muscles to generate extensor moments during lifting, *Journal of Biomechanics*, 27(8), 1077-1087.
- Dolan, P., Adams, M.A. (2001). Recent advances in lumbar spinal mechanics and their significance for modelling, *Clin. Biomech.*, 16(Suppl.), S8-S16, 2001.
- Duan, Y., Seeman, E., Turner, C.H. (2001). The biomechanical basis of vertebral body fragility in men and women, *Journal of Bone and Mineral Research*, 16(12), 2276-2283.
- Dvir, Z. (2000). *Clinical Biomechanics*, Churchill Livingstone, New York, Edinburgh, London, Philadelphia.
- Eberlein, R., Holzapfel, G.A., Fröhlich, M. (2004). Multi-segment FEA of the human lumbar spine including the heterogeneity of the annulus fibrosus, *Computational Mechanics*, 34(2), 147-163.
- Elliott, D.M., Setton, L.A. (2001). Anisotrope and inhomogeneous tensile behaviour of the human annulus fibrosus: experimental measurements and material model predictions, *Journal of Biomechanical Engineering*, 123, 256-263.
- Fagan, M.J., Julian, S., Siddall, D.J., Mohsen, A.M. (2002). Patient specific spine models. Part 1: finite element analysis of the lumbar intervertebral disc – a material sensitivity study, *Proceedings of the Institution of Mechanical Engineers*, Part H: 216(5), 299-314.
- Ferguson, S.J., Steffen, T. (2003). Biomechanics of the aging spine, *European Spine Journal*, Suppl 2, S97-S103.
- Ferguson, J., Ito, K., Nolte, L.P. (2004). Fluid flow and convective transport of solutes within the intervertebral disc, *Journal of Biomechanics*, 37(2), 213-221.
- Goel, V.K., Kong, W., Han, J.S., Weinstein, J.N., Gilbertson, L.G. (1993). A combined finite element and optimization investigation of lumbar spine mechanics with and without muscles. *Spine*, 18, 1531-1541.
- Goel, V.K., Monroe, B.T., Gilbertson, L.G., Brinckmann, P. (1995a). Interlaminar shear stresses and laminae separation in the disc. Finite element analysis of the L3-L4 motion segment subjected to axial compressive loads. *Spine*, 20(6), 689-698.
- Goel, V.K., Gilbertson, L.G. (1995b). Applications of the finite element method to thoracolumbar spinal research – past, present and future, *Spine*, 20(15), 1719-1727.
- Gong, H., Zhang, M., Yeung, H.Y., Qin, L. (2005). Regional variations in microstructural properties of vertebral trabeculae with aging. *Journal of Bone and Mineral Metabolism*, 23(2), 174-180.
- Goto, K., Tajima, N., Chosa, E., Totoribe, K., Kuroki, H., Arizumi, Y. (2002). Mechanical analysis of the lumbar vertebrae in a three-dimensional finite element method in which intradiscal pressure in the nucleus pulposus was used to establish the model, *Journal of Orthopedic Science*, 7(2), 243-246.
- Gu, W.Y., Mao, X.G., Foster, R.J., et al., (1999). The anisotropic hydraulic permeability of human lumbar annulus fibrosus – influence of age, degeneration, direction and water content, *Spine*, 24(23), 2449-2455.

- Iatridis, J.C., Setton, L.A., Wedenbaum, M., Mow, V.C. (1997). Alterations in the mechanical behavior of the human lumbar nucleus pulposus with degeneration and aging. *Journal of Orthopaedic Research*, 15(2), 318-322.
- Iatridis, J.C., Setton, L.A., Foster, R.J., Rawlins, B.A., Weidenbaum, M., Mow, V.C. (1998). Degeneration affects the anisotropic and nonlinear behaviours of human annulus fibrosus in compression, *Journal of Biomechanics*, 31(6), 535-544.
- Johannessen, W., Elliott, D.M. (2005). Effects of degeneration on the biphasic material properties of human nucleus pulposus in confined compression, *Spine*, 30(24), E724-E729.
- Keaveny, T.M., Yeh, O.C. (2002). Architecture and trabecular bone - toward an improved understanding of the biomechanical effects of age, sex and osteoporosis. *Journal of Musculoskeletal and Neuronal Interactions*, 2(3), 205-208.
- Kim, Y.E., Goel, V.K., Weinstein, J.N., Lim, T.H. (1991). Effect of disc degeneration at one level on the adjacent level in axial mode, *Spine*, 16(3), 331-335.
- Kumaresan, S., Yoganandan, N., Pintar, F.A. (1999). Finite element analysis of the cervical spine: a material property sensitivity study, *Clinical Biomechanics*, 14(1), 41-53.
- Kumaresan, S., Yoganandan, N., Pintar, F.A. et al. (2001). Contribution of disc degeneration to osteophyte formation in the cervical spine: a biomechanical investigation, *Journal of Orthopaedic Research*, 19(5), 977-984.
- Kurowski, P., Kubo, A. (1986). The relationship of degeneration of the intervertebral disc to mechanical loading conditions on lumbar vertebrae, *Spine*, 11(7), 726-731.
- Kurutz, M., (2006a). Age-sensitivity of time-related in vivo deformability of human lumbar motion segments and discs in pure centric tension, *Journal of Biomechanics*, 39(1), 147-157.
- Kurutz, M., (2006b). In vivo age- and sex-related creep of human lumbar motion segments and discs in pure centric tension, *Journal of Biomechanics*, 39(7), 1180-9.
- Kurutz, M., Bene É., Lovas, A., (2003). In vivo deformability of human lumbar spine segments in pure centric tension, measured during traction bath therapy, *Acta of Bioengineering and Biomechanics*, 5(1), 67-92.
- Kurutz, M., Oroszváry, L. (2010). Finite element analysis of weightbath hydrotraction treatment of degenerated lumbar spine segments in elastic phase, *Journal of Biomechanics*, 43(3), 433-441.
- Lavaste, F., Skalli, W., Robin, S., Roy-Camille, R., Mazel, C. (1992). Three-dimensional geometrical and mechanical modelling of lumbar spine, *Journal of Biomechanics*, 25(10), 1153-1164.
- Little, J.P., Percy, M.J., Adam, C.J. (2008). Are coupled rotations in the lumbar spine largely due to the osseo-ligamentous anatomy? - a modeling study. *Computer Methods in Biomechanics and Biomedical Engineering*, 11(1), 95-103.
- Lodygowski, T., Kakol, W., Wierszycki, M. (2005). Three-dimensional nonlinear finite element model of lumbar intervertebral disc, *Acta of Bioengineering and Biomechanics*, 7(2), 29-37.
- Lu, Y.M., Hutton, W.C., Gharpuray, V.M. (1996a). Can variation in Intervertebral disc height affect the mechanical function of the disc? *Spine*, 21(19), 2208-2217.
- Lu, Y.M., Hutton, W.C., Gharpuray, V.M. (1996b). Do bending, twisting and diurnal fluid changes in the disc affect the propensity to prolapse? A viscoelastic finite element model. *Spine*, 21(22), 2570-2579.

- Malandrino, A., Planell, J.A., Lacroix, D., (2009). Statistical factorial analysis on the poroelastic material properties sensitivity of the lumbar intervertebral disc under compression, flexion and axial rotation, *Journal of Biomechanics*, 42(3), 341-348.
- McGill, S.M. (200). Biomechanics of the thoracolumbar spine, In: Dvir, Z. (Ed.), *Clinical Biomechanics*, Churchill Livingstone, New York, Edinburgh, London, Philadelphia, pp. 103-139.
- McNally, D.S., Adams, M.A. (1992). Internal intervertebral disc mechanics as revealed by stress profilometry, *Spine*, 17(1), 66-73.
- Moramarc, V., Palomar, A.P., Pappalettere, C., Doblaré, M. (2010). An accurate validation of a computational model of human lumbosacral segment, *Journal of Biomechanics*, 43(2), 334-342.
- Mosekilde, L. (2000). Age-related changes in bone mass, structure, and strength effects of loading, *Zeitschrift für Rheumatologie*, 59(Suppl.1), 1-9.
- Nachemson, A. (1966). The load on lumbar discs in different positions of the body, *Clinical Orthopaedics and Related Research*, 45, 107-122.
- Nachemson, A.L. (1981). Disc pressure measurements, *Spine*, 6(1), 93-97.
- Nachemson, A.L., Morris, J.M. (1964). In vivo measurements of intradiscal pressure, *Journal of Bone and Joint Surgery Am.*, 46, 1077-1092.
- Natarajan, R.N., Williams, J.R., Andersson, G.B. (2004). Recent advances in analytical modeling of lumbar disc degeneration. *Spine*, 29(23), 2733-2741.
- Natarajan, R.N., Williams, J.R., Andersson, G.B., 2006. Modeling changes in intervertebral disc mechanics with degeneration, *Journal of Bone and Joint Surgery Am.* 88(4), Suppl. 2, 36-40.
- Noailly, J., Lacroix, D., Planell, J.A. (2005). Finite element study of a novel intervertebral disc substitute, *Spine*, 30(20), 2257-2264.
- Noailly, J., Wilke, H.J., Planell, J.A., Lacroix, D. (2007). How does the geometry affect the internal biomechanics of a lumbar spine bi-segment finite element model? Consequences on the validation process, *Journal of Biomechanics*, 40(11), 2780-2788.
- Panjabi, M.M., Oxland, T., Takata, K., Goel, V., Duranceau, J., Krag, M., (1993). Articular facets of the human spine, quantitative three dimensional anatomy, *Spine*, 18, 1298-1310.
- Pintar, F.A., Yoganandan, N., Myers, T. Elhagediab, A., Sances Jr, A. (1992). Biomechanical properties of human lumbar spine ligaments, *Journal of Biomechanics*, 25(11), 1351-1356.
- Rohlmann, A., Zander, T., Bergmann, G. (2006a). Spinal loads after osteoporotic vertebral fractures treated by vertebroplasty or kyphoplasty, *European Spine Journal*, 15(8), 1255-1264.
- Rohlmann, A., Bauer, L., Zander, T., Bergmann, G., Wilke, H.J., (2006b). Determination of trunk muscle forces for flexion and extension by using a validated finite element model of the lumbar spine and measured in vivo data, *Journal of Biomechanics*, 39(6), 981-989.
- Rohlmann, A., Zander, T., Schmidt, H., Wilke, H.J., Bergmann, G. (2006c). Analysis of the influence of disc degeneration on the mechanical behaviour of a lumbar motion segment using the finite element method, *Journal of Biomechanics*, 39(13), 2484-2490.

- Rohlmann, A., Burra, N.K., Zander, T., Bergmann, G. (2007). Comparison of the effect of bilateral posterior dynamic and rigid fixation devices on the loads in the lumbar spine: a finite element analysis, *European Spine Journal*, 16(8), 1223-1231.
- Ruberté, L.M., Natarajan, R.N., Andersson, G.B. (2009). Influence of single-level lumbar degenerative disc disease on the behavior of the adjacent segments--a finite element model study, *Journal of Biomechanics*, 42(3), 341-348.
- Sato, K., Kikuchi, S., Yonezawa, T. (1999). In vivo intradiscal pressure measurement in healthy individuals and in patients with ongoing back problems, *Spine*, 24(23), 2468-2474.
- Schmidt, H., Heuer, F., Wilke, H.J. (2009). Which axial and bending stiffnesses of posterior implants are required to design a flexible lumbar stabilization system? *Journal of Biomechanics*, 42(1), 48-54.
- Schroeder, Y., Wilson, W., Huyghe, J.M., Baaijens, P.T. (2006). Osmoviscoelastic finite element model of the intervertebral disc, *European Spine Journal*, 15(Suppl 3), 361-371.
- Sharma, M., Langrana, M.A., Rodriguez, J. (1995). Role of ligaments and facets in lumbar spinal stability, *Spine*, 20(8), 887-900.
- Shirazi-Adl, S.A., Shrivastava, S.C., Ahmed, A.M., (1984). Stress analysis of the lumbar disc-body unit in compression. A three-dimensional nonlinear finite element study, *Spine*, 9(2), 120-34.
- Shirazi-Adl, S.A., Ahmed, A.M., Shrivastava, S.C., (1986a). Mechanical response of a lumbar motion segment in axial torque alone and combined with compression, *Spine*, 11(9), 914-927.
- Shirazi-Adl, A., Ahmed, A.M., Shrivastava, S.C., (1986b). A finite element study of a lumbar motion segment subjected to pure sagittal plane moments, *Journal of Biomechanics*, 19(4), 331-350.
- Smit, T., Dodgaard, A., Schneider, E. (1997). Structure and function of vertebral trabecular bone, *Spine*, 22(24), 2823-2833.
- Umehara, S., Tadano, S., Abumi, K. (1996). Effects of degeneration on the elastic modulus distribution in the lumbar intervertebral disc, *Spine*, 21(7), 811-819.
- Wang, J.L., Parnianpour, M., Shirazi-Adl, A., Engin, A.E. (2000). Viscoelastic finite element analysis of a lumbar motion segment in combined compression and sagittal flexion, *Spine*, 25(3), 310-318.
- White, A. A., Panjabi, M. M. (1990). *Clinical Biomechanics of the Spine*, Lippincott Williams and Wilkins, Philadelphia, etc.
- Wilke, H.J., Neef, P., Caimi, M., Hoogland, T., Claes, L.E. (1999). New in vivo measurements of pressures in the intervertebral disc in daily life, *Spine*, 24(8):755-62.
- Wilke, H.J., Rohlmann, A., Neller, S., Graichen, F., Claes, L., Bergmann, G. (2003). ISSLS Prize Winner: A novel approach to determine trunk muscle forces during flexion and extension: a comparison of data from an in vitro experiment and in vivo measurements, *Spine*, 28(23), 2585-2593.
- Williams, J.R., Natarajan, R.N., Andersson, G.B.J. (2007). Inclusion of regional poroelastic material properties better predicts biomechanical behaviour of lumbar discs subjected to dynamic loading, *Journal of Biomechanics*, 40(9), 1981-1987.
- Zander, T., Rohlmann, A., Calisse, J., Bergmann, G., (2001). Estimation of muscle forces in the lumbar spine during upper body inclination, *Clinical Biomechanics*, 16(Suppl1), S73-S80.

- Zander, T., Rohlmann, A., Burra, N.K., Bergmann, G., (2006). Effect of a posterior dynamic implant adjacent to a rigid spinal fixator, *Clinical Biomechanics*, 21(8), 767-774.
- Zhang, Q.H., Zhou, Y.L., Petit, D., Teo, E.C., (2009). Evaluation of load transfer characteristics of a dynamic stabilization device on disc loading under compression, *Medical Engineering and Physics*, 31(5), 533-538.
- Zhong, Z.C., Wei, S.H., Wang, J.P., Feng, C.K., Chen, C.S., Yu, C.H. (2006). Finite element analysis of the lumbar spine with a new cage using a topology optimization method, *Medical Engineering and Physics*, 28(1), 90-98.

Analysis of human pressure ulcer and cushion pads for its prevention

Masataka Akimoto M.D.,Ph.D.
*Nippon Medical School
Japan*

1. Introduction

Pressure sore ulcers can be a serious problem for bedridden patients. The first indication of ulcer formation is redness on the skin surface. However, some practitioners have found evidence that ulcers form in deeper tissue and then spread toward the surface of the skin (Daniel et al., 1982; Koshimura et al., 2004). By the time surface damage is noticed, subcutaneous fat tissue necrosis has already occurred (Fig. 1). It sometimes manifests as an undermining formation (Fig. 2) and it tends to extend.

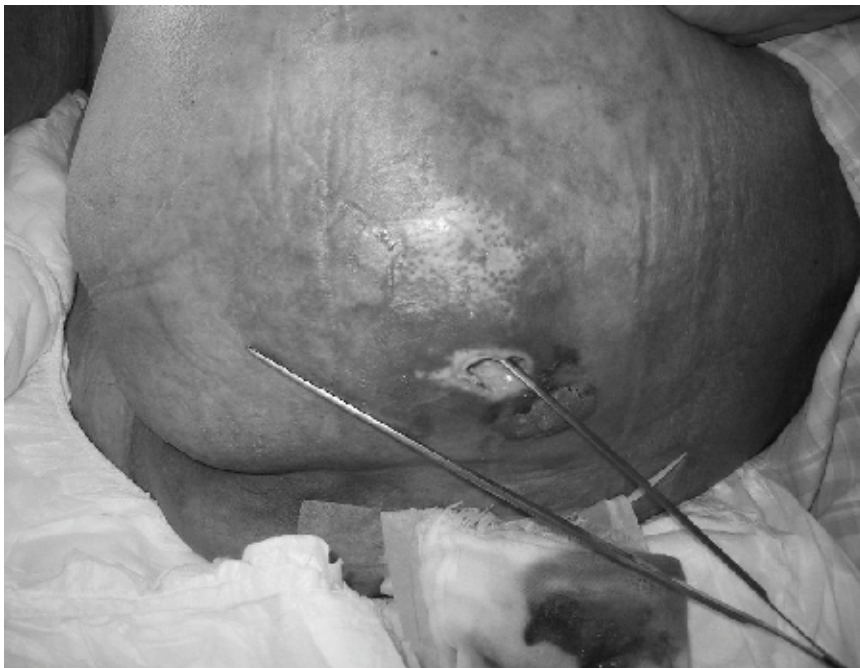


Fig. 1. Pressure ulcer with undermining: Pressure ulcer sometimes found with deeper tissue necrosis.

There have been several studies of finite element analysis (FEA) of pressure ulcers (Chow & Odell, 1978; Mak et al., 1994; Honma & Takahasi, 2001; Ragan et al., 2002). Todd and Thacker demonstrated the consistency of the finite element model in analysis of pressure ulcers (Todd & Thacker, 1994).

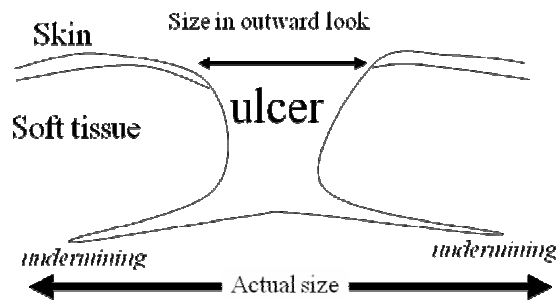


Fig. 2. Pressure ulcer with undermining: Cross sectional schematic of a pressure ulcer. Pressure ulcers are sometimes found with deeper tissue necrosis.

The author hypothesized that structural change (i.e. undermining formation) causes worsening of stress distribution of the wound. The finite element model can be used to characterize the extension mechanism of pressure sore with undermining.

In this chapter, author show some studies to analyze stress distribution that affect damage of soft tissue that is emergence of pressure ulcers. In the first study (Study 1), analysis in different sized undermining with vertical directional load was done.

In actual bedridden patients, load is not only gravity. They are pushed and rotated by nursing staffs. In the second study (study 2), the author added analyses for different directional load that will occur in actual situation.

Thick cushion pads are sometimes used for prevention of pressure ulcers. A thick cushion pad is useful for reduction of stress concentration caused by a vertical directional load. The author hypothesized that even if a cushion pad was vertically thin, it would effectively reduce oblique loads, which can increase stress concentration.

The third study was to analyze the stress distribution of a model of the human body attached to a thin cushion pad with a range of hardness, to assess the effectiveness of the cushion pad in reducing stress concentration.

The purpose of these studies were to describe the stress distribution of pressure ulcers in combination of various geometric and load conditions. The present findings support a hypothesis that after a small necrosis arises in a deep region of the body, it induces structural change that results in a structure that facilitates mechanical extension mechanically of the necrosis (Kuroda & Akimoto, 2005; Akimoto et al., 2007).

All data analysis was performed using a personal computer (Pentium 4: 2.4 GHz with 1 GB RAM) and ADINA analytical software (version 8.3, ADINA R&D, Inc., Massachusetts, U.S.A.)

2. Materials and methods (Study 1)

The first assumption of the model was that the shape of the human trunk is a cylinder. The second assumption was that the human body consists of 2 categories of tissue: soft tissue

and hard tissue. Soft tissue corresponds to structures such as skin, fat and muscle. Hard tissue corresponds to bone. To simplify the calculations, only the lower half of the cylinder model was used. Thus, the cross section of the basic model consists of 2 concentric semicircles. The outer semicircle functions as the soft tissue, and the inner semicircle functions as the hard tissue (Fig. 3).

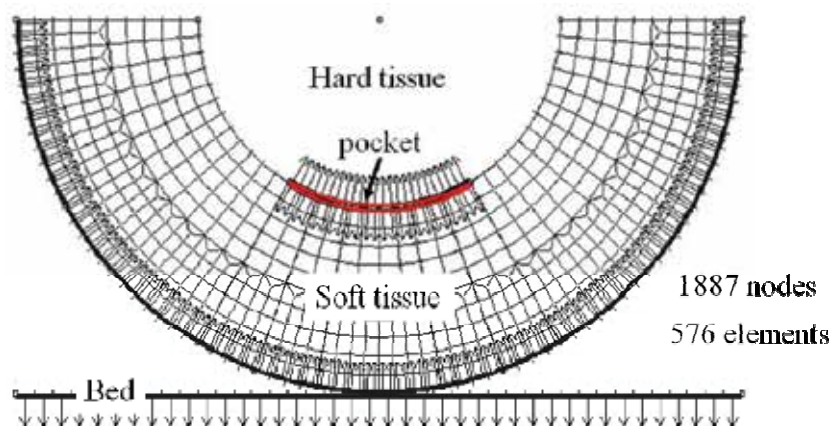


Fig. 3. Basic design of pressure sore model: A hollow half cylinder with a diameter of 20 cm was used to represent the soft tissue. A cylinder with a diameter of 10 cm was used to represent the hard tissue. Undermining was represented by a small gap at the junction of the soft and hard tissue. Model was meshed into 1887 nodes and 576 elements.

For FEA, the soft tissue and the cushion pad were combined into a mesh with a basic geometry of 1887 nodes and 576 elements.

Assuming the simplest clinical state, the following basic parameters of FEA were assigned values: geometry, material properties, loading condition and boundary condition. Analysis was performed in 2 dimensions.

Geometry:

A hollow half cylinder with a diameter of 20 cm was used to represent the soft tissue. A hollow half cylinder with a diameter of 10 cm was used to represent the hard tissue. Undermining was represented by a small gap at the junction of the soft and hard tissue. The upper and lower edges of the undermining were designed so that they comprised a contact pair with no friction. For evaluation, four different models were prepared:

- 1) no undermining (no gap);
- 2) small undermining (gap, 1.7 cm);
- 3) medium undermining (gap, 3.5 cm); and
- 4) large undermining (gap, 5.2 cm)

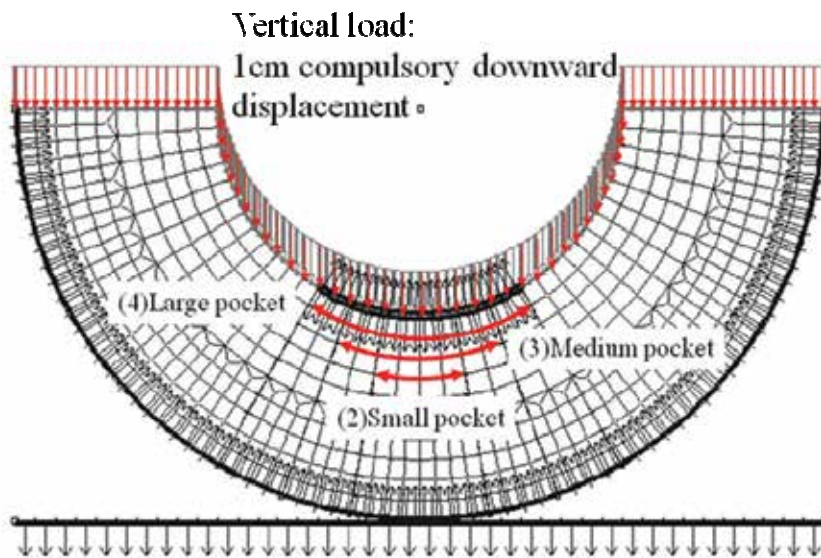


Fig. 4. Undermining size and load condition for study 1: Four different sized models were prepared: 1) No undermining (no gap); 2) Small (gap, 1.7 cm); 3) Medium (gap, 3.5 cm); and 4) Large undermining (gap, 5.2 cm). A 1-cm downward displacement on upper edges to represent gravity.

Material Properties:

Actual biological tissue is nonlinear, anisotropic and visco-elastic. To simplify the calculations, the specific microstructure of the tissue was not taken into account. It was assumed that the soft tissue was linear, isotropic, and time-independent. Young's module of the soft tissue was set to 15 kPa. Poisson's ratio of the soft tissue was set to 0.49.

Loading Conditions:

Vertically directed loading was included to represent gravity. This loading is expressed as a 1 cm downward displacement of the upper edges of the model.

Boundary Conditions:

The patient was assumed to be lying on a flat, hard, non-slip bed. For this purpose, a tangential line was drawn adjacent to the lower edge of the soft tissue. This line was fixed in all directions, and formed a contact pair with the edge of the soft tissue or the edge of the cushion pad. The coefficient of friction was set to 1.0 for these contact pairs.

3. Results (Study 1)

Results of the FEA were summarized and visualized using a von-Mises stress distribution map. On the stress distribution map of the model with no undermining, a large concentration of effective stress was observed at the centre just under the hard tissue, and also at the junction of soft and hard tissue. In both regions, the amount of effective stress was approximately 4.0 kPa (Fig. 5).

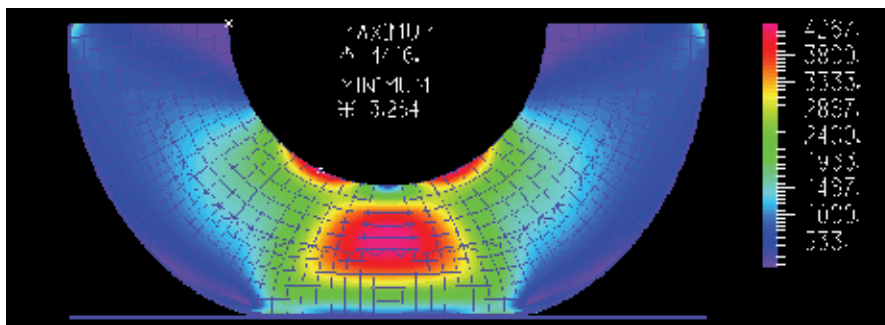


Fig. 5. Stress distribution of no-undermining model

The model with a small undermining had nearly the same pattern of stress distribution as the model with no undermining. Also, the maximum amount of effective stress was the same as that of the model with no undermining (Fig. 6).

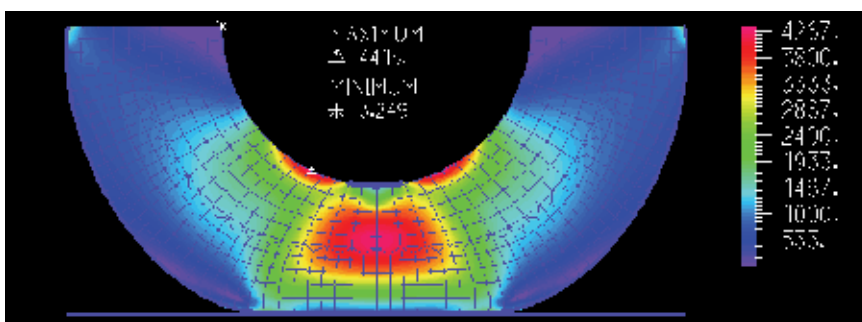


Fig. 6. Stress distribution of small undermining model

In the model with medium undermining, the points of maximum effective stress were at the edge of the undermining. The maximum amount of effective stress was 6168 Pa. Stress was also concentrated immediately under the hard tissue, with a local maximum amount of effective stress of about 4.0 k Pa (Fig. 7).

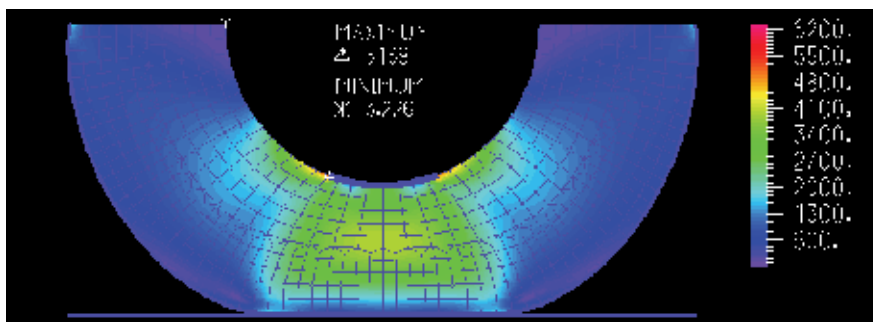


Fig. 7. Stress distribution of medium undermining model

In the model with large undermining, the points of maximum effective stress were at the edge of the undermining. The maximum amount of effective stress was 7.7 k Pa. Stress was

also concentrated under the hard tissue, with a local maximum amount of effective stress of about 4.0 k Pa (Fig. 8).

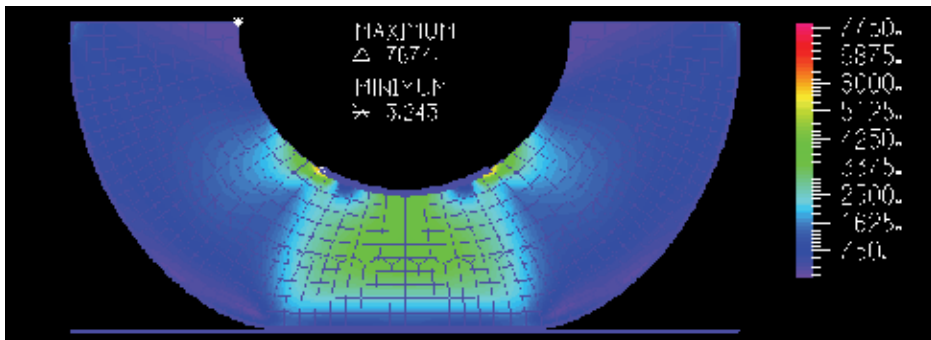


Fig. 8. stress distribution of large undermining model

A series of examinations revealed two main areas of stress concentration:

- 1) At the junction of hard and soft tissues, or the edge of the undermining
- 2) The centre of the soft tissue just under the hard tissue.

The maximum amount of effective stress increased with increasing size of undermining (Fig. 9).

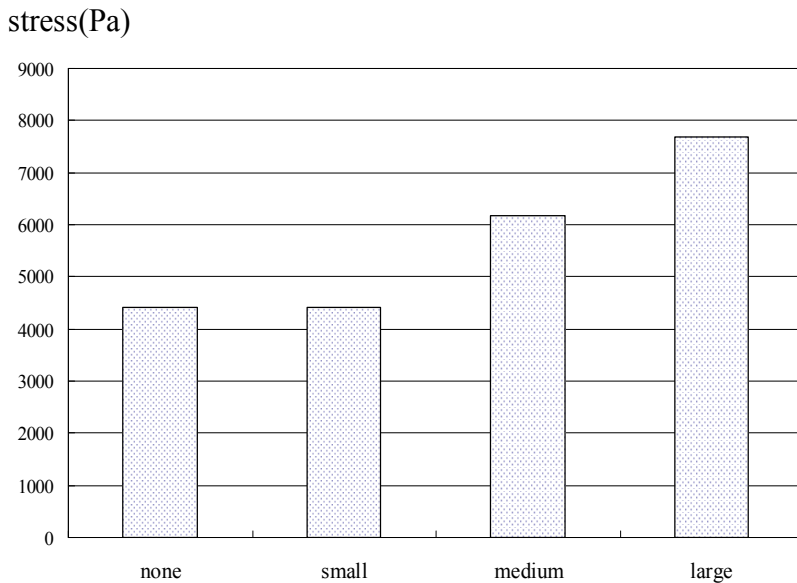


Fig. 9. Maximum effective stress (Study 1): Amount of maximum effective stress increased with increment of undermining size

4. Materials and methods (Study 2)

Basic geometry, material properties and boundary conditions were the same as those of the study 1.

Under actual clinical conditions, patients do not generally remain in the same position at all times. They are moved or rotated during nursing care. To simulate these conditions, the model was subjected to horizontal displacement. The model conditions included 1 cm of horizontal movement and 1 cm of vertical movement, representing horizontal and vertical displacement resulting from horizontal and vertical loads, respectively (Fig. 10).

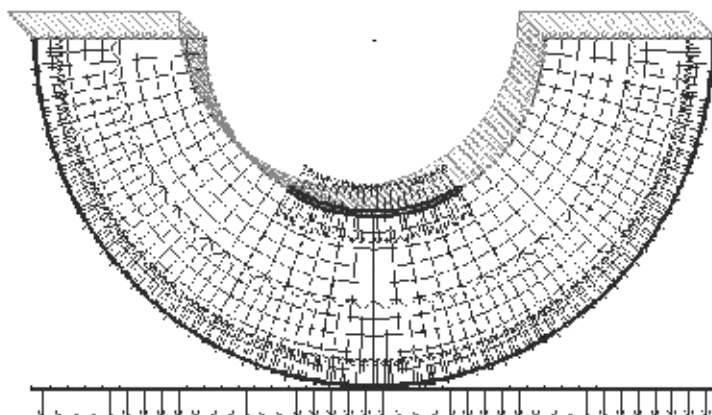


Fig. 10. Model and loading conditions for study 2

5. Results (Study 2)

Results of the FEA were summarized and visualized using a von-Mises stress distribution map. In a condition with additional 0.33 cm horizontal load, distribution of stress had changed. Larger stress concentration emerged at the edge of the undermining in each size of undermining. In a condition with additional 0.66 cm horizontal load, distribution of stress pattern showed more biased stress concentration. The amount of effective stress was increased. Further increment of stress was observed in each undermining patterns of 1.0 cm horizontal load model (Fig. 11-13).

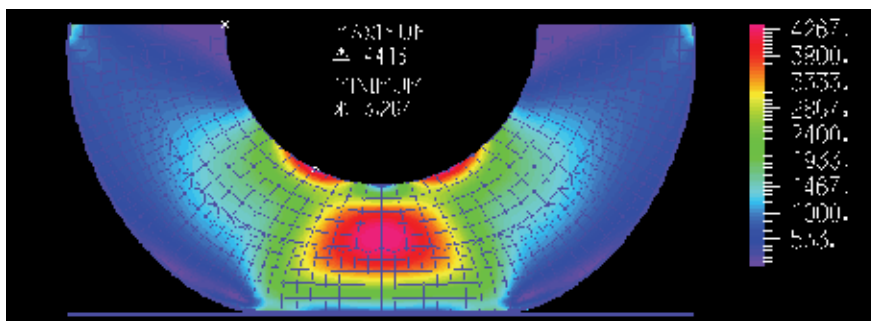


Fig. 11. Stress distribution in no gap model with a straight load

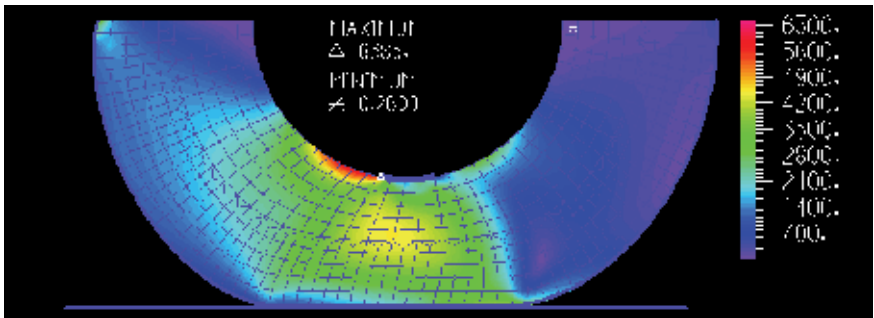


Fig. 12. Stress distribution in small gap model with an oblique load (1.0 cm horizontal + 1.0 cm vertical)

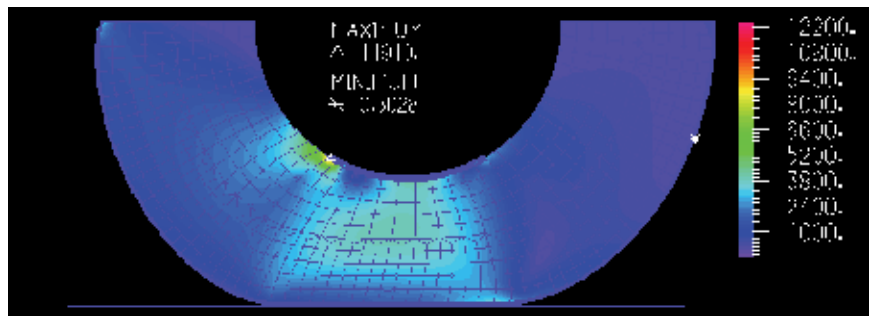


Fig. 13. Stress distribution in large gap model with oblique load (1.0 cm horizontal + 1.0 cm vertical)

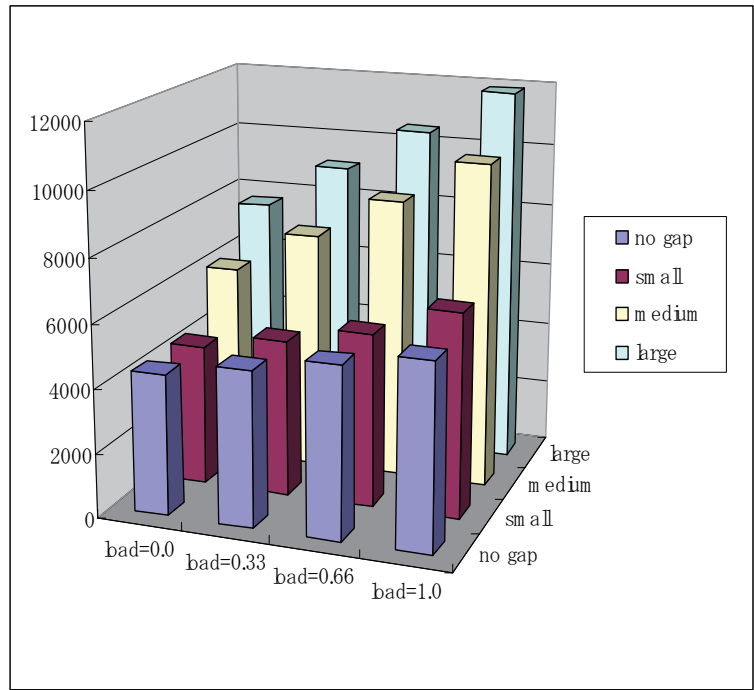


Fig. 14. Study 2: maximum stress in various horizontal load and gap size.

The results of the study 2 were summarized in a 3-D bar graph (Fig. 14). With the increment of horizontal load, biased stress concentration increased. As we noticed the gap size affect the stress distribution of pressure sore in the previous study, the horizontal load is also a big factor of worsening pressure ulcer.

6. Materials and methods (Study 3)

Basic design of FEA models were as the same as previous studies except cushion pads. For the study 3, the cushion pad was represented by a thin additional outer layer of soft tissue (Fig.15). It was assumed that the skin and the pad were in tight contact.

Poisson's ratio of the cushion pad was set to 0.49 (same as the soft tissue value). Five different values were used for Young's module of the cushion pad: 15 kPa (same as the soft tissue value), 7.5 kPa (1/2 the soft tissue value), 3.75 kPa (1/4 the soft tissue value), 1.87 kPa (1/8 the soft tissue value), and 0.93 kPa (1/16 the soft tissue value).

As a control, a model without a cushion pad was also analyzed. Effective stress was evaluated using a distribution map.

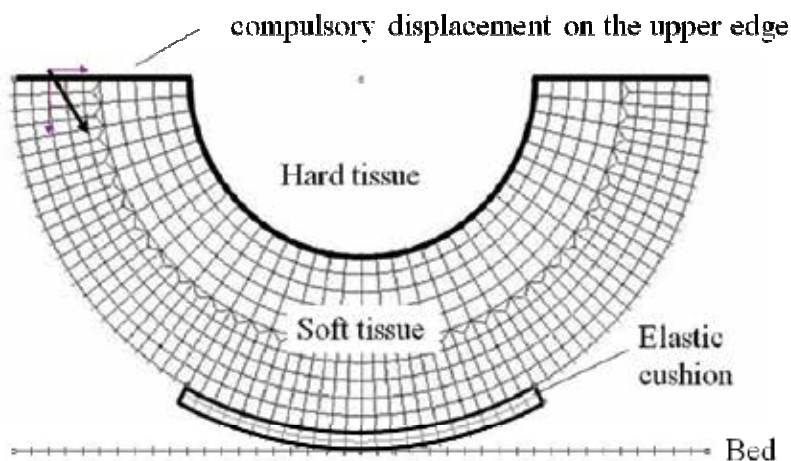


Fig. 15. Design of a model for study 3.

7. Results (Study 3)

Results of the FEA were summarized and visualized using a von-Mises stress distribution map. In the model without a cushion pad, 2 regions of concentration were observed. One region was at the boundary between the soft and hard tissue, and the other was at the centre of the soft tissue immediately below the hard tissue. The maximum values of effective stress were 5.83 kPa for the region at the boundary and 4.64 kPa for the region at the centre. These results are consistent with the previous findings.

All of the models with a cushion pad had a stress concentration pattern that was similar to that of the model without a cushion pad (Fig 16-18). That is, they each had 2 regions of concentration: one at the centre, and one at the boundary. The maximum values of effective stress decreased as Young's module of the cushion pad decreased (Fig 19). Thus, although

the stress concentration pattern was similar to that of the model without a cushion pad, the stress distribution became more diffuse as the cushion pad softness increased (i.e., Young's module decreased).

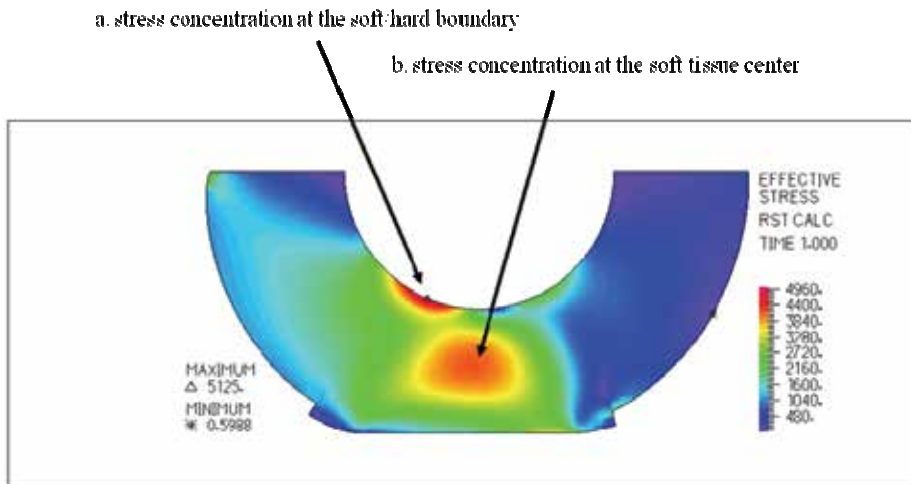


Fig. 16. Stress distribution with a same hardness cushion pad ($E=15000$ Pa)

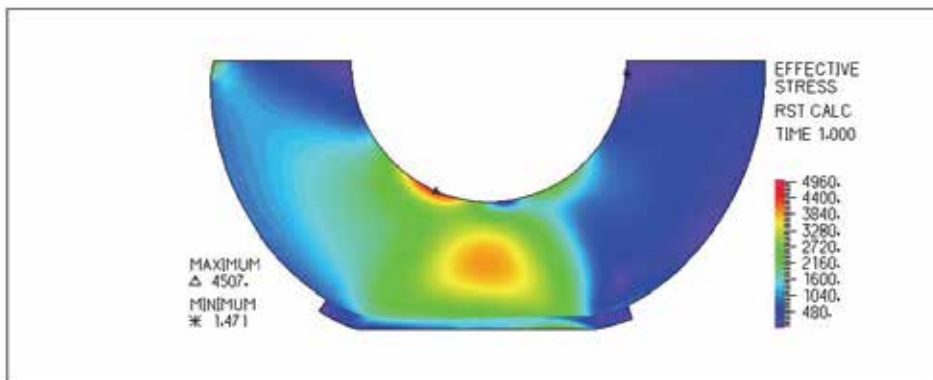


Fig. 17. Stress distribution with a quarter hardness cushion pad ($E=3750$ Pa)

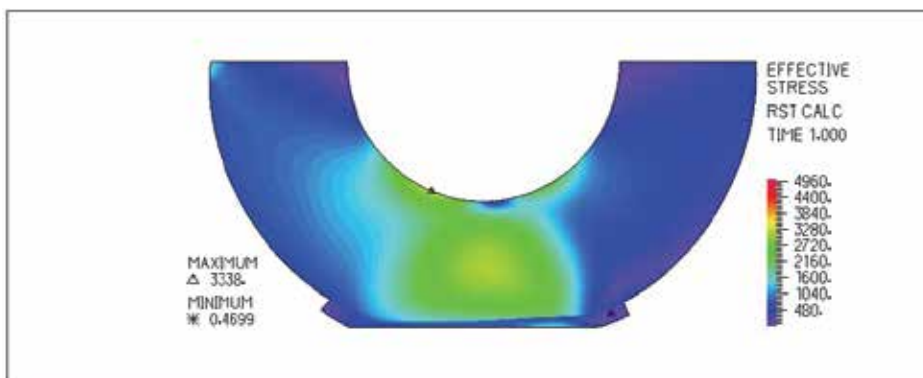


Fig. 18. Stress distribution with an 1/16 hardness cushion pad ($E=973$ Pa)

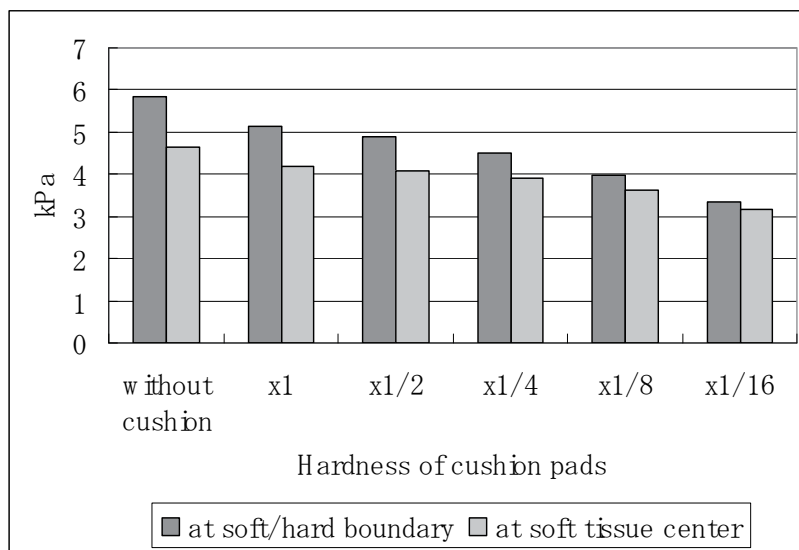


Fig. 19. Study 3: extreme maximum value of effective stress

8. Discussion

Chow and Odell made an axi-symmetric finite element model of a human buttock (Chow & Odell, 1978). The purpose of their study was to characterize stress patterns within the soft tissues of the buttock under different loading conditions. They modelled the buttock as a hemisphere of linear elastic isotropic soft tissue, with a rigid core to model the ischium. Honma and Takahashi evaluated the model of Chow and Odell, using the same conditions but with more precise calculation (Honma & Takahashi, 2001). Although that model was based on a hemisphere and our model was based on a cylinder, the basic design and shape of the 2 models are quite similar.

They mentioned stress concentration in the middle of the soft tissue and at the junction between the soft and hard tissue. In the present studies, we improved that model by converting it from axi-symmetric to non-symmetric, allowing analysis under non axi-symmetric conditions such as with an oblique load.

The results obtained from the first study using our non-undermining model were in good agreement with results obtained using the model of Chow and Odell. One of the purposes of the present analyses was to obtain data for use in further refinement of models of undermining. In the present models with gaps (representing undermining), stress was concentrated at the edges of the undermining. The maximum stress value increased with increasing size of undermining. These results suggest that after undermining develops, stress begins to concentrate at the edges of the undermining, and that this stress erodes the edges of the undermining, thus increasing the size of the undermining. Such a process is consistent with the clinical phenomena associated with formation of undermining.

Our second analyses were done for further modification for additional load from horizontal direction. Models with undermining showed increment of the stress concentration on the edges of the undermining in each combination of horizontal loads. Furthermore, shear stress

increased with horizontal load. In clinical condition, patient sometimes are moved and pushed in bed, so horizontal loads are tend to be exposed to such loads..

This result suggests a vicious circle of pressure sore that once some sized undermining emerges, stress concentration emergence on the edges of the undermining that will destroy the edges and increase the size of undermining by itself, and then more stress concentration occurs (Fig. 20). It may explain clinical phenomenon of formation and extension of undermining.

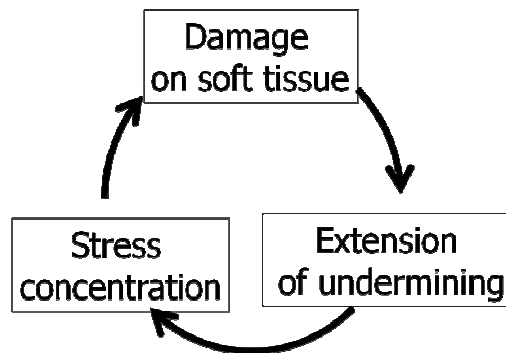


Fig. 20. A vicious circle of a pressure sore: Once some sized undermining emerges, stress concentration emerge on the edges of the undermining that will destroy the edges and increase the size of undermining by itself, then more stress concentration occurs.

A thick cushion pad is useful for reduction of stress concentration. There have been several studies (Souther et al., 1974; Garber & Krouskop, 1982; Minns et al., 1984; Delauter et al., 1984;) of the effectiveness of various cushions in reducing seat-interface pressures. The conclusion of those studies was that there seems to be a maximum effective cushion thickness, in terms of reduction of stress concentration. Those studies dealt mainly with vertical loads and cushion thickness.

Actual loads on patients in clinical settings are not limited to vertical loads due to gravity, but also include horizontal movement. A cushion that is thin in the vertical direction is not necessarily thin in the horizontal direction. In the third study, we simulated conditions in which a human is seated on a thin cushion pad with a variety of values of softness. The results suggest that a thin cushion pad reduces stress concentration, thus supporting the hypothesis. They also indicate that softer material is more effective at reducing stress concentration.

9. Conclusion

The finite element method can be used to calculate relationships between displacements and pressures or stresses. Thus, it can be used to determine the effects of various positions and motions of a patient on pressure sores, and this information can be used to prevent expansion of pressure sores. In the present study, although a very simple model was used for analysis, the results strongly suggest a particular mechanism for formation of large undermining in pressure ulcers.

10. References

- Akimoto M.; Oka T.; Oki K & Hyakusoku H.(2007).Finite element analysis of effect of softness of cushion pads on stress concentration due to an oblique load on pressure sores,*J Nippon Med Sch*,74(3),230-5,1345-4676
- Chow, W.W. & Odell, E.I. (1978). Deformations and stress in soft body tissues of sitting person, *J Biomech Eng*,100.,79-87, 0148-0731
- Daniel,R.K.; Priest,D.L. & Wheatley, D.C. (1981). Etiologic factors in pressure sores: an experimental model, *Arch Phys Med Rehabil*, 62.,492-8, 0003-9993
- Delauter, B.; Borni, R.; Hongladarom, T. & Giaconi, R. (1984). Wheelchair cushions designed to prevent pressure sores: an evaluation, *Arch. Phys Med Rehabil*, 57.,579-83, 0003-9993
- Garber, S. & Krouskop, T. (1982). Body build and its relationship to pressure distribution in the seated wheelchair patient, *Arch Phys Med Rehabil*, 63.,17-20, 0003-9993
- Honma, T. & Takahashi, M. (2001). Stress Analysis on the Sacral Model for Pressure Ulcers, *Jpn J Pressure Ulcer*, 3.,20-26, 1345-0417
- Koshimura, J.; Konya, C.; Sanada, H.; Nakatani, T.; Sugama, J.; Yajima, H. & Tabata, K. (2004). The process of undermining formation in the pressue ulcers, *Jpn J Pressure Ulcer*, 6., 607-615, 1345-0417
- Kuroda, S. & Akimoto, M.(2005).Finite Element Analysis of Undermining of Pressure Ulcer with a Simple Cylinder Model,*J Nippon Med Sch*,72.,174-178,1345-4676
- Mak, A.F.; Huang, L. & Wang, Q. (1994). A biphasic poroelastic analysis of the flow dependent subcutaneous tissue pressure and compaction due to epidermal loadings: issues in pressure sore, *J Biomech Eng*, 116.,421-9, 0148-073
- Minns, R., Sutton, R., Duffus, A. & Mittington, R. (1984). Underseat pressure distribution in the sitting spinal cord patient, *Paraplegia*, 22.,297-304, 0031-1758
- Ragan, R.; Kernozek, T. W.; Bidar, M. & Matheson, J. W. (2002). Seat-interface pressures on various thicknesses of foam wheelchair cushions: a finite modeling approach, *Arch Phys Med Rehabil* , 83.,872-5, 0003-9993
- Souther, S., Carr, S. & Vistnes, L. (1974). Wheelchair cushions to reduce pressure under bony prominences, *Arch Phys Med Rehabil*, 55.,460-4, 0003-9993
- Todd, B. A. & Thacker, J. G. (1994). Three-dimensional computer model of the human buttocks in vivo, *J Rehabil Res Dev*, 31., 111-9, 0748-7711

Microfinite Element Modelling for Evaluating Polymer Scaffolds Architecture and their Mechanical Properties from Microcomputed Tomography

Angel Alberich-Bayarri¹, Manuel Salmerón Sánchez²
M. Ángeles Pérez³ and David Moratal²

¹ *Department of Radiology, Hospital Quirón Valencia, Valencia, Spain.*

² *Center for Biomaterials and Tissue Engineering, Universitat Politècnica de València, Valencia, Spain.*

³ *Group of Structural Mechanics and Materials Modeling, Instituto de Investigación en Ingeniería de Aragón, Universidad de Zaragoza, Zaragoza, Spain.*

1. Introduction

With increasing time of life expectation, problems related to bone loss are one of the major causes of disability. Nowadays, the current standard therapy for treating bone atrophy consists on bone allografts and autografts (Rajan et al., 2006). Metal implants have also been used for the assessment of bone defects. These therapies present some drawbacks, allografts and autografts are limited by availability of material, donor site morbidity and also the possibility of disease transmission or immune rejection in the case of allografts (Geffre et al., 2009). In the case of metal implants, a high availability exists and also the risk of transmittable diseases from donor to host is eliminated. However, additional surgeries can be needed due to complications related with stress shielding, infection or implant failure.

In the last decades, there has been a growing interest in polymer scaffolds which encourage bone regeneration to treat bone defects (Parson, 1985). These scaffolds can be prepared prior to or during surgery, can be modified to alter mechanical strength and resorption rate, can be created in custom shapes unique to each defect site, and can be produced with highly controlled structures (Geffre et al., 2009).

Polymer scaffolds of different architectures are commonly synthesized with the main objective of obtaining effective functional biological responses. Such structures are designed to behave as an extracellular matrix where cells organize into a three dimensional architecture, stimulating the growth of new tissue (Langer & Vacanti, 1993; Huttmacher 2000; Freed et al., 1998; Yang et al., 2001).

Tissue engineering techniques are focused on the design of scaffolds that match biological and physio-chemical properties of the tissue where they are settled. In the case of bone, the bone is formed by osteoblasts that migrate from the adjacent original bone and marrow

cavities, a mechanism known as osteoconduction. Although this osteointegration is mandatory, there are situations (e.g., large defects) where the scaffolds not only are designed to provide the template for tissue regeneration but also need to be osteoinductive, that is, stimulate the migration of undifferentiated cells and induce their differentiation into active osteoblasts in order to boost de young bone formation (Oliveira et al., 2009). In order to improve the growth stimulating properties of grafting materials, they have been combined with growth factors and different cells types with variable results depending on the host regenerative capability (Dupraz et al., 1998; Gauthier et al., 2003).

Several methods have been developed for 3D polymer scaffold synthesis with different resultant structure properties regarding network topology, pores shape and density, determining their biological and mechanical functional response (Gao et al., 2003; Horbett et al., 1985; Zhang & Ma, 1999). The different methods used for scaffolds synthesis define the final 3D architecture. The most common techniques for fabric preparation are porogen leaching (Mikos et al., 1993; Zhou et al., 2005), fibre templates (Thomson et al., 1995), phase separation (Gao et al., 2003), emulsion freeze-drying (Whang et al., 1995), gas foaming (Harris et al., 1993), and solid free fabrication (SFF)-rapid prototyping (RP) techniques (Moroni et al., 2006). However, the control of these synthesized scaffold characteristics in the fabrication process is still the hobbyhorse of this science (Freed et al., 1998; Yang et al., 2001; Hajiali et al., 2010).

The architecture of scaffolds can be evaluated after the fabrication process by different ways. Although there are several techniques for scaffold analysis, a reasonable classification is based on the destructive or non-destructive nature of the methodology. Destructive methods are those in which the sample under analysis is deteriorated and is no longer useful to perform more measurements. The compressive stress-strain measurements using a load cell are performed experimentally and are an example of destructive methodology. An ideal way to evaluate the scaffold 3D structural design after the fabrication process would consist on a non-destructive, non-invasive and quantitative technique. In this sense, the use of micro-computed tomography (μ CT) scanners have allowed for the non-destructive and non-invasive examination of scaffolds at a high spatial resolution (Ho & Hutmacher, 2006). The use of μ CT has also eased the quality control of scaffold fabrication processes, the study of scaffold degradation kinetics and the assessment of bone tissue response (van Lenthe et al., 2007). Even more, the recent application of advanced image processing algorithms and simulation-based computational methods, like the Finite Element (FE) method to micro-computed tomography (μ CT) acquisitions, have allowed for the texture and virtual mechanical analysis of different scaffold 3D architectures (Alberich-Bayarri et al., 2009).

The combination of the FE method with μ CT acquisitions (also called μ FE) and image processing techniques as an alternative instrument for the mechanical properties evaluation of synthetic scaffolds is very recommendable. The geometrical fidelity of the corresponding structure to be analyzed by FE is a very important issue. This geometrical modelling can be obtained with high detail and reliability from μ CT scans. The combination of μ CT reconstructions and FE analysis for scaffolds design has been widely investigated in the literature. In this sense, a relevant study characterized porous phosphate glass and macroporous calcium phosphate bone cement, initially using μ CT to investigate the porosity and then through the μ FE method for the stress-strain analysis and influence of the cells attached to the material (Lacroix et al., 2006). A published work also described a μ CT based FE modelling of native trabecular bone and bone scaffolds (Jaecques et al., 2004).

In this chapter, the process for a non-invasive analysis of synthetic polymer scaffolds using the μ CT acquisition technique in combination with the μ FE method is extensively described. Some results and examples of the application of the methodology to different scaffold architectures are exposed. Finally, the elasticity modulus results obtained in different scaffold architecture are compared to the obtained experimentally in the corresponding stress-strain essays.

2. Image acquisition and processing

Polymer scaffold characteristics can be obtained after proper image processing from micro-computed tomography images (μ CT). The μ CT technique serves to analyze the inner 3D architecture of an object by image analysis. The sample is irradiated with X-rays in a series of 2D slices. The radiation attenuates while crossing the slice and arrives at the detectors with a reduced energy. Then, the corresponding attenuation coefficients can be calculated. These coefficients are directly related to the density of the materials forming the sample. An image is finally calculated with pixel intensities scaled by the densities of the materials forming the sample (Ho & Hutmacher, 2006).

The μ CT is a scanning system that is much higher in resolution than conventional clinical scanners. Clinical tomographic scanners may have resolutions on the order of half a millimetre or less. However, μ CT scanners permit the acquisition at very high spatial resolutions. In concrete, actual μ CT scanners allow spatial resolutions of a few micrometers down to even nanometers if the proper technology is available. With the significant increase in μ CT spatial resolution, there is also an associated decrease in imaging field of view. This is the reason that explains the limitation and use of these scanners to research areas for the analysis of biopsied biological tissue, different materials or small samples.

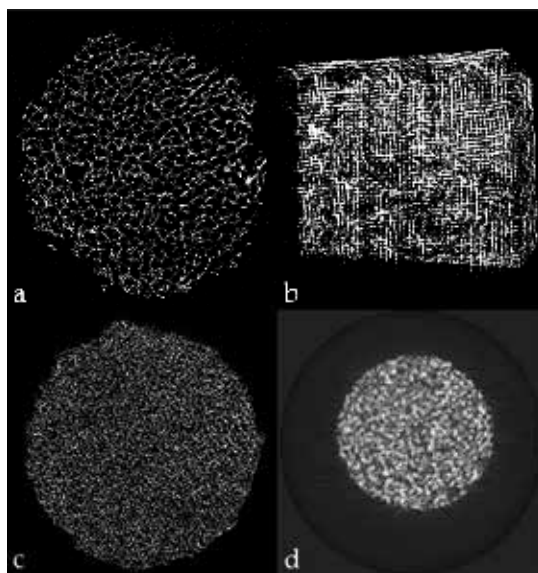


Fig. 1. Images acquired using μ CT. Polymer topology differences may be appreciated. Spherical pores structure was synthesized in scaffolds a), c) and d) while an orthogonal pore mesh is shown in b).

In the process of scaffold analysis, the first step is the μ CT acquisition. In order to explain the methodology of processing and analysis with a practical case, scaffolds of different synthesized topologies (spherical pores, cylindrical orthogonal pore mesh, salt particles) were scanned using a μ CT system (SkyScan, Kontich, Belgium). A set of 2D slices covering the entire sample were obtained. A very high isotropic spatial resolution of $7\mu\text{m}$ was achieved in acquisition. Representative slices of the μ CT images acquired from different scaffolds are shown in figure 1.

Some pre-processing is required after acquisition. First of all, volumes of interest (VOI) are extracted semi-automatically from each dataset. The process consists on initially defining a squared region of interest (ROI) placed in one of the slices of the dataset, the ROI must be inscribed in the circle defined by the cylindrical section of the scaffold (as it can be observed in figure 2-a). Then, the ROI is propagated through the rest of the slices, verifying that the selected volume exclusively contains inner structural data and not surrounding air. A new dataset is created with the segmented region of each slice (figure 2-b).

After segmentation, structure voxels must be differentiated from pore voxels, that is, images must be converted to its binarized form. The thresholding is implemented in a slice by slice basis by means of the Otsu's method (figure 2-c) (Otsu, 1979). As a result, a three-dimensional (3D) binary matrix representing the segmented volume of polymer under study is obtained.

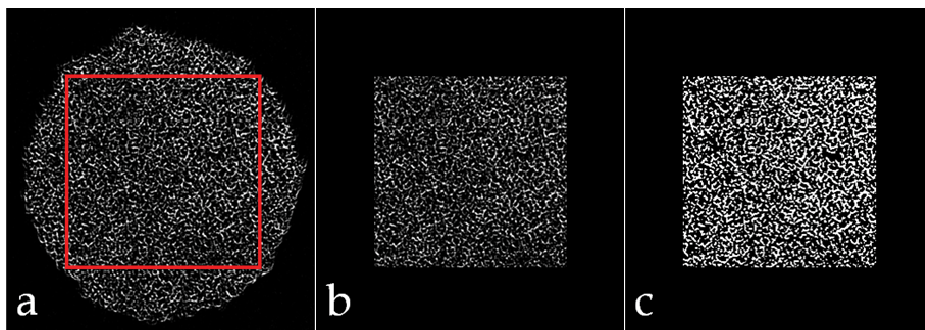


Fig. 2. In a), example of the definition of the inscribed ROI to be propagated through the rest of the slices for the VOI segmentation. In b), new dataset creation with the segmented ROI's for each slice. In c), binary image resulting from the thresholding process applied to the segmented dataset.

3. Volumetric reconstruction and meshing

Once the region under analysis has been binarized, the resulting 3D logical matrix can be visualized by a volumetric reconstruction process. An initial smoothing is performed to the volume in a 3D routine and then, the marching cubes algorithm (Lorensen & Cline, 1987) is applied in order to obtain a volumetric reconstruction of the scaffold structure, as it can be seen in figure 3 for the different scaffold topology.

In order to build the μ FE mesh from the volumetric data of the 3D binary matrix, the voxels corresponding to polymer must be converted to very small structural elements with certain coordinates and dimensions. There are different types of elements for FE volumetric modelling, like tetrahedrons or hexahedrons. In our case, the element used to form the mesh

is an eight-noded hexahedron, also called 'brick' element. However, the meshing process is not trivial and it supposes a high computational cost. To convert the volumetric reconstructions to a μ FE mesh, a fast voxel mesher specially designed for this application is applied (Alberich-Bayarri et al., 2007). The meshed scaffolds have a mean number of 3000000 nodes and 1400000 elements, approximately.

After both nodes and elements lists are defined, they must be converted to a format that can be interpreted by common FE simulation software, like ANSYS (Ansys Inc., Southpointe, PA, USA) or ABAQUS (Simulia, Providence, RI, USA). To do that, a specially designed routine translates the raw data into standardized *.ans, for ANSYS or *.inp for ABAQUS FE analysis software.

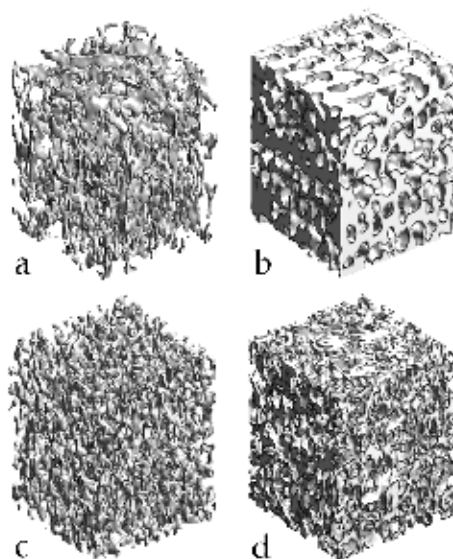


Fig. 3. 3D reconstructions of the different scaffolds after segmentation and thresholding. Same topologies distribution than in figure 1. Spherical pores in a), c) and d); orthogonal pore mesh in b).

Finally, the mesh can be loaded in the corresponding FE analysis platform (figure 4).

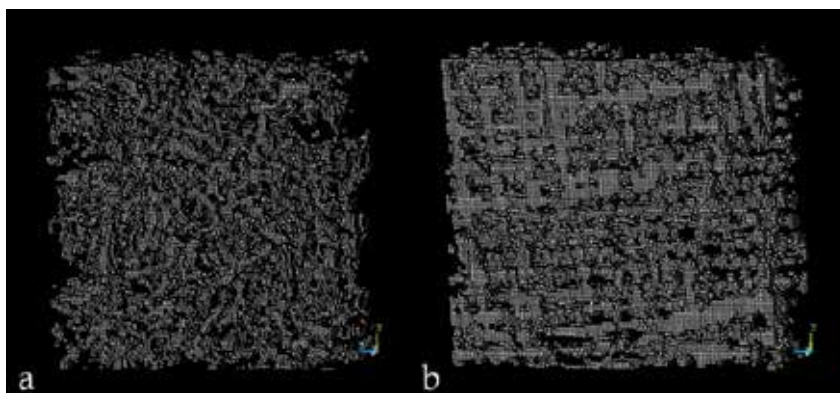


Fig. 4. μ FE meshes corresponding to pore based (a) and fiber based (b) polymer scaffolds.

4. FE model definition

Once the mesh is loaded, the model has to be defined by the specification of the material properties, the boundary conditions and the specific essay to be simulated. Regarding the material properties, each element is assigned linear elastic isotropic properties consisting of a bulk elasticity modulus of $E_b=1\text{MPa}$ and a Poisson's ratio $\nu=0.3$.

The aim of the process is to simulate a compressive stress-strain test of the scaffold. To specify this situation, first, boundary conditions must be defined. A null displacement in the three directions of space is imposed on nodes from one side of the sample. Thus, this side remains fixed when any load is applied to the structure. After definition of boundary conditions, the situation required to simulate compression needs to be specified. Thus, a deformation of 10% of the scaffold edge length is imposed on nodes from the opposite side.

5. Simulation and results calculation

Simulations of the generated μFE models are performed by the application of the generalized structural mechanics theory (Zienkiewicz et al., 2006) to the system of equations defined by the previously imposed conditions. The mathematical steps for the solution calculation are widely explained in chapter "*Finite Element Modeling for a Morphometric and Mechanical Characterization of Trabecular Bone from High Resolution Magnetic Resonance Imaging*" of this book.

In practice, the solution is calculated in efficient software platforms specially designed for large array processing. The identification of one element to each voxel supposes a high computational burden, especially when the number of nodes tends to be greater than 1×10^6 , due to the large stiffness matrices to be assembled and the significant increase in the degrees of freedom. There are three degrees of freedom per node and each element has 8 nodes (24 equations per element). Different methods may be used for solving these large systems of differential equations. In our case, as mentioned in chapter about FE analysis of trabecular bone of this book, for trabecular bone simulations, systems are solved by Gaussian elimination using a standard sparse solver. All the nodal forces and displacements of the structure are obtained and the nodal stresses and strains can be calculated. In addition, the elastic modulus of the porous structure (E) can be efficiently estimated by the homogenization theory approximation, based on linear elastic behaviour and small deformations theory (Hollister et al., 1991; Hollister & Kikuchi, 1992).

The different solutions can be represented in a parametric 3D reconstruction in order to graphically evaluate regions with certain mechanical conditions. If a compressive essay is simulated in the different scaffolds shown in figures 1 and 2, it is observed that the normal stress distributions are not similar between the different architectures. In this sense, Figure 5a-c shows that the scaffolds with spherical pores have a uniform stress distribution and pores do not act as stress concentrators as it would happen if only one of spherical pore were surrounded by the scaffolding material (Alberich-Bayarri et al., 2009). However, stress tends to concentrate around the channels of the scaffold with crossed fibers, as it can be observed in figure 5b. This could be explained by the inherent anisotropy of the scaffold and by the lower porosity in comparison to spherical pores scaffolds.

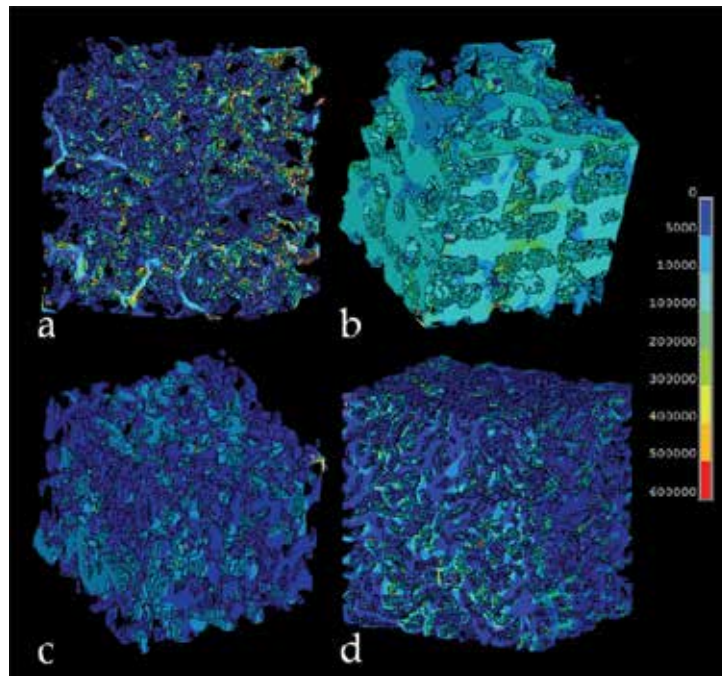


Fig. 5. Distribution of normal stresses. Same topologies distribution than in figures 1 and 2. Spherical pores in a), c) and d); orthogonal pore mesh in b).

Results can be quantitatively compared between different architectures. Table 1 shows the stresses and Young's modulus obtained with μ FE. It can be observed that the scaffold based on crossed channels (which corresponds to Figure 5-b) shows a high maximum stress and apparent Young's modulus in comparison to other topologies.

	PLLA (salt particles)	PMMA (sphere porogens)	Fiber templates
S_{MAX} [MPa]	0,82	0,83	1,02
E_{app} [kPa]	2.53	19.99	152.44

Table 1. Maximum stresses and apparent Young's modulus obtained for three different scaffold architectures (Poly-L-Lactide Acid, PLLA; Poly(methyl methacrylate), PMMA; fiber templates).

If the results obtained for the apparent Young's modulus parameter are compared to structure porosity, that can be directly calculated from the 3D reconstructions, it is appreciated that a dependence of the normalized modulus (to the bulk) with the square of porosity exists. Both the experimental measurements and the μ FE results are well correlated with the proposed exponential dependence for the compressive modulus on porosity (Gibson & Ashby, 2001), as it can be observed in figure 6.

$$\frac{E_{app}}{E_b} = C \left(\frac{\rho}{\rho_b} \right)^2 \quad (1)$$

Where ρ and ρ_b are the apparent density of the scaffold and bulk density of the constituting material, respectively.

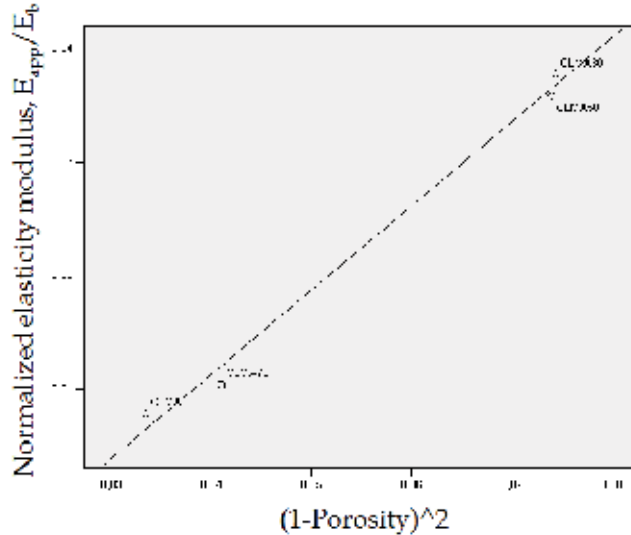


Fig. 6. Relationship ($r^2=0.92$) between normalized Young's modulus and the square of porosity in four different PMMA scaffolds analyzed.

The existence of a general relationship (in the elastic regime) between the modulus and porosity is evidenced with the analysis of these results. This has been also suggested in previous works in which interconnected spherical pores of different sizes and interconnected throats have been investigated (Diego et al., 2007). Even more, a similar behaviour was observed for Young's modulus analysis of trabecular bone in a population of healthy patients (Alberich-Bayarri et al., 2008; Gibson & Ashby, 2001).

6. Validation with experimental measurements

The μ FE simulations and measurements are performed in a non-destructive way, that is, mechanical properties are extracted without compromising the structure to real mechanical loads. Although the FE methods have been widely extended and are essential in any engineering process because of their accuracy and reliability, the last affirmation is relatively ambitious and the indirect measurements performed by FE method need to be validated with the standard reference technique, that is, with experimental compression stress-strain measurements. This technique performs a continuous sampling of the stress-strain relationship during the essay to obtain curves showing the mechanical behaviour of the specimen being analyzed. An example of a real stress-strain curve can be observed in figure 7.

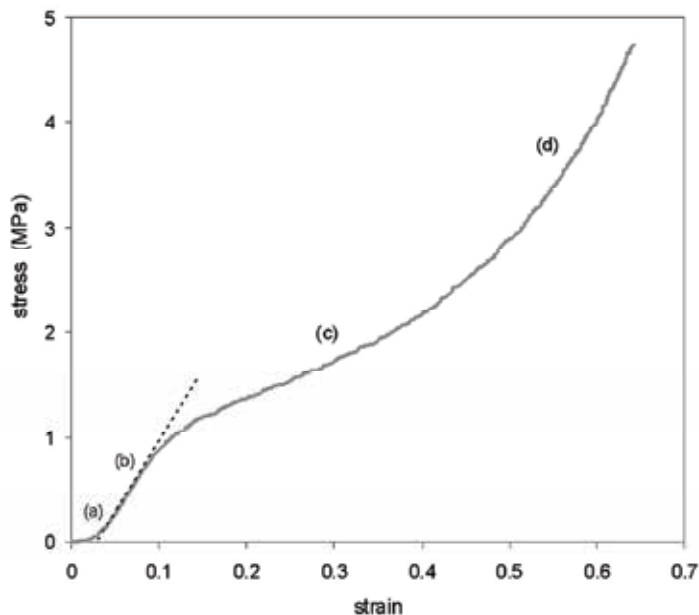


Fig. 7. Experimental stress-strain curve of a scaffold with interconnected spherical pores. The letters on the curve are related to different mechanical features. Briefly: (a) contact region between the sample and the device; (b) the linear elastic region (the elastic modulus was calculated from the slope of the dotted line); (c) buckling phenomenon leading to the so-called plateau; and (d) final collapse of the structure increasing the compressive modulus. (Adapted from Alberich-Bayarri et al., 2009).

In order to evaluate the measurements of scaffolds Young's modulus, the measurements performed in two different scaffolds topologies (interconnected spherical pores, crossed fibers) using both μ FE and experimental techniques were compared. Experimental compression stress-strain measurements were performed on a Microtest system with a load cell of 15N. The samples consisted on cylinders with an approximated radius of 5mm and 3mm height. The initial range of strains (0–0.1) was underestimated due to the non-conformity of the contact between the machine plate and the specimen. Thus, the experimental compressive modulus of the scaffold was not easy to measure. The test was carried out until microstructure was completely collapsed. Finally, the compressive modulus was determined from the initial linear slope of the curve after full contact between plate and specimen was ensured.

In figure 8, the mentioned dependence of the normalized modulus (to the bulk) as a function of the square of porosity can be observed. The μ FE and the experimental results can be also compared. Both methods are accomplishing the proposed exponential relationship between the compressive modulus and scaffold porosity. The effectiveness of the μ FE method used for the calculation of mechanical characteristics is demonstrated. The obtained values of elasticity modulus by the μ FE method are well related to those obtained experimentally. The agreement between the experimental results and the μ FE simulations supports the feasibility of the technique as a tool for scaffold design and non-invasive analysis on real fabricated scaffolds before physical experiments are planned.

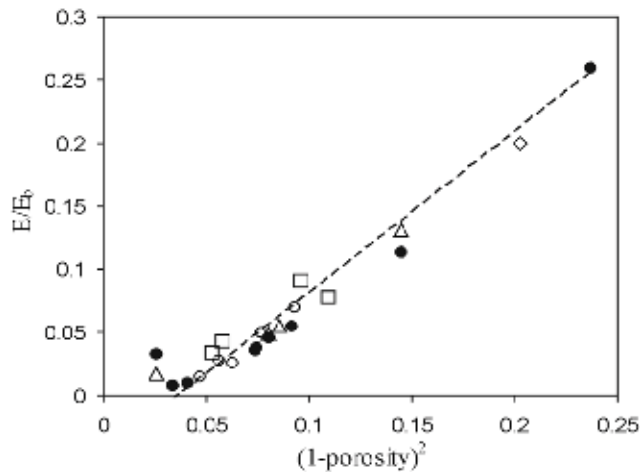


Fig. 8. μ FE calculated (black dots) and experimentally measured normalized (to the bulk E_b) moduli for the different structures: (squares) constructs with spherical pores, (diamonds) construct with orthogonal cylindrical pores, (delta) construct based on salt particles. Additional points (circle) from (Diego et al. 2007) have been included so as to reinforce the universality of the relationship between the (reduced) elastic modulus and porosity. (Adapted from Alberich-Bayarri et al., 2009).

7. Conclusions and future challenges

During the last decades, treatment of bone defects has been directed to the research in polymer scaffolds of different architectures. Scaffolds play a key role in bone regeneration and the design of their 3D architecture is crucial to develop their function.

The use of high spatial resolution acquisitions from μ CT scanners in combination with advanced image processing methods represent a powerful tool to develop a structural and mechanical characterization of the synthesized polymer scaffolds and evaluate the achievement of the desired properties at the manufacture stage.

The algorithms employed for the image analysis and fast meshing processes allows for the detailed numerical simulations of the mechanical properties at a micro scale level, also called μ FE. These simulations suppose a high computational burden that can be optimized by proper μ FE model definition. μ FE results obtained for the apparent Young's modulus are highly close to the experimental results.

All the simulations performed in the synthesized scaffolds using the μ FE method have been done in the linear behaviour domain. However, in future it would be of high interest to improve in the knowledge of the non-linear behaviour of the scaffolds when strong compressive conditions are considered and buckling processes begin.

Also, the analysis of the scaffolds using μ CT combined with μ FE before and after cell seeding would help to evaluate the non-invasive method proposed as a reliable way to quantify the levels of new tissue deposition in the scaffold.

Acknowledgements

The support of the Spanish Ministry of Science and Innovation through project TEC2009-14128 and of the Generalitat Valenciana through projects GV/2009/126 (grups d'investigació emergents) and ACOMP/2010/022 (ajudes complementàries) is acknowledged.

The authors like to thank the 3B's Research Group-Biomaterials, Biodegradables and Biomimetics of the University of Minho (Braga, Portugal) for the μ CT acquisitions.

8. References

- Alberich-Bayarri, A.; Moratal, D.; Martí-Bonmatí, L.; Salmerón-Sánchez, M.; Vallés-Lluch, A.; Nieto-Charques, L. & Rieta JJ. (2007). Volume mesh generation and finite element analysis of trabecular bone magnetic resonance images. *Proceedings of: Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 1603-1606, 1557-170X, Lyon, France, August 2007, *IEEE Engineering in Medicine and Biology Society*.
- Alberich-Bayarri, A.; Martí-Bonmatí, L.; Sanz-Requena, R.; Belloch, E. & Moratal, D. (2008). In vivo trabecular bone morphologic and mechanical relationship using high-resolution 3-T MRI. *American Journal of Roentgenology*, 191, 3, (Sep 2008), 721-726, 0361-803X.
- Alberich-Bayarri, A.; Moratal, D.; Ivirico, J.L.; Rodríguez Hernández, J.C.; Vallés-Lluch, A.; Martí-Bonmatí, L.; Estellés, J.M.; Mano, J.F.; Pradas, M.M.; Ribelles, J.L. & Salmerón-Sánchez, M. (2009). Microcomputed tomography and microfinite element modeling for evaluating polymer scaffolds architecture and their mechanical properties. *Journal of biomedical materials research. Part B, Applied biomaterials*, 91, 1, (Oct 2009), 191-202, 1552-4973.
- Diego, R. B.; Estelles, J. M.; Sanz, J. A.; García-Aznar, J. M. & Sánchez, M. S.. (2007). Polymer scaffolds with interconnected spherical pores and controlled architecture for tissue engineering: fabrication, mechanical properties and finite element modeling. *Journal of biomedical materials research. Part B, Applied biomaterials*, 81, 2 (May 2007), 448-455, 1552-4973.
- Dupraz, A.; Delecrin, J.; Moreau, A.; Pilet, P. & Passuti, N. (1998). Long-term bone response to particulate injectable ceramic. *Journal of biomedical materials research*, 42, 3 (Dec 1998), 368-375, 0021-9304.
- Freed, L.E.; Hollander, A.P.; Martin, I.; Barry, L.R. & Vunjak-Novakovic, G. (1998). Chondrogenesis in a cell-polymer-bioreactor system. *Experimental cell research*, 240, 1 (Apr 1998), 58-65, 0014-4827.
- Gao, C.Y.; Wang, D.Y. & Shen, J.C. (2003). Fabrication of porous collagen/chitosan scaffolds with controlling microstructure for dermal equivalent. *Polymers for advanced technologies*, 14, 6 (May 2003), 373-379, 1042-7147.
- Gauthier, O.; Khairoun, I.; Bosco, J.; Obadia, L.; Bourges, X.; Rau, C.; Magne, D.; Bouler, J.M.; Aguado, E. & Daculsi, G. (2003). Noninvasive bone replacement with a new injectable calcium phosphate biomaterial. *Journal of biomedical materials research. Part A*, 66, 1 (Jul 2003), 47-54, 1549-3296.

- Geffre, C.P.; Margolis, D.S.; Ruth, J.T.; DeYoung, D.W.; Tellis, B.C. & Szivek, J.A. (2009). A Novel Biomimetic Polymer Scaffold Design Enhances Bone Ingrowth. *Journal of biomedical materials research. Part A*, 91, 3 (Dec 2009), 795–805, 1549–3296.
- Gibson, L.J. & Ashby M. (2001). The mechanics of foam: Basic results. In: *Cellular Solids: Structure and Properties*, Clarke, D.R.; Suresh, S. & Ward, I.M., 175–234, Cambridge University Press, 0521499119, Cambridge, UK.
- Hajiali, H.; Karbasi, S.; Hosseinalipour, M. & Rezaie, H.R. (2010). Preparation of a novel biodegradable nanocomposite scaffold based on poly (3-hydroxybutyrate)/bioglass nanoparticles for bone tissue engineering. *Journal of Materials Science. Materials in Medicine*, 2010. [Epub ahead of print], 0957–4530.
- Harris, L.D.; Kim, B.S. & Mooney, D.J. (1998). Open pore biodegradable matrices formed with gas foaming. *Journal of biomedical materials research*, 42, 3, (Dec 1998), 396–402, 0021–9304.
- Ho, S.T. & Hutmacher, D.W. (2006). A comparison of micro CT with other techniques used in the characterization of scaffolds. *Biomaterials*, 27, 8 (Mar 2006), 1362–1376, 0142–9612.
- Hollister, S.J.; Fyhrie, D.P.; Jepsen, K.J. & Goldstein, S.A. (1991). Application of homogenization theory to the study of trabecular bone mechanics. *Journal of biomechanics*, 24, 9, 825–839, 0021–9290.
- Hollister, S.J. & Kikuchi, N. (1992). A comparison of homogenization theory and standard mechanics analyses for periodic porous composites. *Computational Mechanics*, 10, 2 (Mar 1992), 73–95, 0178–7675.
- Horbett, T.A.; Schway, M.B. & Ratner, B.D. (1985). Hydrophilic-hydrophobic copolymers as cell substrates-effect on 3T3 cell-growth rates. *Journal of colloid and interface science*, 104, 1 (Mar 1985), 28–39, 0021–9797.
- Hutmacher, D.W. (2000). Scaffolds in tissue engineering bone and cartilage. *Biomaterials*, 21, 24 (Dec 2000), 2529–2543, 0142–9612.
- Jaecques, S.V.N.; Van Oosterwyck, H.; Muraru, L.; Van Cleynenbreugel, T.; De Smet, E.; Wevers, M.; Naert, I. & Vander Floten, J. (2004). Individualised, micro CT-based finite element modelling as a tool for biomechanical analysis related to tissue engineering of bone. *Biomaterials*, 25, 9 (Apr 2004), 1683–1696, 0142–9612.
- Lacroix, D.; Chateau, A.; Ginebra, M.P. & Planell, J.A. (2006). Micro-finite element models of bone-tissue engineering scaffolds. *Biomaterials*, 27, 30 (Oct 2006), 5326–5334, 0142–9612.
- Langer, R. & Vacanti, J.P. (1993). Tissue Engineering. *Science*, 260, 5110 (May 1993), 920–926, 0193–4511.
- Lorensen, W.E. & Cline, H.E. (1987). Marching Cubes: A high resolution 3D surface construction algorithm. *Computers & Graphics*, 21, 4 (July 1987), 163–169, 0097–8493.
- Moroni, L.; de Wijn, J.R. & Blitterswijk, C.A. (2006). 3D fiber-deposited scaffolds for tissue engineering: Influence of pores geometry and architecture on dynamic mechanical properties. *Biomaterials*, 27, 7 (Mar 2006), 974–985, 0142–9612.
- Oliveira, S.M.; Mijares, D.Q.; Turner, G.; Amaral, I.F.; Barbosa, M.A. & Teixeira, C.C. (2009). Engineering endochondral bone: in vivo studies. *Tissue engineering. Part A*, 15, 3 (Mar 2009), 635–643, 1937–3341.
- Otsu, N. (1979). A threshold selection method from gray-level histogram. *IEEE Transactions on Systems, Man, and Cybernetics*, 9, 1, 82–86, 1083–4427.

- Mikos, A.G.; Sarakinos, G.; Leite, S.M.; Vacanti, J.P. & Langer, R. (1993). Laminated 3-dimensional biodegradable foams for use in tissue engineering. *Biomaterials*, 14, 5 (Apr 1993), 323-330, 0142-9612.
- Parsons, J.R. (1985). Resorbable materials and composites. New concepts in orthopedic biomaterials. *Orthopedics*, 8, 7 (Jul 1985), 907-915, 0147-7447.
- Rajan, G.P.; Fornaro, J.; Trentz, O. & Zellweger, R. (2006). Cancellous allograft versus autologous bone grafting for repair of comminuted distal radius fractures: a prospective, randomized trial. *The Journal of trauma*, 60, 6 (Jun 2006), 1322-1329, 0022-5282.
- Thomson, R.C.; Wake, M.C.; Yaszemski, M. & Mikos, A.G. (1995). Biodegradable polymer scaffolds to regenerate organs. *Advances in Polymer Science*, 122, 1995, 247-274, 0065-3195.
- van Lenthe, G.H.; Hagenmüller, H.; Bohner, M.; Hollister, S.J.; Meinel, L. & Müller, R. (2007). Nondestructive micro-computed tomography for biological imaging and quantification of scaffold-bone interaction in vivo. *Biomaterials*, 28, 15 (May 2007), 2749-2490, 0142-9612.
- Whang, K.; Thomas, C.H.; Healy, K.E. & Nuber, G. (1995). A novel method to fabricate bioabsorbable scaffolds. *Polymer*, 36, 4 (Feb 1995), 837-842, 0032-3861.
- Yang, S.; Leong, K.F.; Du, Z. & Chua, C.K. (2001). The design of scaffolds for use in tissue engineering. Part I. Traditional factors. *Tissue engineering*, 7, 6 (Dec 2001), 679-689, 1076-3279.
- Zhang, R. & Ma P.X. (1999). Poly(alpha-hydroxyl acids)/hydroxyapatite porous composites for bone-tissue engineering. I. Preparation and morphology. *Journal of biomedical materials research*, 44, 4 (Mar 1999), 446-455, 0021-9304.
- Zhou, Q.L.; Gong, Y.H. & Gao, C.Y. (2005). Microstructure and mechanical properties of poly(L-lactide) scaffolds fabricated by gelatin particle leaching method. *Journal of applied polymer science. Applied polymer symposium*, 98, 3 (Aug 2005), 1373-1379, 0271-9460.
- Zienkiewicz, O.C.; Taylor, R.L. & Zhu, J.Z. (2006). *The finite element method, its basis & fundamentals*. 6th edition. Elsevier, Oxford, UK. 2006. 0750663200

Computational Modelling of Auxetics

Bogdan Maruszewski and Tomasz Strek

Poznan University of Technology

Institute of Applied Mechanics

Poland

Artur A. Pozniak

Poznan University of Technology

Institute of Physics

Poland

Krzysztof W. Wojciechowski

PWSZ im. Prezydenta St. Wojciechowskiego w Kaliszu

Poland

1. Introduction

Modern technologies require new materials of special properties. One of the reasons for interest in materials of unusual mechanical properties comes from the fact that they can be used (either as inclusions or as matrices) to form composites of required properties.

There is a number of physical properties that we implicitly assume to be positive. However, one may be surprised to discover that they can also be negative. Negative materials include, amongst other ones, those having negative stiffness (Lakes et al., 2001), negative thermal expansion (Hartwig, 1995), negative refractive index (Sang & Li, 2005), negative permittivity (Ruppin, 2000) and/or negative permeability (Ruppin, 2000). It is worth to add that in presence of some constraints even the compressibility can be negative (Lakes & Wojciechowski, 2008).

A new field of challenge are studies of materials exhibiting negative Poisson's ratio. The latter is a negative ratio of relative transverse dimension change to relative longitudinal dimension change of a body when an infinitesimal change of a stress acting along the longitudinal direction occurs whereas the other stress components remain unchanged. Such materials, first manufactured by Lakes (Lakes, 1987) and coined *auxetics* by Evans (Evans, 1991), are a subject of intensive studies both in the context of fundamental research and applications (Remillat et al., 2009).

The aim of this chapter is to demonstrate recently discovered anomalous deformation of an auxetic plate, constrained by fixing two opposite sides, which is loaded by uniform tension (or compression) applied perpendicularly to two other opposite sides of the plate. The problem was studied both in three dimensions (3D) by Strek et al. (Strek et al., 2008) and in two dimensions (2D) by Pozniak et al. (Pozniak et al., 2010) by finite element methods. In all the cases studied it has been assumed that the material was isotropic.

The paper (Strek et al., 2008) dealt with computer simulations of mechanical behaviour of a thick elastic plate. Simulations have been done for Poisson's ratio from interval $-1 < \nu < 0.5$ using COMSOL (Comsol, 2007). An anomalous feature of the plate deformation for negative Poisson's ratio values compared to classical positive values has been observed at strongly negative Poisson's ratios, $\nu < -0.7$. For such values of ν the displacement vector has components which are anti-parallel to the direction of loading.

2D version of this system, described in (Pożniak et al., 2010), allowed one for more precise computations using much finer meshes than those used in the 3D case. In consequence, the 2D simulations performed with FEniCS (Logg & Wells 2010) revealed the anomalous behaviour of the displacement vector already at $\nu < -0.25$.

The anomalous behaviour of the displacement vector, which in some parts of the plate has components opposite to the direction of the applied force can be thought of as locally negative compliance. Systems with negative compliance have been recently studied by Lakes and co-workers (Lakes, 2001; Lakes et al., 2001; Jaglinski et al., 2007). The reason is that combination of such (negative) materials with common ones (of positive compliance) of the same absolute value offers composites of zero compliance, i.e. of infinite elastic moduli.

In the present chapter we briefly review the results obtained in (Strek et al., 2008) and (Pozniak et al., 2010). By studying larger meshes in 3D and finer ones in 2D we extend those investigations to computationally 'larger' systems. This allows one to study, respectively, thicker plates in 3D and more subtle effects both in 3D and 2D cases. In consequence, we get a better insight in the unusual phenomenon under study.

2. Modelling methods and tools

A great deal of computational research has been undertaken and published in the field of computational mechanics since the advent of the digital computer. Before 1970, the Finite Difference Method (FDM) was almost universally used as a computer based numerical method in modeling dynamics process. Since then there has been a revolution in the general area of mathematical modeling. Highly sophisticated and detailed analysis of many engineering problems has become possible. However, it can be argued that the last three decades have in many ways belonged to the Finite Element Method (FEM) as the method of choice among the currently available numerical methods for solving mathematical equations (Huebner, 1975; Hinton and Owen, 1979).

All mechanical problems considered in this work are governed by equations with appropriate boundary and initial conditions. Numerical results for 3D systems are obtained using standard computational code COMSOL Multiphysics (Comsol, 2004; Comsol, 2007). As COMSOL implicitly simulates 3D systems, to study 2D cases another package, known as FEniCS, was applied. ABAQUS was used to test the obtained results both in 3D and 2D.

2.1 Comsol Multiphysics

Theory in this section is based on COMSOL Multiphysics manual (Comsol, 2007). COMSOL Multiphysics is a powerful interactive environment for modelling and solving all kinds of scientific and engineering problems based on partial differential equations (PDEs) using the finite element method. One can access the power of COMSOL Multiphysics as a standalone product, by script programming in the COMSOL Script language or in the MATLAB language (Comsol, 2007).

A general time-dependent PDE problem in the coefficient form used by COMSOL results in the following equation system (Comsol, 2007)

$$\mathbf{e}_a \frac{\partial^2 \mathbf{u}}{\partial t^2} + \mathbf{d}_a \frac{\partial \mathbf{u}}{\partial t} - \nabla \cdot (\mathbf{c} \nabla \mathbf{u} + \boldsymbol{\alpha} \mathbf{u} - \boldsymbol{\gamma}) + \boldsymbol{\beta} \cdot \nabla \mathbf{u} + \mathbf{a} \mathbf{u} = \mathbf{F} \quad (1)$$

with boundary conditions

$$\mathbf{n} \cdot (\mathbf{c} \nabla \mathbf{u} + \boldsymbol{\alpha} \mathbf{u} - \boldsymbol{\gamma}) + \mathbf{q} \mathbf{u} = \mathbf{g} - \mathbf{h}^T \boldsymbol{\mu} \quad (2)$$

and

$$\mathbf{h} \mathbf{u} = \mathbf{r} . \quad (3)$$

The first equation (1) is satisfied inside the domain, whereas the second (2), representing so called generalized Neumann boundary condition, and the third (3) – so called Dirichlet boundary condition, are both satisfied on the boundary of domain. In this work all governing equations obey a general time-dependent PDE problem in the coefficient form reduced to equation

$$-\nabla \cdot (\mathbf{c} \nabla \mathbf{u}) = \mathbf{F} \quad (4)$$

where the diffusive term flux is defined as

$$\mathbf{c} \nabla \mathbf{u} = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix} \nabla \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix} \cdot \begin{bmatrix} \nabla u_1 \\ \nabla u_2 \\ \nabla u_3 \end{bmatrix} = \begin{bmatrix} c_{11} \frac{\partial u_1}{\partial x} + c_{12} \frac{\partial u_2}{\partial x} + c_{13} \frac{\partial u_3}{\partial x} \\ c_{11} \frac{\partial u_1}{\partial y} + c_{12} \frac{\partial u_2}{\partial y} + c_{13} \frac{\partial u_3}{\partial y} \\ c_{11} \frac{\partial u_1}{\partial z} + c_{12} \frac{\partial u_2}{\partial z} + c_{13} \frac{\partial u_3}{\partial z} \\ c_{21} \frac{\partial u_1}{\partial x} + c_{22} \frac{\partial u_2}{\partial x} + c_{23} \frac{\partial u_3}{\partial x} \\ c_{21} \frac{\partial u_1}{\partial y} + c_{22} \frac{\partial u_2}{\partial y} + c_{23} \frac{\partial u_3}{\partial y} \\ c_{21} \frac{\partial u_1}{\partial z} + c_{22} \frac{\partial u_2}{\partial z} + c_{23} \frac{\partial u_3}{\partial z} \\ c_{31} \frac{\partial u_1}{\partial x} + c_{32} \frac{\partial u_2}{\partial x} + c_{33} \frac{\partial u_3}{\partial x} \\ c_{31} \frac{\partial u_1}{\partial y} + c_{32} \frac{\partial u_2}{\partial y} + c_{33} \frac{\partial u_3}{\partial y} \\ c_{31} \frac{\partial u_1}{\partial z} + c_{32} \frac{\partial u_2}{\partial z} + c_{33} \frac{\partial u_3}{\partial z} \end{bmatrix} \quad (5)$$

where ∇u_i are column vectors. The flux matrix or flux tensor is a column vector in this work. For anisotropic materials, each of the components of \mathbf{c} can be a matrix.

2.2 FEniCS

To study 2D systems FEniCS was applied. FEniCS Project (Logg & Wells 2010) is a software suite dedicated to Finite Element Analysis laying emphasis to partial differential equations. DOLFIN may be regarded as its central part being responsible for dealing with the FEM issues. All components of FEniCS are released under GNU General Public License or GNU Lesser General Public License. The sources as well as compiled packages are freely available through <http://www.fenics.org>.

As mentioned, FEniCS is not a monolithic project but consists of a few components. FIAT is responsible for finite element basis function evaluation. Variational forms coming from weak formulations of PDEs are handled by the Unified Form Language and the FEniCS Form Compiler.

Dealing with FEniCS requires knowledge of the weak forms written in the Unified Form Language in order to let the FEniCS form compiler generate the low level code. Here the bilinear form is written as $a = \text{inner}(\epsilon(v), \sigma(u)) dx$ which stands for the integral

$\int_X \epsilon \cdot \sigma dx$, the linear form is $L = \text{inner}(v, f) dx$, representing $\int_X v \cdot f dx$, v and u are trial and test

functions respectively defined in UFL $v = \text{TrialFunction}(\text{element})$, $u = \text{TestFunction}(\text{element})$. Element definition is as simple as previous ones, namely $\text{element} = \text{VectorElement}(\text{"Lagrange"}, \text{"triangle"}, 1)$ for the first order Lagrange element. Putting $a = L$ one gets the weak form of differential equation of elasticity.

3. Elastic deformations

3.1 3D case

It is possible to completely describe the strain conditions at a point with the deformation components, (u, v, w) in 3D and their derivatives (Landau and Lifshits, 1986). One can express the shear strain in a tensor form, ϵ_{xy} , ϵ_{yz} , ϵ_{xz} or in an engineering form, γ_{xy} , γ_{yz} , γ_{xz} . Following the small-displacement assumption, the normal strain components and the shear strain components are given from the deformation as follows

$$\begin{aligned} \epsilon_x &= \frac{\partial u}{\partial x}, \quad \epsilon_{xy} = \frac{\gamma_{xy}}{2} = \frac{1}{2} \left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \right), \\ \epsilon_y &= \frac{\partial v}{\partial y}, \quad \epsilon_{yz} = \frac{\gamma_{yz}}{2} = \frac{1}{2} \left(\frac{\partial v}{\partial z} + \frac{\partial w}{\partial y} \right), \\ \epsilon_z &= \frac{\partial w}{\partial z}, \quad \epsilon_{xz} = \frac{\gamma_{xz}}{2} = \frac{1}{2} \left(\frac{\partial u}{\partial z} + \frac{\partial w}{\partial x} \right). \end{aligned} \tag{6}$$

The symmetric strain tensor, $\epsilon = \frac{1}{2}(\nabla \mathbf{u} + (\nabla \mathbf{u})^T)$, consists of both normal and shear strain components

$$\boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_x & \varepsilon_{xy} & \varepsilon_{xz} \\ \varepsilon_{yx} & \varepsilon_y & \varepsilon_{yz} \\ \varepsilon_{zx} & \varepsilon_{zy} & \varepsilon_z \end{bmatrix}. \quad (7)$$

The stress in a material is described by the symmetric stress tensor $\boldsymbol{\sigma}$

$$\boldsymbol{\sigma} = \begin{bmatrix} \sigma_x & \tau_{xy} & \tau_{xz} \\ \tau_{yx} & \sigma_y & \tau_{yz} \\ \tau_{zx} & \tau_{zy} & \sigma_z \end{bmatrix} \quad (8)$$

consisting of three normal stresses $(\sigma_x, \sigma_y, \sigma_z)$ and six, or if symmetry is used, three shear stresses $(\tau_{xy}, \tau_{yz}, \tau_{xz})$. The stress-strain relationship (the constitutive law) for linear conditions reads

$$\boldsymbol{\sigma} = \mathbf{D}\boldsymbol{\varepsilon} \quad (9)$$

where \mathbf{D} is a 6×6 elasticity matrix, and the stress and strain components are described in vector form with the six stress and strain components in column vectors defined as

$$\begin{aligned} \boldsymbol{\sigma} &= [\sigma_x, \sigma_y, \sigma_z, \tau_{xy}, \tau_{yz}, \tau_{xz}]^T, \\ \boldsymbol{\varepsilon} &= [\varepsilon_x, \varepsilon_y, \varepsilon_z, \gamma_{xy}, \gamma_{yz}, \gamma_{xz}]^T. \end{aligned} \quad (10)$$

For static conditions Navier's equation reads (Landau and Lifshits, 1986)

$$-\nabla \cdot \boldsymbol{\sigma} = \mathbf{F} \quad (11)$$

where \mathbf{u} denotes the displacement and \mathbf{F} denotes the volume forces (body forces).

Lamé's constants λ and μ in terms of Young's modulus, E , and Poisson's ratio, ν , are the following

$$\begin{aligned} \lambda &= \frac{E\nu}{(1+\nu)(1-2\nu)}, \\ \mu &= \frac{E}{2(1+\nu)}, \end{aligned} \quad (12)$$

then the elastic matrix \mathbf{D} reads

$$\mathbf{D} = \frac{E}{(1+\nu)(1-2\nu)} \begin{bmatrix} 1-\nu & \nu & \nu & 0 & 0 & 0 \\ \nu & 1-\nu & \nu & 0 & 0 & 0 \\ \nu & \nu & 1-\nu & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1-2\nu}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1-2\nu}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1-2\nu}{2} \end{bmatrix}. \quad (13)$$

The use of (7)–(10) in (6) leads us to Navier's equation - the final form of the equation (Comsol, 2004; Comsol, 2007)

$$-\nabla \cdot (\mathbf{c} \nabla \mathbf{u}) = \mathbf{F}, \quad (14)$$

where \mathbf{c} is the flux matrix. The flux matrix \mathbf{c} reads

$$\mathbf{c} = \begin{bmatrix} \begin{bmatrix} D_{11} & D_{14} & D_{16} \\ D_{14} & D_{44} & D_{46} \\ D_{16} & D_{46} & D_{66} \end{bmatrix} & \begin{bmatrix} D_{14} & D_{12} & D_{15} \\ D_{44} & D_{24} & D_{45} \\ D_{46} & D_{26} & D_{56} \end{bmatrix} & \begin{bmatrix} D_{16} & D_{15} & D_{13} \\ D_{46} & D_{45} & D_{34} \\ D_{66} & D_{56} & D_{36} \end{bmatrix} \\ \begin{bmatrix} D_{14} & D_{44} & D_{46} \\ D_{12} & D_{24} & D_{26} \\ D_{15} & D_{45} & D_{56} \end{bmatrix} & \begin{bmatrix} D_{44} & D_{24} & D_{45} \\ D_{24} & D_{22} & D_{25} \\ D_{45} & D_{25} & D_{55} \end{bmatrix} & \begin{bmatrix} D_{46} & D_{45} & D_{34} \\ D_{26} & D_{25} & D_{23} \\ D_{56} & D_{55} & D_{35} \end{bmatrix} \\ \begin{bmatrix} D_{16} & D_{46} & D_{66} \\ D_{15} & D_{45} & D_{56} \\ D_{13} & D_{34} & D_{36} \end{bmatrix} & \begin{bmatrix} D_{46} & D_{26} & D_{56} \\ D_{45} & D_{25} & D_{55} \\ D_{34} & D_{23} & D_{35} \end{bmatrix} & \begin{bmatrix} D_{66} & D_{56} & D_{36} \\ D_{56} & D_{55} & D_{35} \\ D_{36} & D_{35} & D_{33} \end{bmatrix} \end{bmatrix}. \quad (15)$$

D_{ij} in the \mathbf{c} matrix is referring to the component in the elasticity matrix (13) in the stress-strain relation for 3D.

In this case, the diffusive flux, reads

$$\mathbf{c} \nabla \mathbf{u} = \begin{bmatrix} \begin{bmatrix} D_{11} & D_{14} & D_{16} \\ D_{14} & D_{44} & D_{46} \\ D_{16} & D_{46} & D_{66} \end{bmatrix} & \begin{bmatrix} D_{14} & D_{12} & D_{15} \\ D_{44} & D_{24} & D_{45} \\ D_{46} & D_{26} & D_{56} \end{bmatrix} & \begin{bmatrix} D_{16} & D_{15} & D_{13} \\ D_{46} & D_{45} & D_{34} \\ D_{66} & D_{56} & D_{36} \end{bmatrix} \\ \begin{bmatrix} D_{14} & D_{44} & D_{46} \\ D_{12} & D_{24} & D_{26} \\ D_{15} & D_{45} & D_{56} \end{bmatrix} & \begin{bmatrix} D_{44} & D_{24} & D_{45} \\ D_{24} & D_{22} & D_{25} \\ D_{45} & D_{25} & D_{55} \end{bmatrix} & \begin{bmatrix} D_{46} & D_{45} & D_{34} \\ D_{26} & D_{25} & D_{23} \\ D_{56} & D_{55} & D_{35} \end{bmatrix} \\ \begin{bmatrix} D_{16} & D_{46} & D_{66} \\ D_{15} & D_{45} & D_{56} \\ D_{13} & D_{34} & D_{36} \end{bmatrix} & \begin{bmatrix} D_{46} & D_{26} & D_{56} \\ D_{45} & D_{25} & D_{55} \\ D_{34} & D_{23} & D_{35} \end{bmatrix} & \begin{bmatrix} D_{66} & D_{56} & D_{36} \\ D_{56} & D_{55} & D_{35} \\ D_{36} & D_{35} & D_{33} \end{bmatrix} \end{bmatrix} \begin{bmatrix} \nabla u_1 \\ \nabla u_2 \\ \nabla u_3 \end{bmatrix}. \quad (16)$$

After some mathematical calculations one can write equation (16) in the following form

$$\mathbf{c} \nabla \mathbf{u} = \begin{bmatrix} \begin{bmatrix} D_{11} & D_{14} & D_{16} \\ D_{14} & D_{44} & D_{46} \\ D_{16} & D_{46} & D_{66} \end{bmatrix} \nabla u_1 + \begin{bmatrix} D_{14} & D_{12} & D_{15} \\ D_{44} & D_{24} & D_{45} \\ D_{46} & D_{26} & D_{56} \end{bmatrix} \nabla u_2 + \begin{bmatrix} D_{16} & D_{15} & D_{13} \\ D_{46} & D_{45} & D_{34} \\ D_{66} & D_{56} & D_{36} \end{bmatrix} \nabla u_3 \\ \begin{bmatrix} D_{14} & D_{44} & D_{46} \\ D_{12} & D_{24} & D_{26} \\ D_{15} & D_{45} & D_{56} \end{bmatrix} \nabla u_1 + \begin{bmatrix} D_{44} & D_{24} & D_{45} \\ D_{24} & D_{22} & D_{25} \\ D_{45} & D_{25} & D_{55} \end{bmatrix} \nabla u_2 + \begin{bmatrix} D_{46} & D_{45} & D_{34} \\ D_{26} & D_{25} & D_{23} \\ D_{56} & D_{55} & D_{35} \end{bmatrix} \nabla u_3 \\ \begin{bmatrix} D_{16} & D_{46} & D_{66} \\ D_{15} & D_{45} & D_{56} \\ D_{13} & D_{34} & D_{36} \end{bmatrix} \nabla u_1 + \begin{bmatrix} D_{46} & D_{26} & D_{56} \\ D_{45} & D_{25} & D_{55} \\ D_{34} & D_{23} & D_{35} \end{bmatrix} \nabla u_2 + \begin{bmatrix} D_{66} & D_{56} & D_{36} \\ D_{56} & D_{55} & D_{35} \\ D_{36} & D_{35} & D_{33} \end{bmatrix} \nabla u_3 \end{bmatrix} = \begin{bmatrix} \alpha_{11} \\ \alpha_{12} \\ \alpha_{13} \\ \alpha_{21} \\ \alpha_{22} \\ \alpha_{23} \\ \alpha_{31} \\ \alpha_{32} \\ \alpha_{33} \end{bmatrix}. \quad (17)$$

For example, the component α_{11} reads

$$\begin{aligned} \alpha_{11} = & D_{11} \frac{\partial u_1}{\partial x_1} + D_{14} \frac{\partial u_1}{\partial x_2} + D_{16} \frac{\partial u_1}{\partial x_3} \\ & + D_{14} \frac{\partial u_2}{\partial x_1} + D_{12} \frac{\partial u_2}{\partial x_2} + D_{15} \frac{\partial u_2}{\partial x_3} \\ & + D_{16} \frac{\partial u_3}{\partial x_1} + D_{15} \frac{\partial u_3}{\partial x_2} + D_{13} \frac{\partial u_3}{\partial x_3}. \end{aligned} \quad (18)$$

The remaining components α_{ij} one can calculate similarly.

For the flux terms the divergence operator works on each row separately. The divergence of the conservative flux source reads

$$\nabla \cdot \boldsymbol{\alpha} = \nabla \cdot \begin{bmatrix} \alpha_{11} \\ \alpha_{12} \\ \alpha_{13} \\ \alpha_{21} \\ \alpha_{22} \\ \alpha_{23} \\ \alpha_{31} \\ \alpha_{32} \\ \alpha_{33} \end{bmatrix} = \begin{bmatrix} \nabla \cdot \begin{bmatrix} \alpha_{11} \\ \alpha_{12} \\ \alpha_{13} \end{bmatrix} \\ \nabla \cdot \begin{bmatrix} \alpha_{21} \\ \alpha_{22} \\ \alpha_{23} \end{bmatrix} \\ \nabla \cdot \begin{bmatrix} \alpha_{31} \\ \alpha_{32} \\ \alpha_{33} \end{bmatrix} \end{bmatrix} = \begin{bmatrix} \frac{\partial \alpha_{11}}{\partial x_1} + \frac{\partial \alpha_{12}}{\partial x_2} + \frac{\partial \alpha_{13}}{\partial x_3} \\ \frac{\partial \alpha_{21}}{\partial x_1} + \frac{\partial \alpha_{22}}{\partial x_2} + \frac{\partial \alpha_{23}}{\partial x_3} \\ \frac{\partial \alpha_{31}}{\partial x_1} + \frac{\partial \alpha_{32}}{\partial x_2} + \frac{\partial \alpha_{33}}{\partial x_3} \end{bmatrix}. \quad (19)$$

3.2 2D case

In consequence of the dimension reduction from 3D to 2D one has to modify eq. (12). In a two-dimensional world λ takes the following form

$$\lambda = \frac{E\nu}{(1+\nu)(1-\nu)}, \quad (20)$$

as Poisson's ratio fits the range $\nu \in (-1; 1)$. Expression for μ remains unchanged.

Obviously, in 2D, vectors have two components, instead of three, and one works with 2x2 matrices, instead of 3x3.

4. Numerical results

4.1 3D case

The object of our interest is a box in 3D, fixed at two parallel lateral surfaces (see Fig. 1C) and loaded at front and back (see Fig. 1A) parallel opposite surfaces. The top and the bottom (parallel) walls are free (Fig. 1B). Four cases of box shape have been considered (see Table 2). Boxes were made either of classic (Poisson's ratio: 0 and +0.3) or auxetic material (Poisson's ratio: -0.999999 and -0.7), isotropic and elastic loaded. In all cases extensions of the box were considered. Numerical data necessary to perform the calculations are collected in Table 1.

All the calculations have been done by Comsol Multiphysics code (Comsol, 2007). Finite element calculations are made with second-order tetrahedral Lagrange elements with mesh statistics collected in Table 2.

Firstly, static Navier's equation is analyzed in this chapter. The following boundary conditions are assumed:

- Loaded boundary - for $x = 0$: $\sigma \cdot \mathbf{n} = -\mathbf{P}$, $\mathbf{P} \neq \mathbf{0}$,
- Loaded boundary - for $x = L$: $\sigma \cdot \mathbf{n} = \mathbf{P}$,
- Fixed boundary - for $y = 0$ and $y = d$: $\mathbf{u} = \mathbf{0}$,
- Free boundary - for $z = 0$ and $z = h$,

where \mathbf{n} is the normal unit vector to boundary.

There is no initial stress and strain in the considered boxes.

Results concerning the simulations of the stretched boxes are shown in Fig 3-12. An anomalous feature of the box deformation for negative Poisson's ratio values compared to classical positive values has been observed. At extremely negative Poisson's ratios the displacement vector has components which are anti-parallel to the direction of loading. This feature is present for all considered boxes with different ratios of height to depth.

Quantity	Symbol	Unit	Value
Density	ρ	kg/m^3	7850
Young's modulus	E	Pa	$2.1 \cdot 10^{11}$
Poisson's ratio	ν	-	-0.999999, -0.7, 0, 0.3
Pressure – force per area (stretch)	$ \mathbf{P} $	N/m^2	10^4

Table 1. Numerical data

Quantity	Symbol	Unit	Box-1	Box-2	Box-3	Box-4
Height (z-direction)	h	m	2	1	0.5	0.005
Width (x-direction)	L	m	1	1	1	1
Depth (y-direction)	d	m	1	1	1	1
Number of mesh points	-	-	27642	12186	6096	139386
Number of elements	-	-	143682	61126	29605	523183
Number of DOF	-	-	607137	262377	129219	2675307

Table 2. Dimensions of boxes and mesh statistics

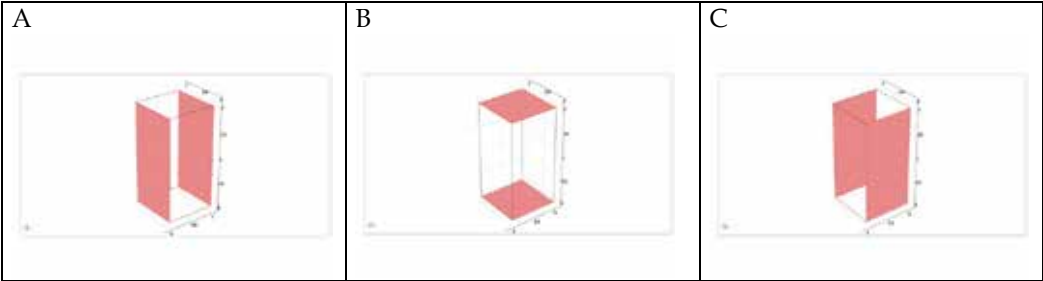


Fig. 1. Boundary conditions on box surfaces: A) loaded, B) free, C) fixed

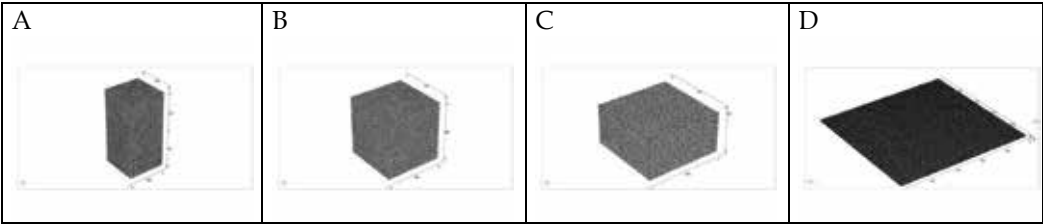
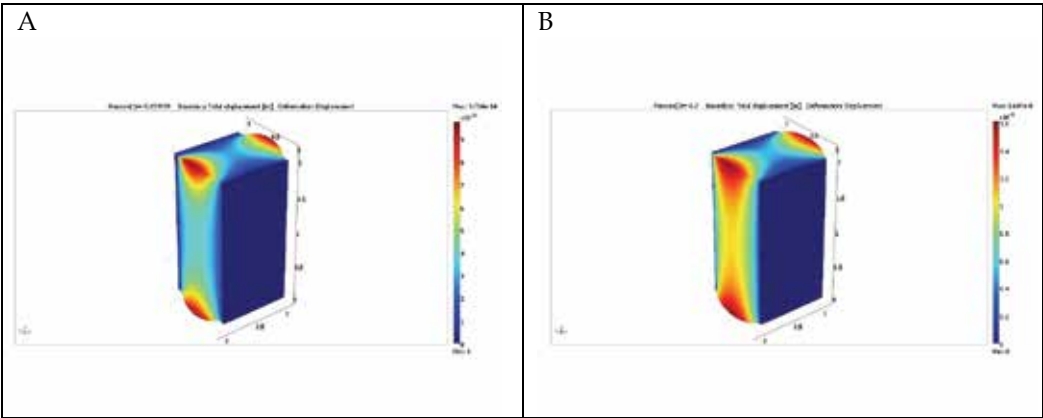


Fig. 2. Shapes of meshed boxes: A) Box-1, B) Box-2, C) Box-3, D) Box-4



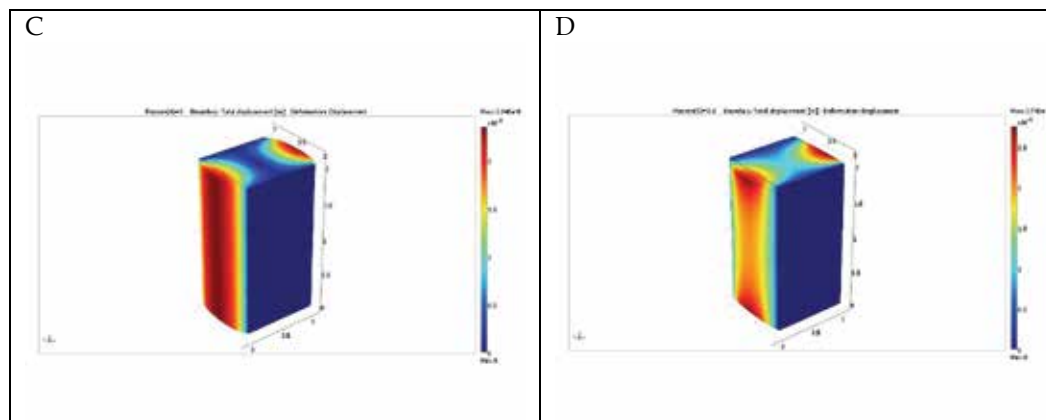


Fig. 3. Total displacement of the box-1 for different Poisson's ratios ν : A) $\nu = -0.999999$, B) $\nu = -0.7$, C) $\nu = 0$, D) $\nu = 0.3$. The initial shape of the plate is marked by a thin continuous line

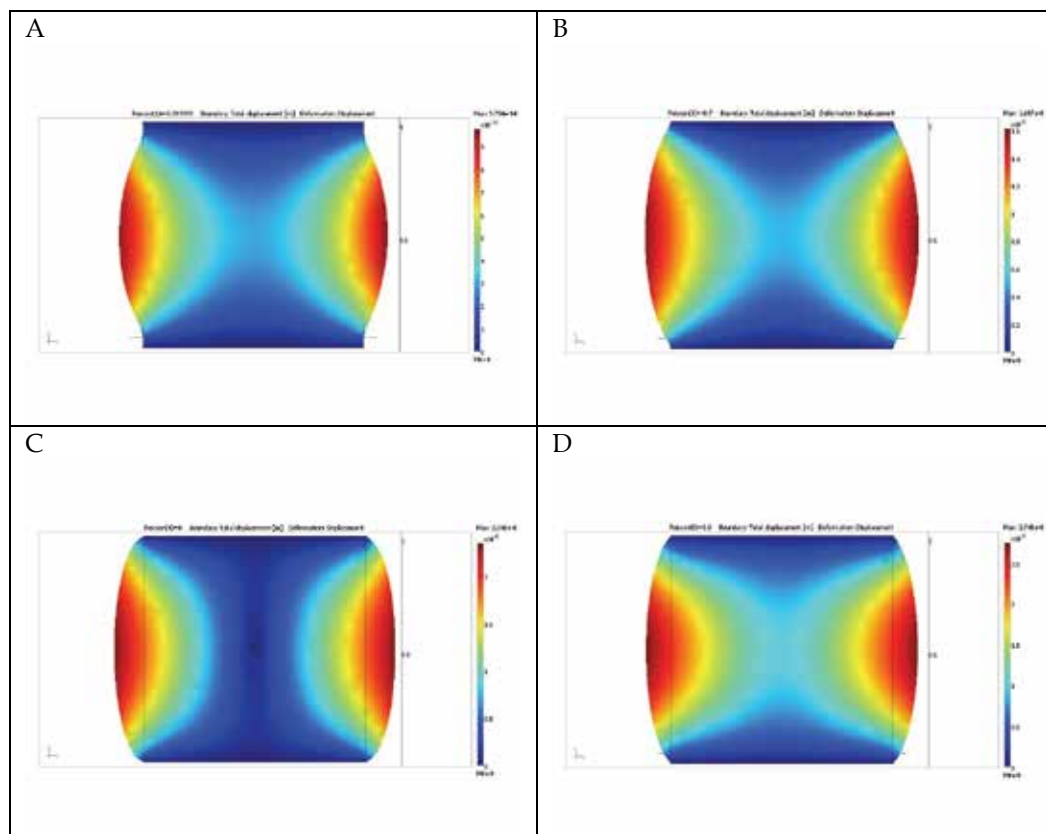


Fig. 4. XY-view of total displacement of the box-1 for different Poisson's ratios ν : A) $\nu = -0.999999$, B) $\nu = -0.7$, C) $\nu = 0$, D) $\nu = 0.3$

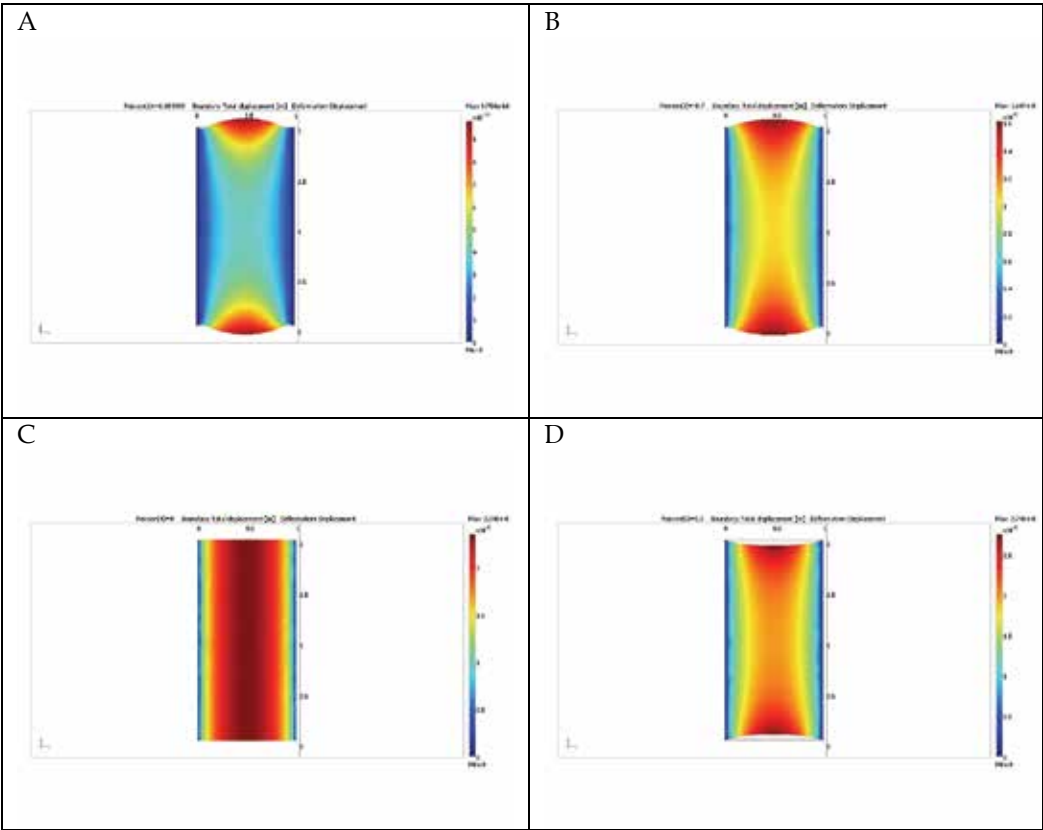
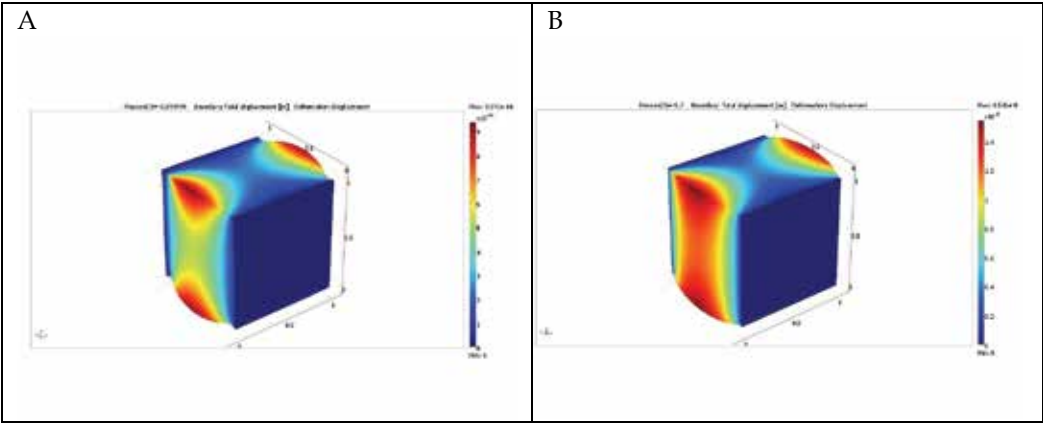


Fig. 5. YZ-view of total displacement of the box-1 for different Poisson's ratios ν : A) $\nu = -0.999999$, B) $\nu = -0.7$, C) $\nu = 0$, D) $\nu = 0.3$



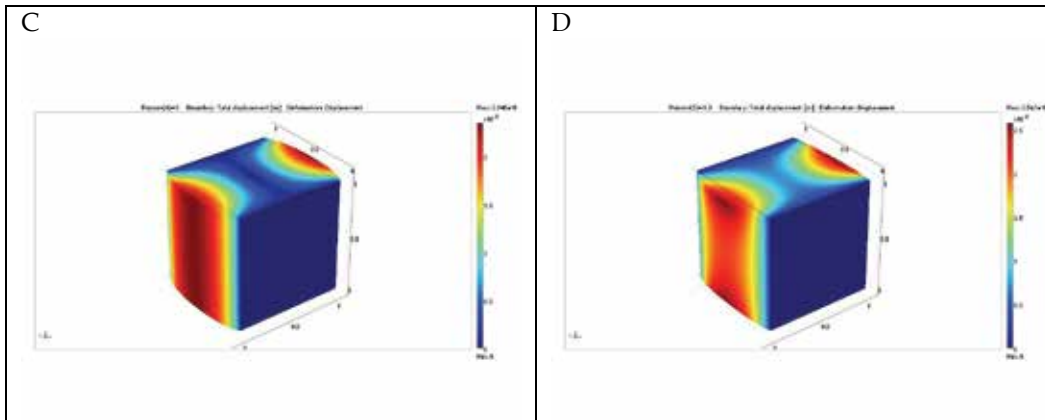


Fig. 6. Total displacement of the box-2 for different Poisson's ratios ν : A) $\nu = -0.999999$, B) $\nu = -0.7$, C) $\nu = 0$, D) $\nu = 0.3$

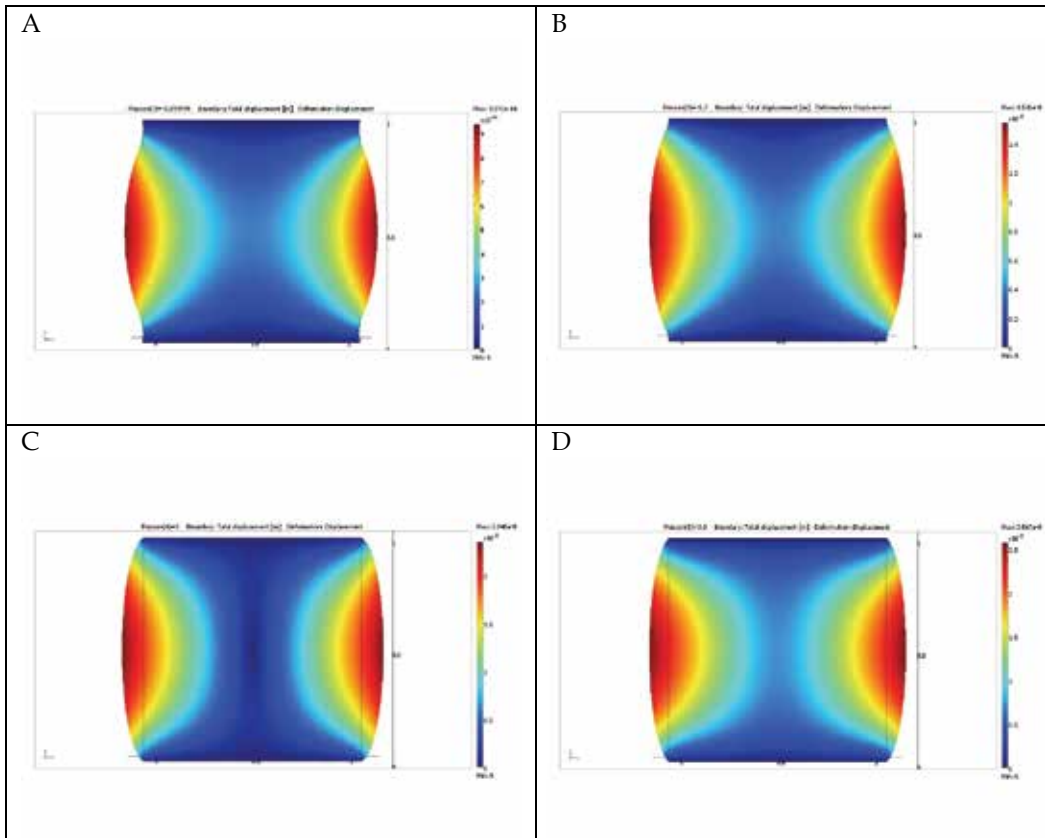


Fig. 7. XY-view of total displacement of the box-2 for different Poisson's ratios ν : A) $\nu = -0.999999$, B) $\nu = -0.7$, C) $\nu = 0$, D) $\nu = 0.3$

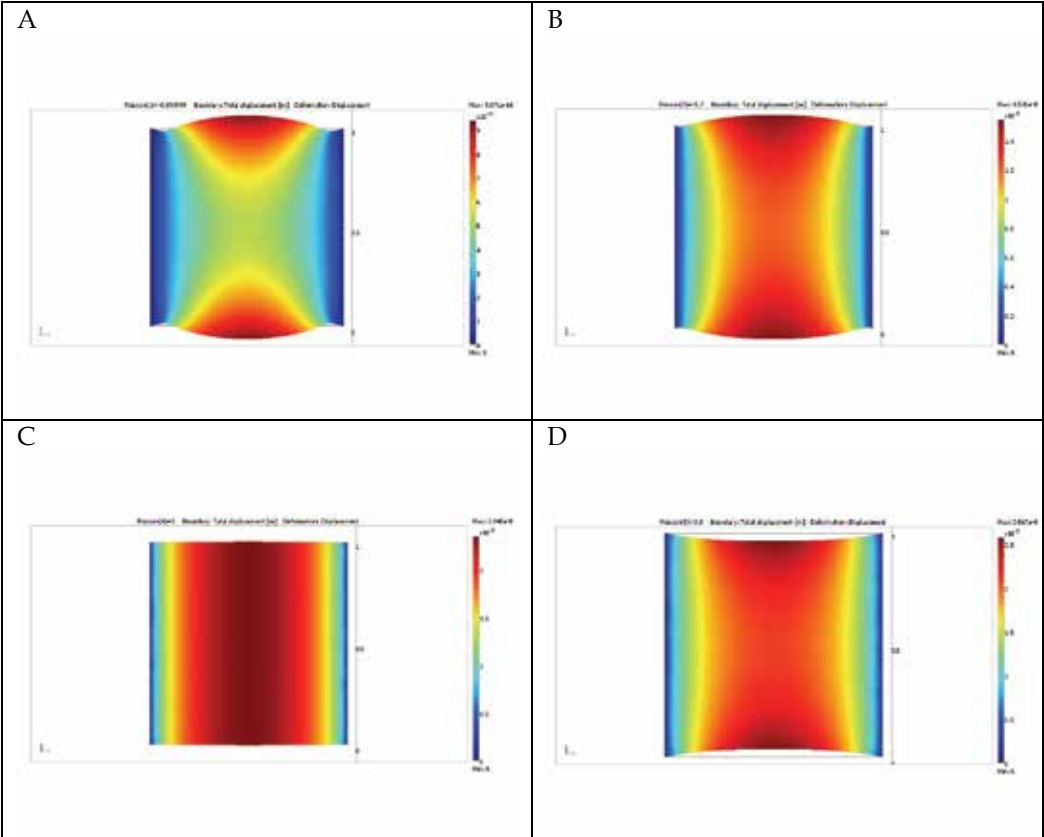
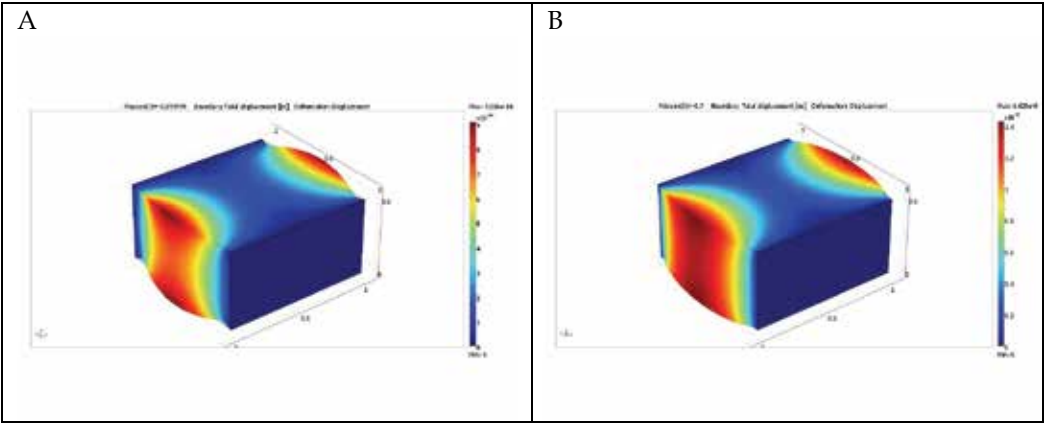


Fig. 8. YZ-view of total displacement of the box-2 for different Poisson's ratios ν : A) $\nu = -0.999999$, B) $\nu = -0.7$, C) $\nu = 0$, D) $\nu = 0.3$



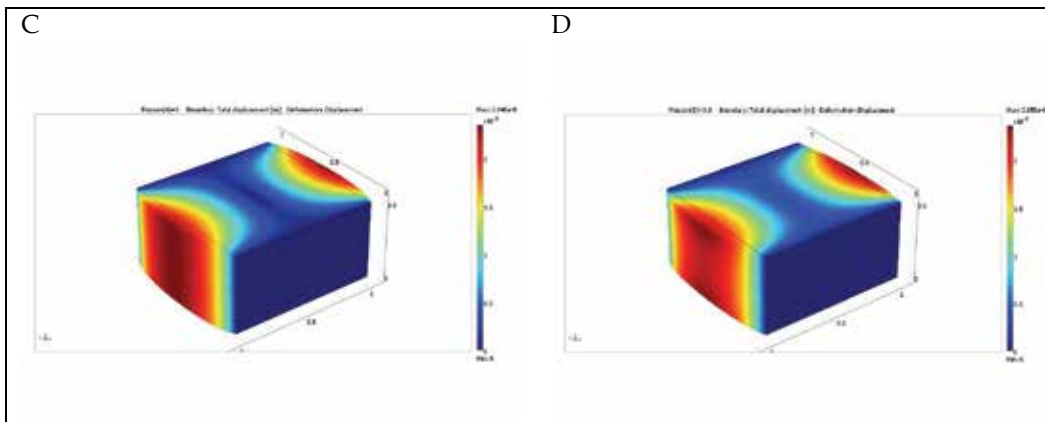


Fig. 9. Total displacement of the box-3 for different Poisson's ratios ν : A) $\nu = -0.999999$, B) $\nu = -0.7$, C) $\nu = 0$, D) $\nu = 0.3$

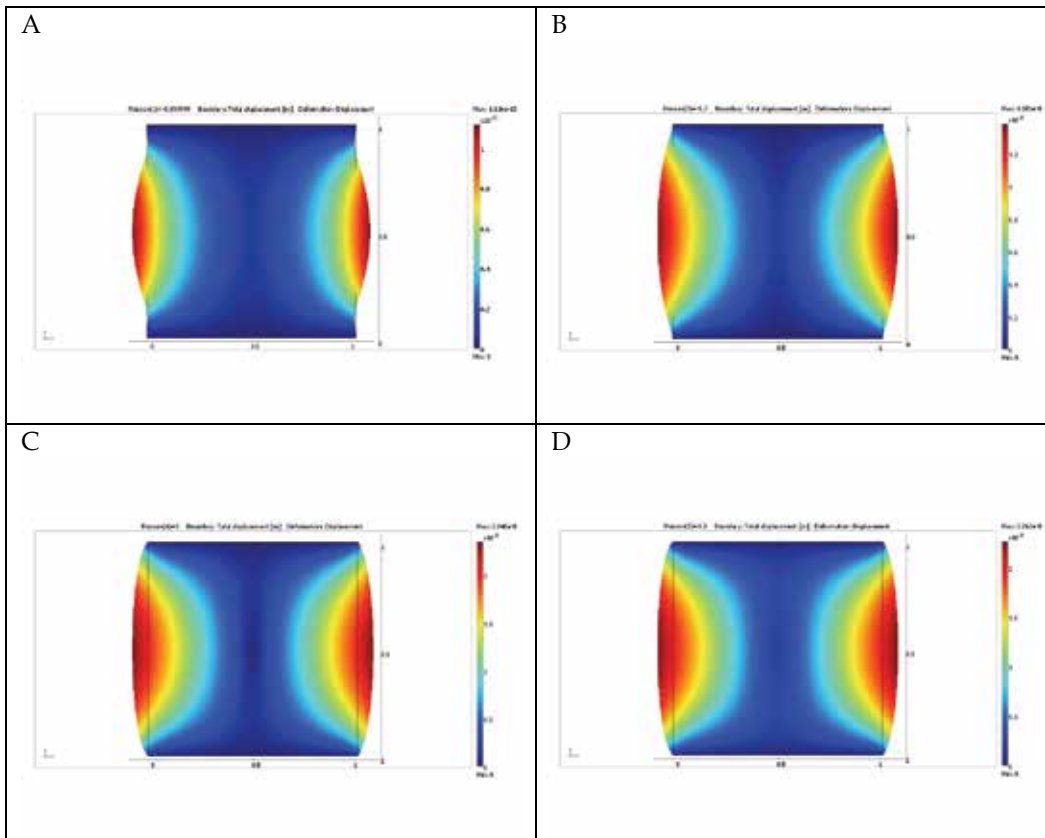


Fig. 10. XY-view of total displacement of the box-3 for different Poisson's ratios ν : A) $\nu = -0.999999$, B) $\nu = -0.7$, C) $\nu = 0$, D) $\nu = 0.3$

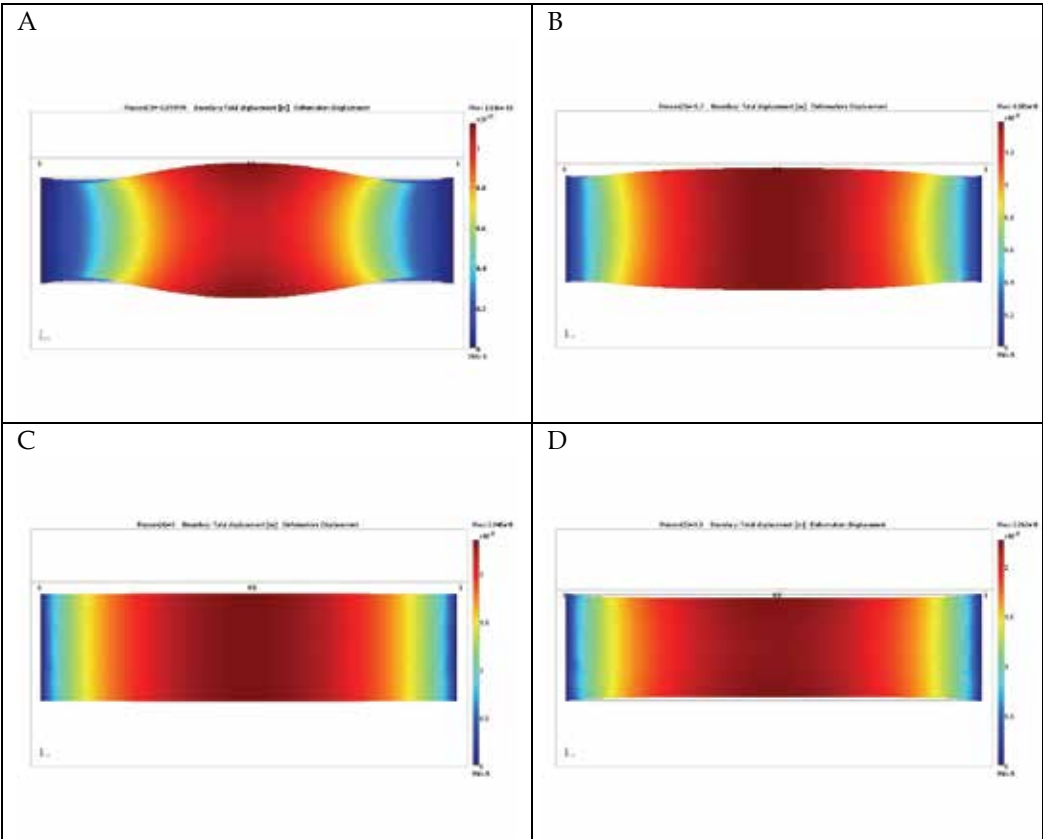


Fig. 11. YZ-view of total displacement of the box-3 for different Poisson's ratios ν : A) $\nu = -0.999999$, B) $\nu = -0.7$, C) $\nu = 0$, D) $\nu = 0.3$

The table below (Table 3) shows maximum value of the deformation on the surface to which the load is applied and average value of the deformation for different Poisson's ratios. One can see that with growing height, h , of the box for given Poisson's ratio and pressure the maximum deformation increases and the average deformation decreases. For a given shape of the box the maximum deformation increases with increasing Poisson's ratio.

	$\nu = -0.999999$		$\nu = -0.7$		$\nu = 0$		$\nu = 0.3$	
	Max.	Avg.	Max.	Avg.	Max.	Avg.	Max.	Avg.
Box-1	9.754e-14	3.557e-14	1.607e-8	8.585e-9	2.348e-8	1.772e-8	2.743e-8	1.635e-8
Box-2	9.371e-14	4.272e-14	1.536e-8	9.360e-9	2.348e-8	1.783e-8	2.567e-8	1.707e-8
Box-3	9.086e-14	4.898e-14	1.425e-8	9.729e-9	2.348e-8	1.783e-8	2.355e-8	1.713e-8

Table 3. Maximum and average total displacements of loaded boundaries of boxes

In Fig. 12 the deformation of the thinnest of the studied boxes is shown for the lowest Poisson's ratio and the finest mesh studied. Some oscillations of the loaded surface can be seen there. This new phenomenon, shown in more detail in Fig. 13, will be even better seen in the 2D case discussed in the next subsection.

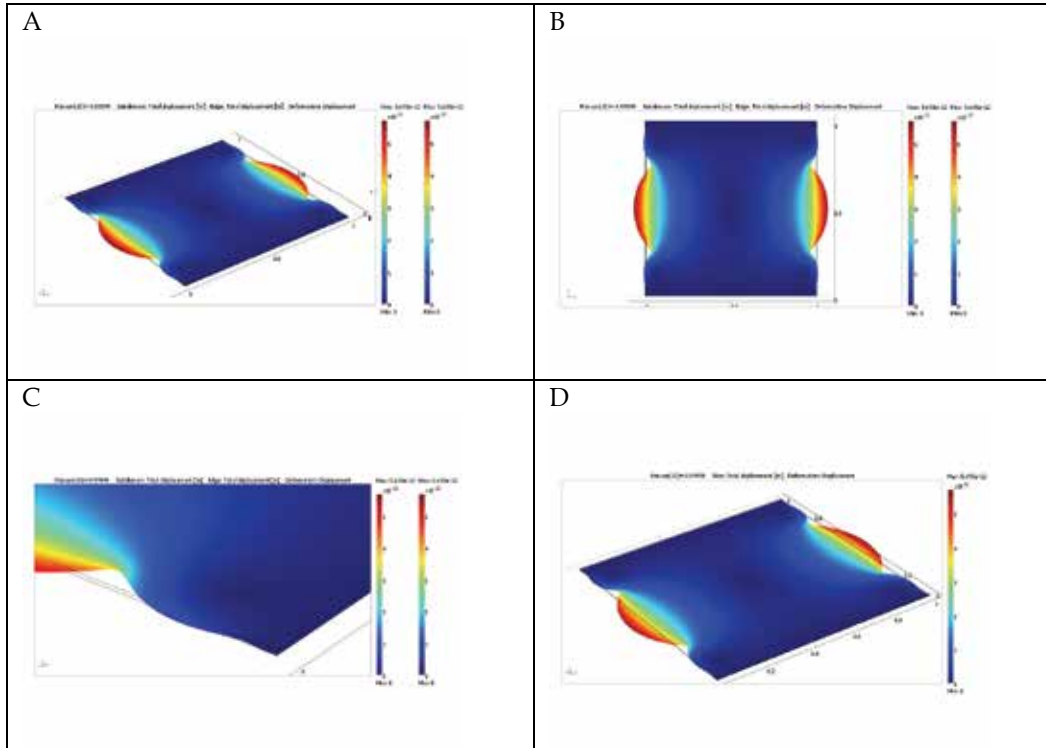


Fig. 12. Total displacement of the box-4 with different Poisson's ratio $\nu=-0.999999$: A) XYZ-view, B) XY-view, C) corner view, D) slice at height $h/2$

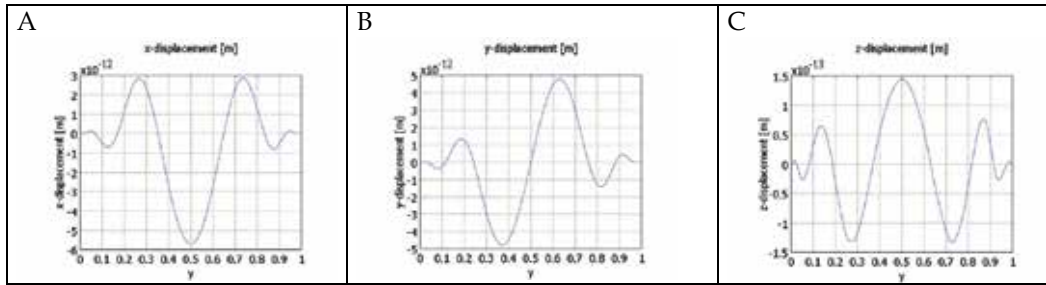


Fig. 13. Displacement of loaded edge (contact: loaded with free boundary) of box-4 with different Poisson's ratio $\nu=-0.999999$: A) x-displacement, B) y-displacement, C) z-displacement

4.2 2D case

The object of our interest is a 2D linear elastic continuum, given through Young modulus E and Poisson's ratio ν . The domain is simply a unit square and is subjected to mixed boundary conditions. First type is kinematic type, often called Dirichlet type and the other is Neumann (or natural) type boundary condition responsible for the traction. The boundary conditions of the system (see Fig. 14) are the following:

- Loaded boundary $\Gamma_R : \boldsymbol{\sigma} \cdot \mathbf{n} = \mathbf{P}$,

- Loaded boundary $\Gamma_L : \boldsymbol{\sigma} \cdot \mathbf{n} = -\mathbf{P}$,
- Fixed boundaries Γ_T and $\Gamma_B : \mathbf{u} = \mathbf{0}$.

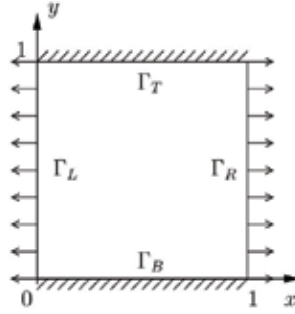


Fig. 14. Geometry of the system. Arrows indicate the uniform stretching force applied. Oblique lines indicate the fixed edges

To solve the problem, functions of the first order were taken as test and basis functions. It means that first order triangle Lagrange finite elements were taken. As the accuracy of finite element method depends on density of the mesh as well as on the degree of interpolating polynomial, the considered model was simulated with various mesh densities. Some calculations with higher order polynomials were also conducted and showed very good agreement with results obtained for first degree polynomials for dense enough meshes.

The most limiting factor in FEM computation is the memory, needed to store and solve systems of algebraic equations. Here, the upper limit of mesh density due to the limit of 32 GB RAM available was 15 600 000 triangle elements of first order with 7 845 601 points, what equals $N=2800$ intervals at a square side. The constants defining material and simulation conditions are as follows: $E=2.1 \times 10^{11}$ [N/m], $|\mathbf{P}| = |\alpha_{xx}| = 10^4$ [N/m] and ν varying from 0.7 down to -0.999.

The point of our interest will be only one component of the displacement field revealing the effect of interest. Fig. 15 shows details of the x -component of the displacement field u as function of y being position of the point on edge Γ_R . The x -component of displacement field is denoted by u_x .

As long as ν is non-negative (0.7, 0.0) the sign of u_x is the same as the sign of acting force, so the system behaves in a common way. But when ν takes the negative values (-0.7, -0.999) it is clear to observe that the u_x on some regions has negative values, what means that the body moves in opposite direction to the acting force!

In the case of $\nu = -0.7$ there is only one region on the boundary where the counterintuitive behavior is observed, but as ν tends to its lower limit (-1) and takes value -0.999 it is easy to notice that depending on mesh density the number of isolated negative-valued u_x varies from one (in case of $N=250$) to at least two (for $N=2000$). Obviously lower Poisson's ratio values need finer meshes to precisely track the behavior near corners of the square.

One would expect that there exist some critical value of ν when the counterintuitive effect occurs. In terms of FEM simulation this value strongly depends on the density of the mesh. To estimate convergence of the critical ν as function of the mesh density values, the values of $\nu_c(N)$ were computed for certain N and plotted in Fig. 16. It can be seen that $\nu_c(N)$ is an increasing function of N and is convex as a function of N^{-1} . The representation of $\nu_c(N)$ as

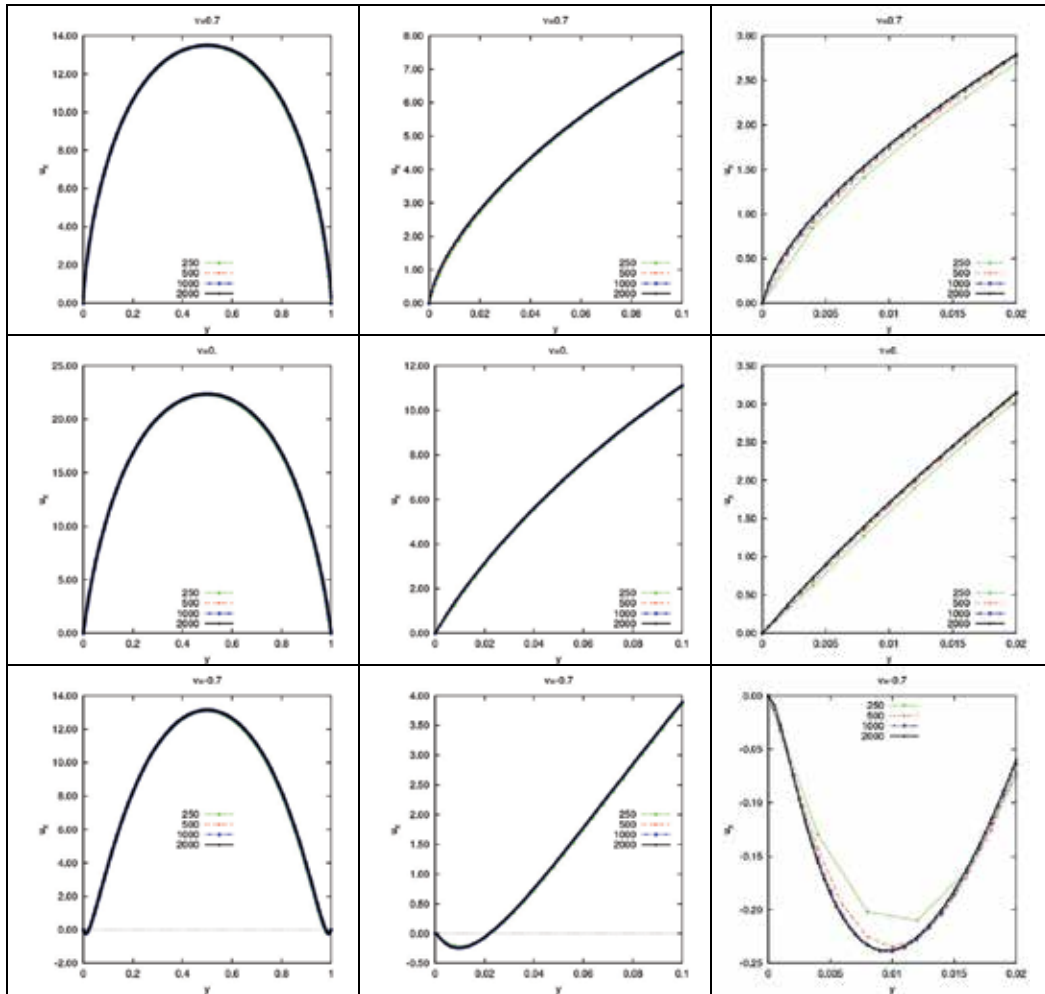
function N^{-1} takes the advantage of bringing infinity to zero and lets one to have a better view on the convergence. Looking at Fig. 16, one can state that

$$-.2 < \nu_c \equiv \lim_{N \rightarrow \infty} \nu_c(N). \quad (21)$$

Observing the behavior of u_x for the case of $\nu = -0.999$ one expects existence of more points where u_x changes sign between positive and negative values when ν tends to -1.

It should be also stated that for the lowest studied ν the closer to the corner, the worse the convergence is. To study models with extremely low ν values, very dense meshes are essential and higher order interpolating polynomials should be also considered.

To make clear that counterintuitive results are reliable some cases were checked using another FEM libraries - GETFEM++, FREEFEM and ABAQUS. These showed exactly the same unusual behaviour of the system.



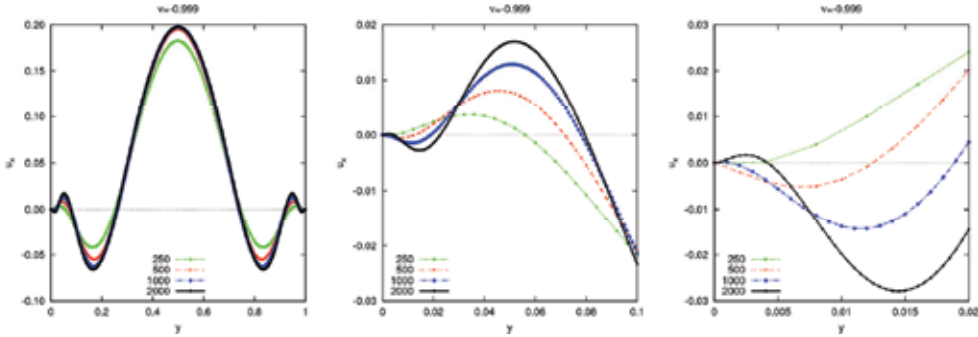


Fig. 15. u_x being the x -component of the displacement vector multiplied by 10^9 (with exception of the last figure in which the factor was 10^{10}) as a function of y for different Poisson's ratios. Figures show the dependence on the whole edge and its details near to the upper-right corner. The numbers in the legend describe the values of N for the meshes.

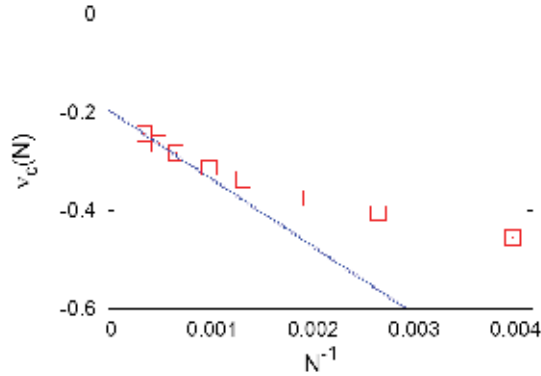


Fig. 16. The N^{-1} dependence of $v_c(N)$. The dotted line goes through the points corresponding to the two largest values of $N=2000, 2800$.

5. Conclusions

In this chapter, results of studies described in (Strek et al., 2008) and (Poźniak et al., 2010) have been reviewed and extended. By considering simple 2D and 3D examples, it has been shown that constrained auxetics can exhibit (locally) negative compliance. This unusual effect – material moving in the direction opposite to the force acting – is observed in 2D near the corners of the deformed square whereas in 3D it can be seen near the edges of the fixed walls of the box. In 2D case, the calculations performed prove that the maximum Poisson's ratio for which such a behaviour is found is not less than -0.2. In 3D, the corresponding Poisson's ratio is not less than -0.7.

The obtained results do not contradict the closure presented in (Poźniak et al., 2010) that the critical Poisson's ratio for which such a behaviour can be observed is zero, i.e. *any* auxetic material in 2D and 3D may show locally negative compliance. Work is in progress to prove that hypothesis.

In 3D case the authors have determined the maximum deformation and the average deformation of the box for different shapes of the box and for different values of the

Poisson's ratio. It has been found that for a given Poisson's ratio the maximum deformation grows and the average deformation decreases with increasing height of the box. For a given box shape the maximum deformation increases with increasing Poisson's ratio.

6. Acknowledgements

This work was partially supported by the grant NN202 261438 MNiSW. Part of the simulations was carried out at the Poznan Supercomputing and Networking Center (PCSS).

7. References

- Comsol (2004). Comsol Structures Mechanics Module, Femlab 3.1 User's Guide, Comsol AB.
- Comsol (2007). Comsol Multiphysics 3.4 User's Guide, Comsol AB.
- Evans K.E. (1991). Auxetic polymers – a new range of materials, *Endeavour* 15, pp. 170-174.
- Hartwig G. (1995). Support elements with extremely negative thermal expansion, *Cryogenics* Volume: 35, Issue: 11, *Nonmetallic Materials*, pp. 717-718.
- Hinton E., Owen D.R.J. (1979). *An Introduction to Finite Element Computations*, Pineridge, Swansea.
- Huebner K.H. (1975). *The Finite Element Method for Engineers*, Wiley, Toronto.
- Jaglinski T., Kochmann D., Stone D., Lakes R. S. (2007). Materials with viscoelastic stiffness greater than diamond, *Science* 315, pp. 620-622.
- Landau L. D., Lifshits E. M. (1986). *Theory of elasticity*, Pergamon Press, London.
- Lakes R.S. (1987). Foam structures with a negative Poisson's ratio, *Science*, 235 pp.1038-1040.
- Lakes R.S. (2001). Extreme damping in composite materials with a negative stiffness phase, *Phys. Rev. Lett.* 86, 2897-2900.
- Lakes R. S., Lee T., Bersie A., Wang Y. C. (2001). Extreme damping in composite materials with negative-stiffness inclusions, *Letters to Nature*, *Nature* 410, pp. 565-567.
- Lakes R., Wojciechowski, K. W. (2008). Negative compressibility, negative Poisson's ratio, and stability, *Phys. Stat. Solidi b*, 245, pp. 545-551.
- Logg A., Wells G. N. (2010). DOLFIN: Automated finite element computing, *ACM Transactions on Mathematical Software*, 37(2).
- Poźniak A.A., Kamiński H., Kędziora P., Maruszewski B., Stręk B., and Wojciechowski K.W. (2010). Anomalous deformation of constrained auxetic square, *Reviews on Advanced Materials Science*, 23, pp. 169-174.
- Remillat C., Scarpa F, Wojciechowski K.W. (2009). Preface, *Physica Status Solidi b*, 246, pp. 2007-2009; (see also references there).
- Ruppin R. (2000). Extinction properties of a sphere with negative permittivity and permeability, *Solid State Communications* 116, pp. 411-415.
- Sang Z.-F., Li Z.-Y. (2005). Effective negative refractive index of graded granular composites with metallic magnetic particles, *Physics Letters A*, Volume: 334, Issue: 5-6, January 24, pp. 422-428.
- Strek T., Maruszewski M., Narojczyk J., Wojciechowski K.W. (2008). Finite element analysis of auxetic plate deformation, *Journal of Non-Crystalline Solids*, 354, pp. 4475–4480.
- Zienkiewicz O. C., Taylor R. L. (2000). *The Finite Element Method*, Volume 1-3, (Fifth edition) Butterworth-Heinemann, Oxford.

Modelling of polymeric fibre-composites and finite element simulation of mechanical properties

Robert A. Shanks

CRC for Polymers, Applied Sciences, RMIT University

GPO Box 2476V Melbourne Victoria

Australia

robert.shanks@rmit.edu.au

Summary

Polymer composites are formed as random or anisotropic fibre dispersions, sheets with long fibre felted or woven fabrics and laminates of such sheets. A thermoset polymer in liquid form is mixed with the fibres and then cured. Thermoplastics must be intermingled with the fabric and the matrix phase united by melting of the thermoplastic and compaction of the composition. The composite sheets are thermoformable by virtue of the thermoplastic matrix phase. These composites present many compositional, structural and processing variables that contribute to properties. These composites are less uniform than typical thermoset resin composites where fibres are wet with a liquid resin that is then solidified by chain extension and crosslinking. Typical characterisation and mechanical performance tests are available to investigate and optimise the composites. Finite element analysis (FEA) enables a theoretical approach to understanding of the structure–property relationships and confirmation of interpretation of measured properties. A displacement field is suited to identifying and quantifying stress intensities in local regions of the composite to determine parameters critical to the performance of the composites. This chapter reviews the application of FEA to various composite types, stress situations and failure mechanisms. The FEA model design and simulation method are evaluated and compared.

1. Introduction

1.1 Types of composite

Composites based upon thermoplastics such as polypropylenes, polyethylenes, poly(ethylene terephthalate), poly(butylene terephthalate), various polyamides, polystyrene and its copolymers with butadiene and acrylonitrile, poly(vinyl chloride) thermoplastic polyurethanes, thermoplastic elastomers and biopolyesters such as poly(lactic acid), poly(hydroxybutyrate) and its copolymers with hydroxyvaleric acid. Thermosetting polymers include epoxy resins, unsaturated polyesters, vinyl esters, epoxy-acrylates, polyurethanes, polyisocyanurates, polybismaleimides, polysiloxanes, formaldehyde based

resins such as phenolic, melamine and urea, and many synthetic elastomers. Epoxy resins are the most commonly studied polymers using finite element methods and they are frequently cited as examples in this chapter.

Fibrous reinforcements include cellulosic fibres such as flax, hemp and others, glass fibres, carbon fibres, mineral fibres, synthetic fibres related to the matrix polymer. Table 1 shows Common fibre reinforcements with brief comments. Short fibres are chosen for direct addition to a polymer in extrusion or injection moulding. Common short fibres are chopped glass and carbon fibres, wood flour and other plant derived cellulosic fibres. Long fibres may be used in pultrusion, woven mats or felted (non-woven such as prepared by needle punching) mats. Mats require inclusion of polymer so that composites can be prepared by thermoforming since the structure of the mat must be maintained. Nano-fibres are subject to much recent investigation and they are of increasing commercial importance, such as carbon nano-tubes, microcrystalline cellulose or nano-cellulose, boron nitride and alumina whiskers. The composites discussed mainly include fibres, however other fillers such as glass spheres and particulate or platelet shapes are studied using FEA since they can be represented in two- or three-dimensional models of polymer composites. Polymer can be included as another fibre or as a powder, that are subsequently melted to form a uniform matrix between and around fibres, as a plastisol (polymer dispersed in plasticiser) that is thermally gelled to form a solid matrix, or as a chain-extendable or cross-linking pre-polymer. Sometimes a third component is included as an interphase surrounding the dispersed phase or when the layers of a laminate have different properties. A more complex model may include a density gradient within the matrix, often to distinguish surface layers from the interior.

Fibre	Composition	Description
Glass fibre	E-glass or S-glass	Chemically resistant, with size coating, need coupling agent
Carbon fibre	Graphite (formed by polymer pyrolysis)	High modulus and strength
Carbon nanotubes	Graphite cylinders	High modulus, strength, conductivity
Kevlar	Poly(phenylene terephthalamide)	High modulus / high strength types
Bast fibre (hemp, flax, rami)	Cellulose	Native or textile crystals, moisture sensitive, long fibres can be woven
Wood fibre or flour	Cellulose	As above, short fibres
Nano-cellulose	Partially hydrolysed cellulose	Crystalline perfection, high properties
Polymer fibres	Polypropylene, Polyamides, poly(ethylene terephthalate)	Moderate modulus and strength, used in special purpose composites
Mineral fibres	Rock wool, boron nitride, alumina	High performance

Table 1. Fibres for reinforcement of polymer composites

The thermoplastic composites under consideration contain long fibres that are structured by felting, typically using a needle-punch method, or woven. The composites are planar, with application as sheets or panelling, such as in automotive liner use or in building panels. The composite sheets are thermo-formable because the matrix is a thermoplastic. In automotive application they are formed into the many complex shapes required by the interior shape and contours. Fibre weave and/or entanglement will increase tensile strength across the plane of the composite. Fibres crossing the plane in either felted or woven composites can increase the interlayer strength thus preventing peeling of layers from sheets. The composite sheets can be laminates with one or two decorative or protective layers. In automotive use the lamination is typically an upholstered, leather-like or wood-like finish. The thermosetting composites may be readily formed into any shape, however the shape cannot be changed after formation, in contrast to thermoplastic composites.

1.2 Composite properties

Modulus that is orientation dependent and calculated as a volume fraction weighted mean of the matrix (polymer) and filler (fibre) in series or parallel (Equ. 1, where V = volume fraction, E = modulus, c = composite, f = fibre, m = matrix). An efficiency factor (g) is usually included to account for interfacial interaction and variables such as voids, inefficient dispersion of fibre bundles and variations in fibre alignments, etc (Equ. 2). The modulus can be measured using a universal test instrument in tensile mode (parallel with fibre orientation) that emphasises the properties of the fibres. The modulus measured in shear mode emphasises the matrix-fibre interface, while in flexure mode a combination of matrix and fibre modulus is measured, with the upper bended surface in tension and the lower surface in compression (Fig. 1). The Von Mises stress contour diagram (Fig. 1, lower image) shows a stress maximum at the fixture end and a compressive stress maximum in the lower centre surface. The same beam in three-point bend mode emphasises the stress in the load region with compressive stress concentrated near the upper surface and tensile stress concentrated near the lower surface (Fig. 2). An important observation for composite design is that the centre of the beam cross-section is relatively stress free.

$$E_c = V_f E_f + V_m E_m = V_f E_f + (1 - V_f) E_m \quad (1)$$

$$E_c = g V_f E_f + V_m E_m = g V_f E_f + (1 - V_f) E_m \quad (2)$$

Where V_f can be calculated from the length (l) and diameter (d) of fibres and their longitudinal (L) and lateral (S) spacing in the composite.

$$V_f = \frac{\pi l d^2}{4 L S^2} \quad (3)$$

The Halpin, Tsai and, Kardos composite model (Equ. 4 and Equ. 5) is a refinement that contains a geometric fitting parameter, A , where $A = 2(l/d)$ for tensile configuration with E as the tensile modulus, and the aspect ratio (l/d) of length (l) and diameter (d). A numerical solution is obtained for Equ. 4 and Equ. 5 to model a composite modulus.

$$E_c = \frac{E_m(1 + ABV_f)}{1 - BV_f} \quad (4)$$

Where:

$$B = \frac{\frac{E_f}{E_m} - 1}{\frac{E_f}{E_m} + A} \quad (5)$$

Fibre composites do not display a yield strength, unlike the matrix thermoplastic, since fibre pull-out and fibre fracture occur intermittently until fracture. Progressive fragmentation of a composite is not well simulated by FEA since the composite structure must be changing to represent structural rupture. Simulation of modulus is made while the composite is still coherent. Measurement of composite mechanical properties provides an overall modulus without information on the contribution of components and variables within the structure. Three assumptions are made: (a). The fibre-matrix interface has perfect adhesion; (b). The fibre and matrix exhibit an elastic response to stress; (c). No axial load is transmitted through fibre ends.

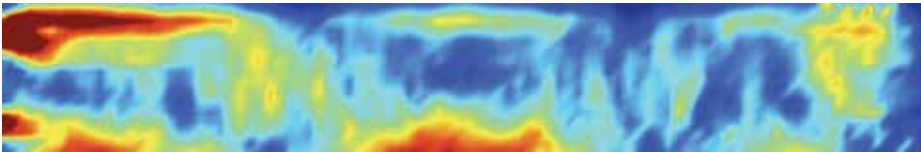
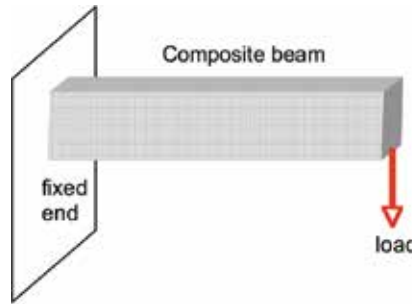


Fig. 1. Supported composite beam with deforming end load (upper) and a Von Mises stress contour plot showing the stress distribution (lower).

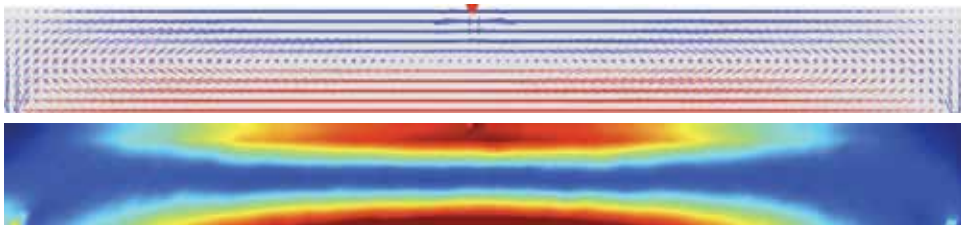


Fig. 2. Beam under three-point bend stress showing compressive stress near upper surface and compressive stress near lower surface, stress distribution (upper image) and Von Mises stress contours (lower image).

FEA requires a proposed model of the composite including the location of fibres within the matrix. The model can be validated by comparison with cross-sectional observation of the composite using optical or scanning electron microscopy. The model should correspond to the actual composite. Inclusion of interfacial interactions within the model is a hypothesis since the interfacial properties cannot be observed. A third phase of strongly adsorbed and immobilised polymer on the surface of fibres can be included to account for interactions (Fig. 3, upper image). The modulus and volume or thickness of this interphase must be estimated as between that of the matrix and fibres. The interphase will be diffused into the continuum matrix. A gradient of modulus can be used instead of a discrete interphase (Chen & Liu, 2001). The nature of the gradient could be linear or non-linear (such as exponential) from the fibre modulus to the matrix modulus. Complex geometry models can be formulated to include various specimen shapes, fibre orientations, woven bundles, and even random voids. The complexity of the model will determine the lowest scale that can be analysed while being representative of the whole composite (Vozkova, 2009). An example of a composite with an enlarged interphase is shown in Fig. 3 (lower image), where the simulation was an application of three-point bend stress. Complex stress fields are associated with the dispersed phase. This type of model is best studied by choice of a single particle or fibre when stresses associated with the interphase are to be considered (see later examples). The concept has been applied to the visco-plastic matrix behaviour of metal composites that reveal similar mechanics to polymer based composites (Shati, Esat, Bahai, 2001).

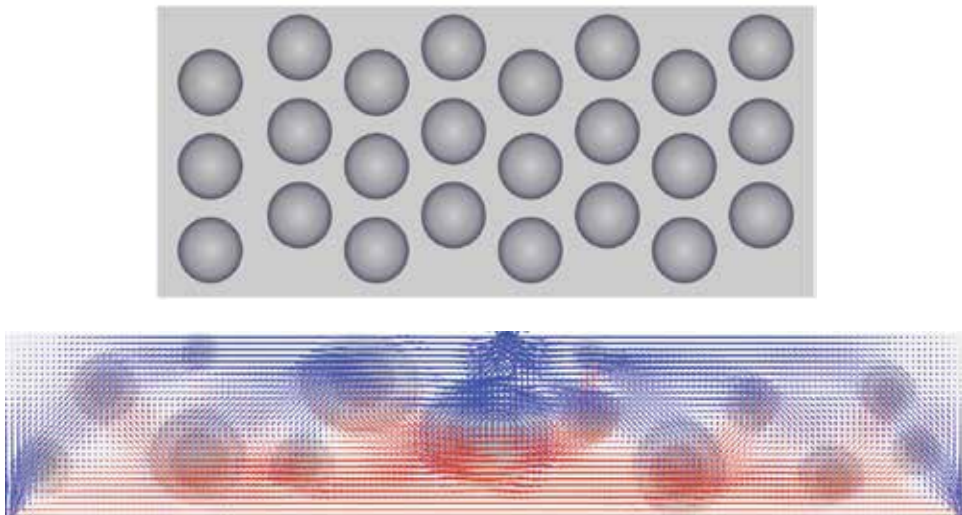


Fig. 3. Fibre composite cross-section schematic showing interphase (darkened) around each fibre (upper) and the lower image shows stress distribution in similar model under three-point bend stress.

The model is then tested by performing a simulation, that is disturbing the model with stress and monitoring the evolution of the strain over time, or applying a strain and monitoring the stress over time. The evolution of the model with time is computed at each node of the finite elements. Application of stress or strain requires that part of the model be fixed. Typically for a beam one end will be fixed and the external stress or strain applied at the other end, for tensile or flexural

(single cantilever bend) testing. Alternatively two ends can be fixed and the stress or strain applied in the centre, equivalent to a dual-cantilever bend. A three-point or four-point bend requires that the ends allow lateral movement or slippage, which is more complex to simulate in comparison to a real situation.

The hypothesis is that the model will behave the same as the real composite. If the hypothesis is correct then the behaviour of the composite will be well understood. If the hypothesis is incorrect then the model needs to be refined and further simulated until the data for the model and real composite converge. Refinement of the model may require inclusion of more variables to better represent the real situation. Unlike a model the real composite is likely to deviate from an ideal structure. Imperfections can be included as discrete entities such as voids, or as a general decrease in the model parameters analogous to an efficiency factor.

An analogy to an anisotropic long fibre composite is reflected in the model of a group of storage silos. Clusters of circular cylindrical structures are found in silos, chimneys and water towers. They are subjected to flexural stress by wind induced oscillations. The cylinders deform in similar fashion to a continuous fibre composite. Five rows of eight silos were modelled using ANSYS with the modulus, Poisson ratio and density as inputs using a harmonic two-node element. The model was validated from data measured in a high wind from ten accelerometers attached to the silo walls. The data revealed harmonic frequencies of the silos relative to the wind velocity (Dooms et al, 2006). The model is structurally similar to a polymer-fibre composite that can display a dimension dependent resonant frequency under modulated force mechanical testing.

1.3 Composite morphological structures investigated using FEA

Important considerations for the design of composites that are assessed using FEA are: The bulk mechanical properties of the composite are determined by the modulus of each phase and their respective volume fractions. Consolidation of the composite is complete and the composite density is a volume fraction average of the component densities. Where consolidation is incomplete this can be represented by a density gradient across the thickness of the composite sheet. Fibre concentration and fibre diameter determine the distribution of fibres since a fixed volume fraction of fibres may be due to few of large diameter or many with small diameter, with a varying total fibre-matrix surface area. The fibre-matrix interface may be implicitly defined such that dewetting or voiding can occur, or be omitted so that the interface will remain intact. Fibre orientation can be included in the model but this will require a larger mesh size to generate sufficient distribution. Fibres are generally included in parallel orientation in models constructed since a random distribution is difficult to model unless the model is large (Fig. 4). The surface versus bulk properties often vary in composites: there may be matrix rich or fibre rich surfaces, with density gradients from the surface to the bulk



Fig. 4. 3D model of a oriented long fibre composite

A 2D model similar to the fibre composite of Fig. 4 is shown in Fig. 5. The fibres are shown from an aligned view so in the 2D perspective they could be particles. Other similar composite models are shown later in this chapter where the fibres are aligned with the longitudinal axis of the beam. This view of fibre ends is best for observing the stress field throughout the composite and because stress transfer between matrix is better tested laterally for long fibres. In this case the stress is in three-point bend configuration pulling downwards from the right side, while the left side is fixed. The lower section of the beam is in compression, while the upper section is in tension. Some stress concentrations are visible between fibre ends. The regions of stress concentration are between fibres rather than in surface regions; compare Fig. 5 with included fibres with Figs 1 and 2 without fibres, where the stress concentrations emanate from along the surfaces. Stress concentrations between the fibres mean that fracture is likely to initiate at fibre-matrix surfaces or within the matrix between fibres. Strong interfacial bonding and a strengthened interphase will contribute most to a composite in this circumstance. The maximum stress concentrations, both tensile and compressive are associated with the fixed end of the beam, away from the location of application of the stress. Stresses in the compression zone along the lower part of the beam have concentrated where fibres ends are juxta positioned downwards to the left, even though the model was constructed with the aim of randomising the fibre-end positions and diameters. The Von Mises stress plot is not shown because the colour shading obscured the circular fibres end.

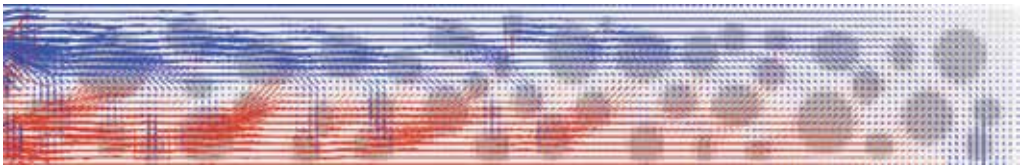


Fig. 5. Composite beam viewed lateral to fibres, fixed on left with bending stress acting down from the right.

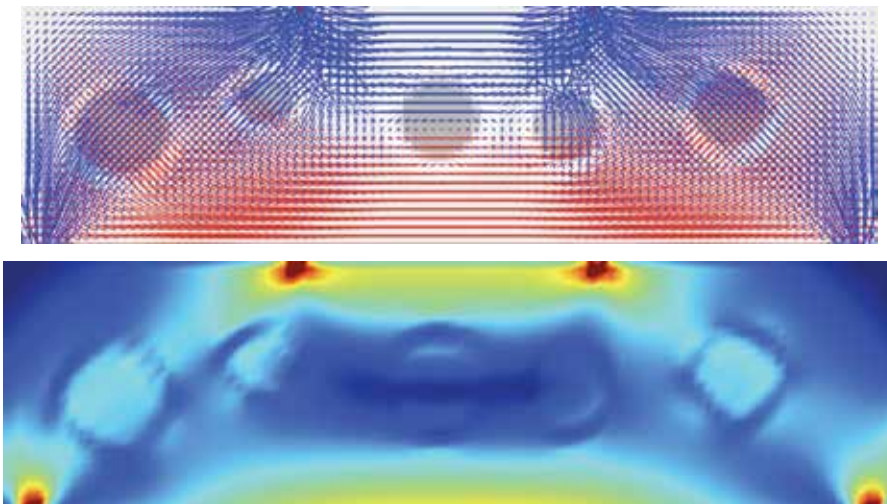


Fig. 6. A simplified model of a polymer with a soft interphase surrounding each hard fibre or particle under four-point stress, stress distribution diagram (upper) and Von Mises contour plot (lower).

A simplified model based on Fig. 5, with fewer hard inclusions surrounded by soft interphase, is shown in Fig. 6. The upper stress distribution diagram shows the compressive stresses passing from the application point through to the supported lower ends. The stress is not concentrated in the vicinity of the fibres, where the lighter shaded areas around each fibre is the soft phase. The lower central section of the beam is under a tensile stress situation. The lower Von Mises contour plot shows the stress reduced zones as darker areas around each fibre. A central stress free zones results from the four-point bend mode radiating stress away from the centre of the beam. Fig. 7 shows the same model as Fig. 6 with both ends fixed to represent a dual cantilever beam instead of a four-point bend configuration. The stress distribution in Fig. 7 is similar though less intense than that of Fig. 6 because some of the stress is redistributed as tensile stress emanating from the top corner fixtures. FEA depends upon the construction of a model to represent a material, such as a composite, and configuring the forces and constraints to best represent either a use situation or to reveal critical zones that are likely to cause performance problems.

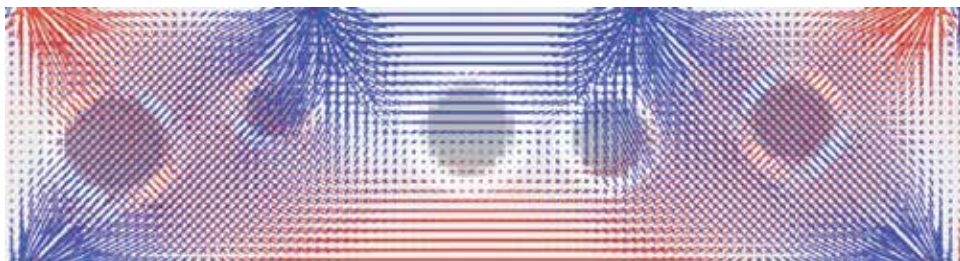


Fig. 7. A four-point stressed dual cantilever simulation of a polymer with a soft interphase surrounding each hard fibre or particle shown as a stress distribution diagram

The contribution of an interphase has been evaluated using a variant of FEA called the advanced boundary element method to model fibre-reinforced composites with consideration of varying thickness boundary layers (Chen and Liu, 2001). Fisher and Brinson (Fisher and Brinson, 2001) have used the Mori-Tanaka model and its extension by Benveniste to study a three-phase composite with either separately dispersed fibre and soft interphase material, that may include voids, or where the fibres were enveloped by the soft interphase material. The Mori-Tanaka model was more effective in predicting the matrix-dominated moduli of the composite. Physical aging was studied by using time and frequency shift factors for these thermorheologically complex materials, with frequency data being preferred. The 2D FEA results demonstrated the importance of the interphase in determining the overall shift rates of the composite. FEA was performed in 2D using a hexagonal array of inclusions with transverse hydrostatic and transverse shear superposition, to obtain the transverse Young's modulus and transverse shear complex moduli.

A composite interphase model was constructed including glass beads, an interphase and polycarbonate matrix, with perfect bonding at each interface and the beads symmetry packed in a cubic array. The Young modulus, stress concentration and stress distribution were simulated. The interphase increased fracture toughness at the expense of elastic modulus, with these observations becoming larger with increase in interphase thickness. A

suitable selection of filler content, interphase stiffness, thickness and Poisson ratio can reduce stress concentration with retention of composite modulus (Tsui et al, 2001). Stress has been imparted in two orthogonal directions simultaneously on a cruciform shaped specimen (Lamkanfi et al, 2010). The cruciform is applicable to real systems such as rotor blades. Strain was concentrated in the conjunction of the two stresses and concentrated in the corners rather than the centroid of the specimen. The numerical model was validated by experiments by means of a digital image correlation technique. Two- and three-dimensional models were evaluated. An orthogonally stressed cruciform has been simulated independently (Fig. 8) with a fixed point at the centre. The stress distribution (Fig. 8, upper) shows the stresses concentrated in the central region near the corner. This is better depicted in the Von Mises contour plot where the darker shading passes from the arms around the centroid.

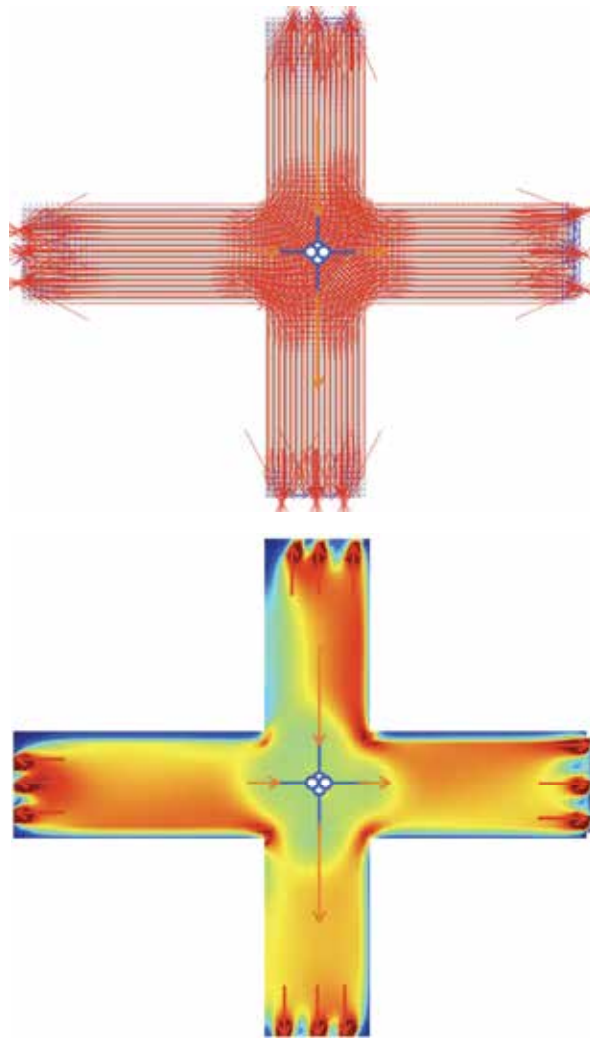


Fig. 8. Cruciform with orthogonal stress, stress distribution (upper), Von Mises contour (lower).

The planar shear bond test has been modelled with consideration of the moduli, bond layer thickness and loading conditions (Dehoff et al, 1995). The model consisted of cylindrical specimens bonded together, and asymmetric elements were used with harmonic stress. Large stress concentrations were confirmed at the bonded interface. The study is directed to assessing whether the shear bond test is useful for understanding the stress states that cause failure and if these stress states exist in clinical dental situations.

2. Computational methods

The finite element method (FEM) is used to solve partial differential equations using approximations that are iterated towards an optimised solution by numerical integration. The process of using FEM to solve practical problems is called FEA. Dedicated software provide the elements for most FEA computations. Examples shown in this chapter were prepared using the ForcePad (version 2.4.2, Division of Structural Mechanics, Lund University, Sweden) two-dimensional FEA program. ForcePad consists of three modes: a sketch mode where the model is prepared, a physics mode where forces and constraints are defined, and an action mode where the FEA calculations are made and the results visualised. The calculations shown in this chapter were performed with a mesh step setting of 6 (fine), vector constraint stiffness scale and force magnitude of 1-10 kN, weight of 1 kN, hard phase elastic modulus of 2 GN, stiffness scale factor of 1000, Young's modulus of 0.35 GN, relative thickness 0.1 and an element threshold of 0.02. Some of the pixelation in figures in this chapter arise from the mesh size used and not the resolution of the images. FEA software often consists of two components, model preparation, such as FEMAP, and the computational engine, such as Nastran.

Commercial software provides a visual and computational environment for application of FEA to problems ranging from simple components to complete complex systems. Typical commercial FEA software are: FEMAP and Nastran, Abaqus, ACTRAN, ANSYS, AcuSolve, ADINA, SLFFEA, LISA, ALGOR, Strand7, AutoFEA, LUSAS, MIDAS, FEAP, JANFEA, CADRE, FEMM, FesaWin, CALCULIX, COMSOL, Visual FEA, FEM-Design, DUNE, FEBio, ForcePAD, JFEM, OOFEM, TOCHNOG, MARC, OpenFEM, SJ MEPLA, FEAMAC, Opera 2D/3D, and many more specialised adaptations of the FEM. Some of the software rely on optional modules for particular and specialised applications. Details of the development and capability of each software package is available from suppliers and any comparison or discussion of performance is a personal choice and beyond the scope of this chapter.

General purpose software such as Mathematica, Matlab, MathCad, Igor, Maple can perform the necessary functions, and some ancillary plug-in programs are available to extend the basic programs. Users design modelling and computational applications within the environment of the software package with the aid of pre-written tools. Many custom written programs are used as described in the literature, which are written from basic principles using languages such as C++ and higher level object tools.

FEA experimental design: Description of the actual material, design of the model, validation of the model preliminary calculations, evolution of the simulation with stress, strain and time, analysis of the data. That is the response of each element to the displacement field associated with the environment of the material. Design of the model: finite element or mesh density and pattern, geometry of modelled composite: selected shape and region or finished article, geometry for application of stress or strain, parameters to be modelled (Fig. 9).

Boundary conditions need to be defined together with anchor or fixture points or surfaces. An element mesh needs to be designed and this mesh established a set of nodes in two-dimensions or three-dimensions. The elements can be triangular, tetrahedral, pyramidal, quadrilateral, hexahedral or other depending on the model and the type of problem. Choice of mesh resolution and geometry is important for each model and most advanced software packages provide a range of options and assistance to the modeller to capture smoothness, computational quality, simulation time, curvature and proximity features. Elastic and viscoplastic properties must be defined for the components. The links between elements are the mathematical expressions and their implementation that describe the theoretical aspects of the problem. The component material assembly can be designed as isotropic, orthotropic, anisotropic or laminate structure. The elements can be linear, such as spring, cable, truss, bar or beam. Surface elements include plane, plate, shear or membrane. Solid elements comprise various shapes that may be pre-stressed or face-stressed. The individual elements are allocated structural stiffness, that is a modulus. The modulus matrix is inverted and multiplied with the stress vector to simulate displacement vectors. Recent software includes wizards to assist the analyst in each step of the FEA process.: model design, validation, optimisation and simulation.

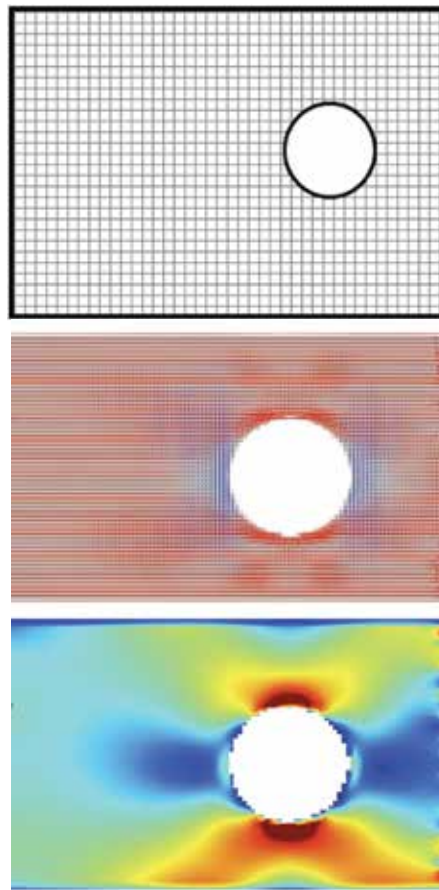


Fig. 9. FEM 2D model of a plaque with a hole for tensile or compression stressing

Validation of the model: Confirmation that the proposed model does portray reality. The model defined in this research used a 3D array of parallel fibres surrounded by matrix. The model did not consider looped entanglements of fibres, though the tensile strength in such cases must be equal to that of a fibre since a fibre must break to allow the matrix to extend. Inefficient wetting and voids are present in real composites, however they are not included in the models because the concentration, shape or distribution of voids is not known. Discrete voids can be replaced by a density gradient in composites that were partially consolidated.

Evolution of the simulation: Input of the material parameters. Digital incrementing of the independent variables. Computation of the dependent variables for each node, within each element, at each increment of stress or strain. The geometry of fixing the position of part/parts of the composite and the location of the displacement field are significant for the evolution of the model.

Analysis of the data. Stress-strain-time curves, analyses of variables between nodes, analysis of variables between composite models. Comparison with continuum composite models.

3. Examples of specific FEA composite simulations

3.1 Laminated composite structures

Damping of laminated carbon and glass fibre plastics is measured as the ratio of energy dissipated to the maximum strain energy stored per strain cycle. A finite damped element model including transverse shear has been used to measure and predict specific damping properties, mode shapes and natural frequencies of the composites (Lin et al, 1984).

Analysis of laminated composites has been performed where the failure mode was delamination of the layers. When a weaker ply fails first, stress will be distributed to the remaining plies and the process will continue giving a progressive failure of the laminate. The modulus of the layers was anisotropic so the strength of the composite depended on the relative orientation of each layer within the overall laminate. Symmetric and anti-symmetric ply laminates with different numbers of layers were formed (Pal & Ray, 2002).

Mechanical properties of a composite depends upon the geometry and aspect ratio of the fibres, while woven fibre composites are distinct from typical unidirectional long fibre composites. A woven fibre composite is usually prepared from multiple layers where the overall composite properties are the sum of the contribution layers. Flexural behaviour is a suitable way to characterise woven or unidirectional composites, since tensile force will be resisted by the modulus of only the continuous fibres. The layers of a woven composite can be divided into unit cells that are the smallest area in which the weave pattern is repeated (Fig. 10). A composite model can be formed by adding units cells laterally and in the thickness direction to form a structure as required for the model. This can be computationally performed by setting periodic boundary conditions. Computations for the property simulation used Abaqus. The FE model used curve beam elements to model warp and weft yarns of the unit cell. This simplified the model yet captured the actual morphology of the woven fibre composite, and hence allowed prediction of flexural modulus. Other weave types and replicating structures are suited to this efficient unit cell model with periodic boundary conditions (Soykasap, 2010).

A model of membrane layers of a laminated fibrous composite (analogous to Fig. 11) addressed the particular issue of interlaminar shear stresses concentrating along an edge region.

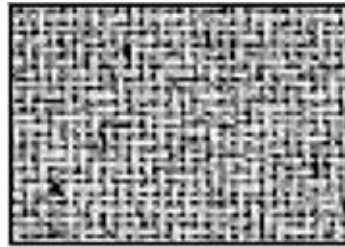


Fig. 10. Standard cross-ply woven fibre mat

The FEA simulation was limited to plane stress, where bending and warping of the laminate have not been considered. The simulated data were favourably compared with corresponding analytical data, with the FE technique presented capable of application in a range of laminate situations involving interlaminar shear stress (Isakson & Levy, 1971).

Through-width delamination of a laminate composite under a buckling compressive stress has been treated using a FEA parametric model. The stress leads to an instability related delamination. Comparison with measured lateral deflections showed that the analysis reflected actual specimen behaviour. Stress transfer was complex and not a simple function of applied stress or lateral deflection, with steep stress gradients at the delamination front suggesting a stress singularity. Hence strain energy release rates were much less responsive than the calculated stresses to mesh refinement. Correlation of the calculated strain energy release rates for mode I and II crack extension with actual delamination growth rates showed that delamination growth was dominated by mode I, even though mode II strain release rate was numerically larger (Whitcomb, 1981).

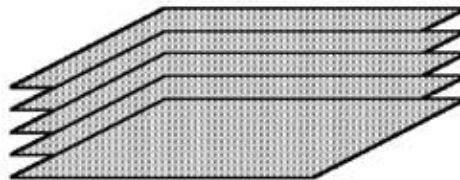


Fig. 11. A multi-layer woven fibre laminate composite

Orthogonal woven composites were studied using a 3D unit cell model using a custom approach followed by Strand6 FEA software. The woven fibre composite was for aerospace application, typically an epoxy-graphite fibre structure. The structural variables included geometric parameters, yarn volume fractions and engineering elastic constants. The in-plane modulus predicted using the unit cell model and a laminate model corresponded with experimental data from the literature (Tan, et al, 1998). Two-dimensional models were presented for the elastic analysis of a plain weave glass or carbon fabric epoxy resin laminate, with the aim of developing a simple though generalised model that will enable interpretation of the two-dimensional extent of the fabric, and the contribution of various fabric geometry parameters. This was intended for prediction of in-plane elastic modulus and selection of fabric geometry for any specific application. The fibre volume fraction

depended on weave geometry for a constant global fibre volume fraction. The elastic modulus predicted increased with fibre undulation, but reduced for 2D prediction and remained constant for a 1D parallel model with lamina thickness increase (Naik & Shembekar, 1992).

An example of a 2D to 3D global/local FEA of a laminated composite plate with a hole is presented. 2D models have shown limitations that can be addressed by using 3D models, though in many cases a 3D model is computationally inefficient for a particular task. An alternative is to perform a detailed 3D analysis on a local region of interest, followed by a more extensive, in geometry, stresses or time, simulation using a 2D model. In the example presented the local region of interest is the hole, around which the interlaminar stresses need to be examined in detail. There appeared to be a critical hole size in the laminate where the interlaminar stresses were maximised. A reduction in hole size for 3D laminates under compression suggested that if the hole contained a fastener, failure could result as the laminate fibres would be crushed into the rigid fastener. The degree of thickening of laminates under compression, around the hole, varied with lay-up, hole size and position about the hole (Muheim & Griffen, 1990).

Progressive failure of plain weave composites subjected to in-plane extension has been simulated using 3D FEA. Stress was parallel to the tow orientation and tow waviness was considered. The predicted strength decreased with tow waviness as progressive delamination occurred (Hyung, et al, 1991). The fibre shaping process in textile composites was evaluated from biaxial tests on cross shaped specimens. The influence of undulation variations in the weave and of interactions between the warp and weft were variables considered. The total energy was calculated as the sum of energies of each elementary cell that can repeat to form the entire composite. The weave undulations and interactions contributed to decline of the mechanical behaviour of the fabric composites (Boisse, et al, 1997).

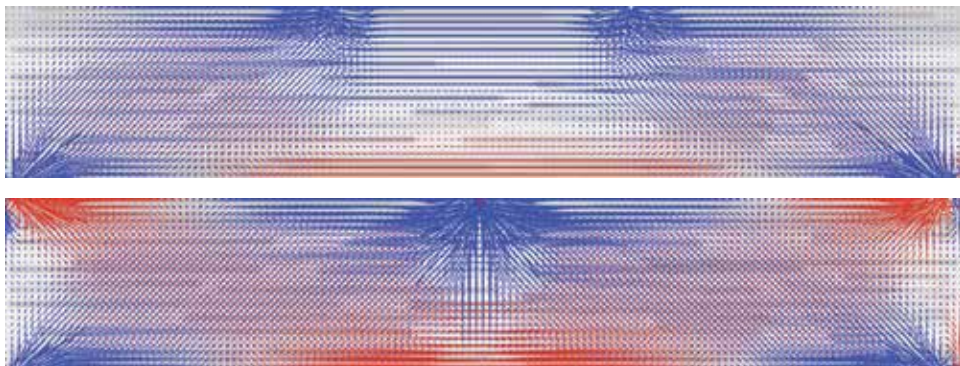


Fig. 12. A model of a laminated fibre-reinforced composite with stiff surface layers (darker shading) and a central softer layer under four-point bend (upper) and three-point dual cantilever (lower).

The laminated composite model shown in Fig. 12 has stiffer surface layer where resistance to stress is most important and a softer central layer where stresses from deformation modes is least. In practice the central layer could be a foam or material with limited consolidation. Each of the layers is reinforced with the same type of fibres. The simulated was performed

under four-point bending (upper image). A compressive stress concentration radiates from the application points to the two supported lower corners. A tensile stress is situated along the higher stiffness layer at the bottom. Stress transfer to fibres is visible throughout the laminated composite, but mainly as expected from regions of higher stress. A three-point dual cantilever simulation of the laminated composite is shown in Fig. 12 (lower). The stress concentrations are more intense with tensile stress regions emanating from the upper corner fixtures and in two localised regions along the bottom layer. Stresses are greater in the stiffer matrix top and bottom layers and stresses are transferred to fibres throughout, though with increased intensity in the higher stress regions of the upper and lower layers of the laminate.

3.2 Anisotropy of thermoplastic composites

Orientations occur during processing operations such as extrusion and injection moulding. Fibrous and platelet fillers are orientated along the extrusion direction or along the more complex flow lines in injection moulds. This directional geometry of the fibres is normally welcome as maximum modulus and strength are required in the machine direction. These orientations simplify the preparation of finite element models and the simulation process. An anisotropic fibre composite is shown in Fig. 13 (left) with the stress applied parallel to the fibres. The stress concentrations are low and concentration in regions where the fibres were made shorter. When the stress was applied transverse to the fibres (right) the stress concentrations were much greater throughout since the lower modulus matrix carried more of the stress than the higher modulus fibres.

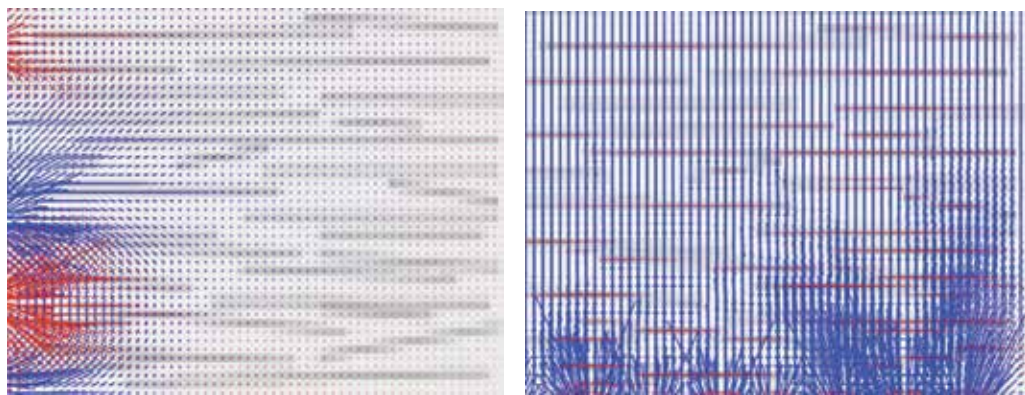


Fig. 13 Anisotropic fibre composite showing stress distribution when stress was applied parallel to the fibres (left) and transverse to the fibres (right).

A more complex consideration is anisotropy of the molecular or supramolecular structure within the matrix. Shear flow causes preferential orientation of molecules, crystallites and crystallite assemblies along the flow contours. Thus a polymer matrix will possess anisotropy of properties such as modulus, Poisson ratio, density fluctuations, residual stress, strength and ultimate properties such as toughness, break stress and strain. The anisotropy arising from these molecular considerations is difficult to accommodate in a model and subsequent simulation of the application of stress or strain. They are difficult to consider when interpreting actual performance properties of polymer composites since any detailed

molecular structure or morphological structure is difficult to measure and usually unknown to the tester and materials designer. The microstructure of the composite is known to be the determinant factor for the properties of fibre reinforced composites (Maligno, Warrior, Long, 2008).

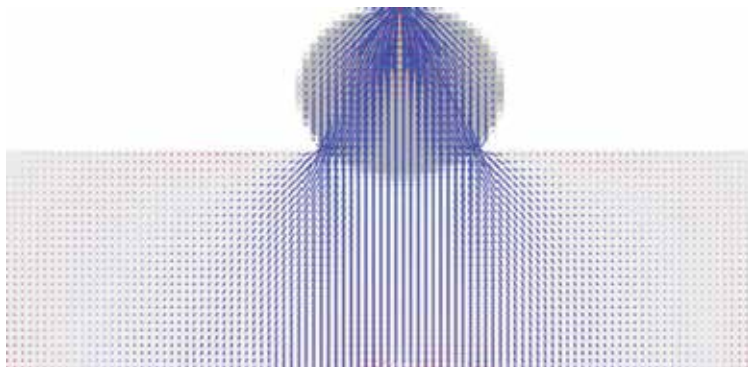


Fig. 14. Stress distribution diagram for indentation of a soft material by a hard round object.

A three-dimensional finite element model was used to evaluate indentation of polypropylene–cellulose fibre composite to determine the elastic modulus and hardness. FEA was conducted using a rigid flat cylindrical disk with a radius of $1\text{ }\mu\text{m}$ with Abaqus. Indentation to 50 or $100\text{ }\mu\text{m}$ depth provided no difference in the unloading values. Analysis of the interphase region showed a $1\text{ }\mu\text{m}$ wide property transition zone, though this zone could not be isolated from the contributions of the adjacent polypropylene and cellulose fibre properties (Lee, et al, 2007). An analogous model is shown in Fig. 14 in 2D and the simulation was performed with stress applied to the hard round object while the soft material was supported along its base. The stress diffuses from the application point with concentrations at the edge where the round object meets the soft material. Indentation by such a blunt object is not as critical as when a crack has been induced and the crack tip provides a focal point for stress concentration.

3.3 Composites with the same material for fibre and matrix

All-polypropylene composites have been prepared from polypropylene fibres with a lower melting temperature polypropylene matrix that may consist of a random polypropylene copolymer, different polypropylene tacticity or a different crystalline allotrope of polypropylene. The different melting temperatures of the polypropylene fibres and matrix are required so that the matrix can melt and form a continuous phase around the fibres without changing the oriented crystalline morphology of the fibres (Houshyar & Shanks, 2003). All-polypropylene composites can be prepared from polypropylene film-fibre layers, felted interwoven fibres of different melting polypropylenes, fibres impregnated with lower melting temperature polypropylene powder, fibres impregnated with polypropylene solution (Houshyar et al, 2005). Adhesion between polypropylene matrix and polypropylene fibres was characterized using a micro-bond test inspired by a fibre pull-out technique. The results showed that adhesion was appreciably increased when polypropylene fibres were used instead of glass fibres in the matrix (Houshyar & Shanks, 2010).

Finite element analysis showed that fibre volume fraction and diameter controlled stress distribution in all-polypropylene composites. The stress concentration at the fibre-matrix interface was greater with decreased fibre volume fraction. Changes in fibre composition were present in high stress regions. Stress concentration at fibres decreased and interfacial shear stress were more severe when higher modulus fibres were included. The relation between matrix and fibre modulus was significant, along with the interfacial stress in decreasing premature interfacial failure and enhancing mechanical properties. The simulations revealed that low fibre volume fraction provided insufficient fibres to disperse the applied stress. Under such conditions the matrix yielded when the applied stress reached the matrix yield stress, resulting in increased fibre orthotropic stress. With high fibre volume fraction there was matrix depletion and stress transfer had diminished capacity (Houshyar et al, 2009).

3.4 Biomaterial composites

A compromise in the design of biomaterials is the need for soft tissue bonding materials in contrast to hard wear resistant materials particularly in restorative dental composites. An approach is to use low modulus adhesive linings to release stress due to contraction during cure. Alternatively polymer without filler or lightly filled can be used as a low modulus relatively thick bonding layer to reduce the gradient to a high modulus restorative material. The FEA study aimed to predict the adhesive lining thickness and required flexibility. The model used a tooth shape connected by spring elements, constructed with 3D CAD using digitised images of a scaled tooth plaster model, and exported to Pro-Engineer, while FEA used ANSYS. The FEA determined optimum adhesive layer thickness for maximum stress release, though the models showed that a thin flexible adhesive layer had the same efficiency as a thicker higher modulus adhesive layer (Ausiello, et al, 2002).

Polymer glass and carbon fibre reinforced dental posts were modelled to compare properties with those of gold alloy cast posts, using a natural tooth restoration as a reference. The models contain continuous, unidirectional fibres that were uniformly packed. The models were formed with reference to the actual geometry of an upper central incisor. The model meshes were created using Mentat software package and they were simulated using MARC solving code. Mechanical data was obtained by three-point bend testing and compared with the FE model results, under various loading conditions. The glass fibre exhibited lowest peak stresses because its modulus was similar to that of dentin, and thus similar to the natural tooth (Pegoretti, et al 2002).

The stress contributions and stress rates within posterior metal-free dental crowns made from new composite materials were investigated. An FEM of a first molar was constructed, and placed under load simulating maximum bite force and mastication force using FEA. The maximum stress concentrated about the loading point. When a load was applied horizontally it was found to be a critical factor determining failure (Nakamura, et al, 2001).

The fatigue and fracture characteristics of human bone subjected to long term dynamic loading may be due to gradual reduction of elastic modulus resulting from various strain environments. The FEM applied to these materials will establish information for the design of joint replacements. Femoral neck fractures were chosen as a situation in which to explore FEA using continuum damage mechanics applied to the fatigue behaviour. Minimal modulus degradation or accumulation of permanent strain was observed until near the end of their fatigue life, and the models were unable to predict the rapid deterioration in the last

stage of this fatigue life. Models subjected to high initial strain, low cycle fatigue showed continuous modulus degradation and permanent strain accumulation throughout the test (Dooms, et al, 2006).

A natural fibre reinforced composite was studied using a shadow moire method to find the whole-field deformation of a cantilever beam under static stress, and the results compared with FEA in the same configuration. The comparison revealed a non-zero stress gradient in the cantilever due to the clamping force (Lim, et al, 2003).

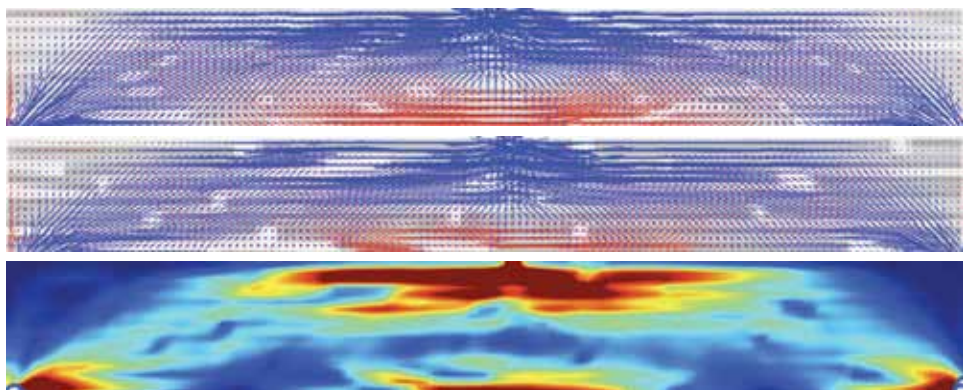


Fig. 15. A composite model with thin fibres (upper) and thick fibres at about the same volume ratio (centre) with Von Mises representation of the thick fibre composite (lower) after three-point bending.

The contribution of thin fibres (3 units) compared with thick fibres (10 units) in a composite with approximately the same volume ratio, is shown in Fig.15. The simulation was in three-point bend configuration with the same stress applied to each composite. Compressive stresses distribute from the upper centre application point to the supports at the bottom ends. The stress intensities are the same with thin and thick fibres. The lower centre region of both composites is under tension and the intensity of stresses appears greater where thin fibres were included. Horizontal streaking of the stresses represents stress transfer to the fibres. The model does not include a interfacial bonding that will be a factor in a real composite. This is probably why there is not a significant distinction between the composite models containing thin or thick fibres. Thin fibres present greater surface to the matrix and will thus provide improved bonding. The thick fibre composite is shown as a Von Mises contour plot (Fig. 15, lower) to emphasise the stress concentrations as mainly near the point of stress application, at the left and right supports and for the tension zone along the lower edge.

3.5 Nanocomposites

Polymer composites with carbon fibre were studied using a homogenisation method that included aspect ratio and concentration of the filler, in a controlled volume finite element method. The composites were assumed to have geometric periodicity at the local length scale and orientation of the carbon nanotubes was considered. The model showed thermal conductivities from numerical prediction were higher and rose more steeply with aspect ratio or volume fraction than those from experiment. Overall the calculated and

experimental values were similar (Sun & Youn, 2006). The mechanical properties of metals discontinuously reinforced with ultra-fine micro-structures have been studied by finite element methods (Huang, Bush, 1997).

Carbon nanotubes (CNT) are finding increasing application in composites because of their high modulus and electrical conduction. Carbon nanotubes are of molecular dimensions and consist of one large molecule so molecular mechanics should be an attractive modelling technique (Fig. 16). Mechanical properties can be better evaluated using FEA. Single-walled (SWCNT) and multi-walled (MWCNT) varieties are available. A CNT is a cylinder of graphitic carbon atoms bonded on the surface in a hexagonal array consistent with a space-frame structure. The hexagonal covalent bonding has been used as beam elements for a mesh for FEA using ANSYS software. The contributions of wall thickness, tube diameter, bonding chirality, on the elastic modulus of SWCNT were investigated. The elastic modulus decreased with wall thickness and increased with tube diameter (Tserpes & Papanikos, 2005). The space-frame model of CNT described merges the real-world scale functions of FEA with those of molecular mechanics at the atomistic scale, equating the modulus of the frame with the force field of the covalent bonds.

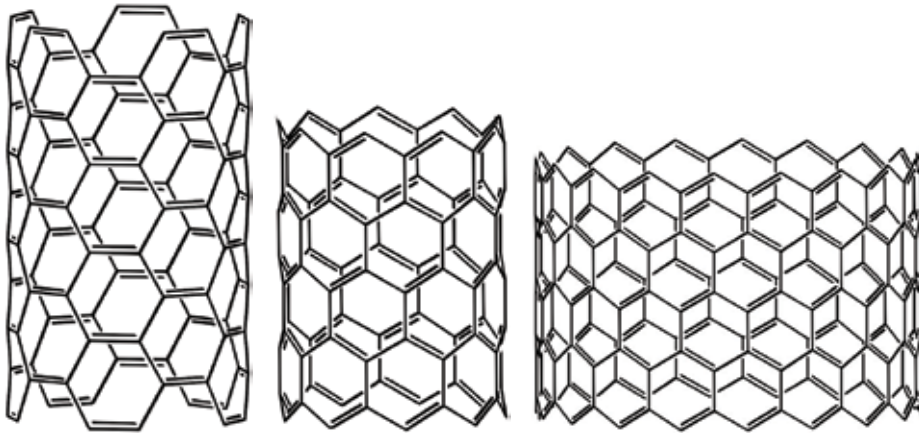


Fig. 16. Carbon nanotube structures, 5,5-armchair (left), 9,0-ZigZag (centre), 18,0-ZigZag (right)

Carbon nanotube reinforced polymers show significant enhancement over the polymer alone with regard to mechanical, electrical and thermal properties. The characteristic waviness of nanotubes embedded in a polymer was investigated in response to impact of their effective stiffness. A 3D FEM of a single infinitely long sinusoidal fibre within an infinite matrix was used to numerically compute a dilute strain concentration tensor. The modulus reduction was dependent on the ratio of sinusoidal wavelength to CNT diameter. As the wavelength ratio increased, the modulus of the composite with randomly oriented wavy CNT converged to that of straight CNT inclusions. The model developed can be used for other CNT-reinforced polymers, including different CNT configurations, viscoelastic response, low matrix-CNT bond strength, thermal and electrical conductivity (Fisher, et al, 2003).

3.6 Composites with particulate filler

A model has been prepared for a glass sphere filled epoxy for comparison with glass fibre fillers in the same polymer. The model considered spheres of equal diameter randomly distributed within an infinite matrix. The section for FEA was a cylinder containing a single sphere at its centre and LUSAS software was used for computations. The FEM was validated by agreement between predictions and experimental measurements that will lead to more efficient use of these materials under varying stress conditions (Guild & Young, 1989). Such a model is illustrated in Fig. 17 (left), while a 2D representation after simulation with compressive stress is shown as a stress distribution (centre) and Von Mises stress contour image (right).

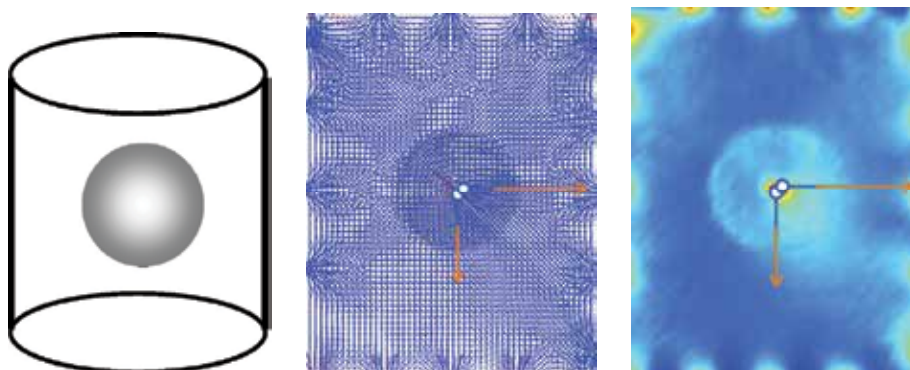


Fig. 17. Sphere in cylinder model of a glass spheres filled epoxy (3D view, left), after simulation under compressive stress showing stress distribution (centre) and Von Mises stress contours (right).

The arrows in Fig. 17 show the anchor points at the centre of the hard particle. An identical Von Mises stress contour plot was obtained when tensile expansive stress was applied.

3.7 Composites containing stress concentrations

Stress concentrations are prevalent in composites due to their formation from two dissimilar materials. Stresses may originate from differential shrinkage during cooling, crystallization of the polymer matrix, solvent evaporation, cure of reactive polymers and viscoelastic memory after polymer matrix flow. During matrix transformations the filler or fibre dispersed phase is typically inert. Stresses in laminates have been analysed by consideration of material non-linearity. Tensile stress was applied in a simulation and damage accumulated was calculated according to a property degradation model, until a failure criteria was reached. The model and simulation data was compared with experimental measurements on laminates containing a circular hole (Chang & Chang, 1987).

A composite design to avoid thermal shock and high strain rates is a functional graded material (FGM). FGM are designed to have a relatively graded variation of structure rather than a sudden change at interfaces that will produce stress concentrations that weaken interfaces. FGM is achieved by the composite preparation technique, such as formation of a continuously varying volume fraction of filler particles in the polymer matrix. FEA analysis of a FGM used Nastran with a plane stress elastostatic model and a symmetric three-point bend test simulation. The mesh discretization used consisted of rectangular eight-node

isoparametric elements. The Young modulus was varied within the model while maintaining a constant Poisson ratio. The FEA fracture analysis demonstrated that the FGM were better than a bimaterial counterpart (Butcher, et al, 1999).

Residual stress concentrations may be due to chemical reaction shrinkage during cure, temperature stress from non-equilibrium cooling particularly past the glass transition temperature, and humidity that will contribute swelling and contraction with variation in water absorption. A method of spring-in, with a right-angle model, was used to identify stress in curved components in composites with woven and chopped fibre reinforcement. The residual stress level in the angled region of the model was calculated and indicated the volume affected by the deviation from structural equilibrium (Wang, et al 2000).

Residual stresses remain in composite as a consequence of thermal relaxation kinetics upon cooling. Stresses are relieved over time by creep and stress relaxation and with annealing temperature treatment to change the thermal history. These considerations applied to many polymer systems, but analogous situations exist within inorganic composites. One study investigated aluminium oxide fibres in a metal alloy matrix of nickel-aluminium. A FEA model was simulated using Abaqus software. The evolution of residual stresses over time was interpreted in response to creep, plastic deformation and cooling/heating rate path dependency (Choo, et al, 2001).

3.8 Fibre-Matrix interfacial adhesion

A rigorous composite model needs to give attention to explicit bonding between the matrix and fillers, which are often fibres, but may be platelets, spheres or irregular particles. Bonded joints provide the connecting system whereby stress is transferred from the relatively low modulus matrix resin to the supporting high modulus fillers. A non-linear model was necessary to achieve accurate results relating a standard adhesion test experimental data with numerical data. Stresses concentrated at the interfaces at the longitudinal and transverse ends of an adhesive layer (Diaz, et al, 2010). Analyses and interpretation of adhesion strength needs to consider the viscoelastic nature of polymers, such that much energy in separating bonded interfaces is expended in the bulk polymer due to viscoelastic processes converting stress energy into heat energy resulting from molecular segmental motion friction. The relationship between stress transfer, the elastic modulus and the mechanical properties depicted by a stress-strain analysis have been simulated by finite element analysis (Kang, Gao, 2002).

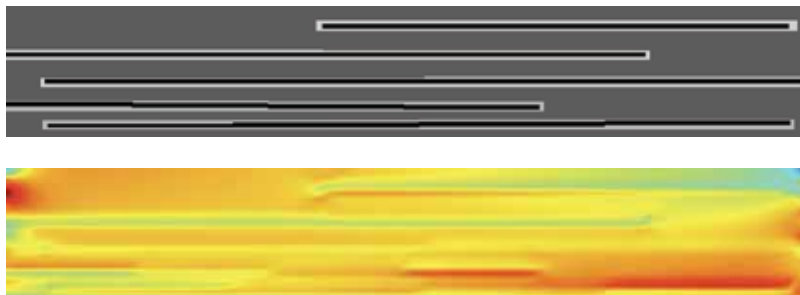


Fig. 18. The model (upper) of composite with fibres with a weak interface, and after simulation showing Von Mises stress (lower).

Fig. 18 shows a representative model of a fibre composite with a weak interface, represented by the lighter shading around each fibre. After simulation of a tensile stress the Von Mises stress distribution shows that stress has not been transferred to the fibres as well as in a composite with a strong fibre-matrix interface (see lower region of Fig. 18).

3.9 Impact initiated delamination

Laminar fibre-matrix composites are subject to failure by delamination under various loading modes and loading rates, including impact forces. The interface strength and deformation behaviour are critical laminate characteristics to be evaluated. Crack growth can be suitably initiated by including interlaminar cracks in the model. Bending, shear and mixed loading modes have been used in the failure simulations. An interface element was used to describe resistance to crack growth (Wisheart & Richardson, 1998).

The resistance to impact has been investigated when composite laminates were subjected to an initial stress. A shear flexible plate composite model was impacted by a mass during the simulation. Finite element analysis was used to solve the contact force and the dynamic response of the composite plate. The initial tensile stress amplified the effect of the impacting mass, while an initial compressive force attenuated the effect of the impacting mass (Sun & Chen, 1985). Barely visible impact damage was predicted using FEM of the larger scale structure. The model included dynamic responses including stress and strain history. The method was validated using experimental data from carbon fibre epoxy composites, and the approach was extended to glass fibre and Kevlar laminates (Davies, et al, 1994).

Low velocity impact damage to composite laminates results in complex cracking of matrix and fibre, and delamination. The complexity of the fractures makes prediction a difficult challenge, since the failure mechanisms are disperse. A FEA of post-impact damage zone with compression buckled delaminated fibres has been undertaken to assess the strength reducing mechanisms. Comparison of the FEM with laminates containing artificially induced damage has demonstrated consistency between prediction and observation (Pavier & Clarke, 1996). Damage mechanisms and mechanics of laminated composites due to low velocity impact has been evaluated (Hyung, et al, 1991).

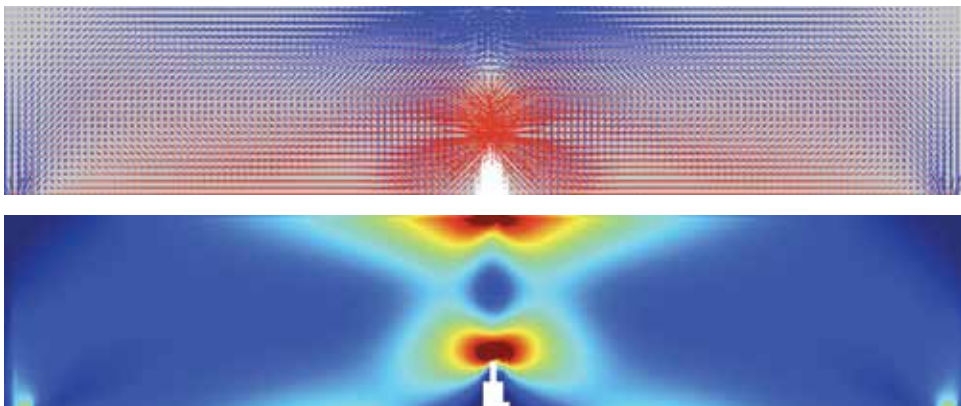


Fig. 19. Stress concentration in a polymer with a crack, induced in three-point bend mode, stress distribution diagram (upper) and Von Mises contour image (lower).

Prediction of damage development to failure in composites is important for assessment of applications. Progressive damage evaluation using FEA can predict potential damage that is often unobserved. 3D FEA of carbon fibre epoxy composites used Abaqus with specially designed meshes to identify criteria to evaluate the onset and reduction in properties due to local damage (Gamble, et al, 1995).

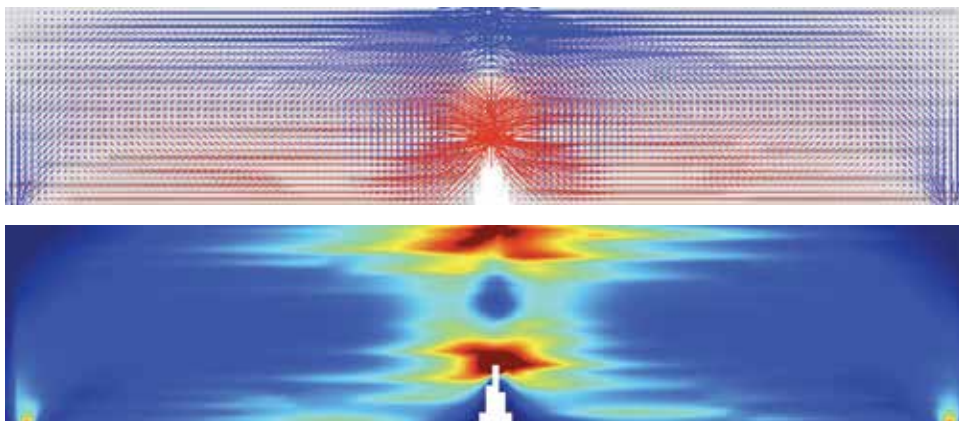


Fig. 20. Stress concentration in a polymer composite with a crack, induced in three-point bend mode, stress distribution diagram (upper) and Von Mises contour image (lower).

Damage development in a polymer is accelerated by the presence of a crack as illustrated in Fig. 19 where application of a stress in three-point bend mode has caused stress tensile concentration at the crack tip and compressive stress concentration at the point of stress application. The stress diagram (Fig. 19, upper) shows the tensile stresses radiating from the crack tip, a situation that is likely to result in crack opening, growth and failure. The Von Mises contour representation (Fig. 19, lower) emphasised the intensity of the stress concentrations. The same FEA simulation of a polymer composite with fibres parallel to the axis of the beam is shown in Fig. 20. The stress concentrations are almost the same as in the plain polymer except that stress has been transferred to the fibres and can be seen to radiate away from the crack tip and application zone along the adjacent fibres. Transfer of stress to the fibres is more clearly seen in the stress distribution diagram (Fig. 20, upper) where the fibres are the darker horizontal lines. The Von Mises contour plot show the stress transfer to fibres are the lighter streaks emanating laterally from the central stress zone.

3.10 Contribution of broken fibre and fibre pullout

Stress concentration at a broken fibre has been simulated by a model including a single broken fibre surrounded by six equally spaced complete fibres in an epoxy resin, based upon a fibre volume fraction of 0.6. The model and separately each component were simulated as a homogeneous, orthotropic material. The stress affected fibre length was half the ineffective broken fibre length (Nedele et al, 1994). The contribution of fibre fracture or a fibre end is typically measured using the single fibre pull-out test. The matrix is fixed in position and a tensile force is applied to the fibre (Wong et al, 2007). A 2D axisymmetric linear finite element model has been developed for fibre pull-out to meet the criteria: energy release rate for interface crack extension, large fibre aspect ratio, singular stress field at the

crack tip, and interface crack initiation for short crack lengths. Interface crack initiation occurred under definite unstable mixed mode conditions, complicating the interpretation of the critical energy release rates (Beckert & Lauke, 1995). A model containing a region of broken fibres is shown in Fig. 21. After simulation under a tensile stress the Von Mises stress contours show that most stress transfer to fibres occurs in the lower region where the fibre ends are overlapped. In the broken fibre (upper) region the stress intensities are low, particularly in the regions between breaks. A consequence is that the broken fibre region is not supporting its share of the stress, and the undamaged region has increased stress concentration that would be likely to cause further fibre breakage in a real situation. Fibre pullout is analogous in that low interfacial bonding will limit the stress that can be transferred to the fibres and the matrix will be more able to deform in the weakly bonded regions.

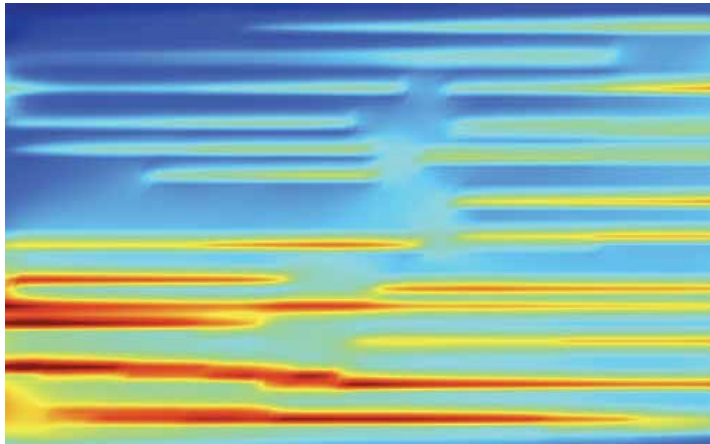


Fig. 21. A fibre composite model including a region of broken fibres (top section) with mostly continuous or overlapping fibre (lower section).

Models for the debonding of a fibre embedded in a brittle matrix have been proposed and assessed. The analysis was in the frictional sliding zone where axial and radial stresses are characterized along the z -axis by a Lamé problem that is illustrated in the article. Steady state stress release rate is evaluated in the absence of friction and the upper limit to interface toughness for mode-2 debonding. Some specific cases considered are: debonding with constant friction, debonding with Coulomb friction, and fibre pullout with constant friction. This method was based upon the use of a continuous distribution of dislocations in an integral equation formulation to replicate a crack. The study was confined to materials having a residual compressive stress across the fibre-matrix interface (Hutchinson & Jensen, 1990).

Matrix dominated failure in polymer composites can be modelled by using cohesive zone interface elements. These elements have been shown to enable simulation of actual failure mechanisms that proceed via damage development through to failure. Cohesive zone elements accurately simulate delamination, in-plane failure of laminates, notch sensitivity, impact fractures through to ultimate debonding failure (Wisnom, 2010).

3.11 Mechanically fastened composites

Composites gain strength from transfer of stress from matrix to high modulus filler, typically fibres. However composites are often mechanically fastened, which requires disruption of structure by a hole with fastener and stress concentration sites at the points of fastening. Observations of failure modes reveal three types: tension, shear-out and bearing. The stress distribution in the model was determined by a finite element method, then the failure stress and failure mode were predicted from a failure criteria hypothesis. The failure mode for a specified stress has been found to depend upon: material properties, joint geometry and fibre ply orientation (Chang et al, 1982). Three-dimensional models of composites with single- and multi-bolt holes with variation in hole clearance were simulated by finite element analysis of shear. Increasing clearance reduced the stiffness of single-bolt joints and changed the load distribution in multi-bolt joints. The simulation results agreed with experiments demonstrating that the finite elements models can become a routine part of the design process (McCarthy, 2002). Fig. 22 demonstrates a mechanically fastenable plaque with three holes that can be potentially bolt holes. The plaque was simulated under compression and the stress distribution diagram (upper image) shows compressive stress lateral to each bolt hole, with tensile stress along the upper and lower edges of each hole. The Von Mises contour diagram (lower image) shows a high stress intensity along the top and bottom edges of each hole, with localised stress along the line of holes especially between pairs of holes.

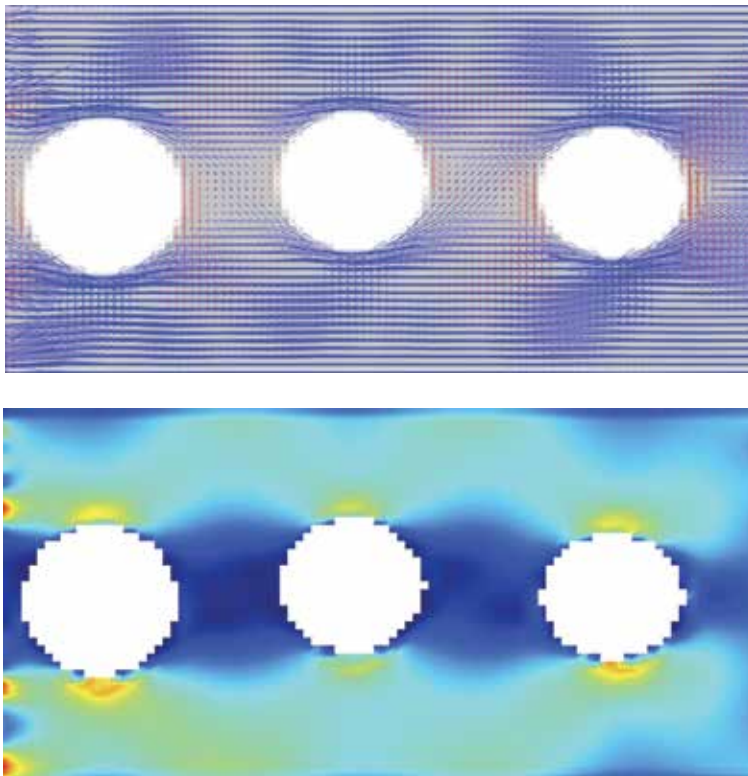


Fig. 22. Plaque with three bolt holes under compression, stress distribution (upper) and Von Mises stress contours (lower).

3.12 Scale and model composite computations

FEA models are mostly of a scale directly relatable to the real system being modelled. The properties of each phase and any discrete interface are global parameters that define the materials and their interactions. The nature of the structure is important for the model to represent the actual composite. However, the components within the composite can be changed by changing the material parameters. For instance, a nylon–glass fibre composite can be changed to a polysulfone–carbon fibre composite by changing the material properties of the respective polymer and fibre, the structure of the composite remains essentially the same. This is because the molecules structure of the polymer and filler is not considered in the model. The same composites could be converted into an epoxy–carbon fibre composite since the model does not differential between a thermoplastic and the highly crosslinked epoxy structure.

A decrease in the scale to micron dimensions brings the model into the realm of meso-materials. At the meso-level discrete atoms are not represented, but molecular are introduced as consisting of beads or chords that are allocated properties consistent with or derived from combinations of real atoms or molecules. Thus the chemical nature of the components of the polymer and filler are included in the model. At the meso-scale a model can only represent a small part of a composite component or article, due to the limitations of computational power and model size. The nodes for the computation can be the meso-scale sub-units instead of a geometric mesh.

Further refinement of scale to the nanometre dimensions brings a model into the molecular domain. In molecular models discrete atomic or in simplified form united atoms are used to build the model. Only a small sub-set of a composite can be modelled because of the complexity of the molecular scale, however discrete chemical properties are represented and assumptions about global properties such as modulus do not need to be made. The properties of the atoms or united atoms are still obtained from sets of parameters that represent the actual molecular properties. These sets of parameters are called force fields, and the values may be derived from actual properties or by ab-initio computational methods from quantum mechanical theory. Application of a suitable force field to the model involves molecular mechanics where the geometries and forces within and between molecules are calculated using Newtonian mechanics, with relationships such as that from Leonard–Jones. Periodic boundary conditions are applied to extend the limited simulation volume to that of a real material (Fig. 23). As an example of this approach, the simulated deformation of a 2D polymer field has been used to calculate a macro-level stress-strain curve, including visualisation of nano-scale void generation, void coalescence and crack formation (Wang, et al, 2003).

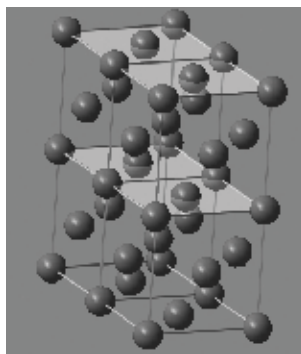


Fig. 23. 3D molecular unit cell model repeatable from each face to create a period boundary system

4. Conclusion

Finite element analysis has provided an implicit means of modelling polymer composites, such as thermoset chopped fibre or woven fabric composites and laminates, thermoplastic-textile and dispersed chopped fibre composites that interprets the stress-strain behaviour in local regions compared with typical volume fraction continuum models. An FEA approach comprising model design, validation, optimisation, simulation and analysis has been described. FEA applied to polymer composites has been reviewed and many examples demonstrated independently using a two-dimensional FEA program. The importance of stress transfer from matrix to fibre is emphasised by stress concentrations. The surface regions and inter-fibre matrix were identified as zones that supported most stress during deformation by bending since one surface is extended while the opposite surface is compressed. Consequently there is a requirement to have fibre rich regions near surfaces and complete compaction, which is readily achieved by a compression moulding process. The interior can be suitably partially compressed to create low density impact resistant composites. A tensile stress gives a uniform deformation field that must be efficiently transferred to the fibres, creating stress concentrations at the fibre-matrix interfaces.

5. References

- Ausiello, P., Apicella, A., Davidson, C. L. (2002). Effect of adhesive layer on stress distribution in composite restorations - a 3D finite element analysis, *Dental Materials*, 18, 295-303.
- Beckert, W., Lauke, B. (1995). Fracture mechanics finite element analysis of debonding crack extension for a single fibre pull-out specimen, *J. Mat. Sci. Lett.*, 14, 333-336.
- Boisse, P., Borr, M., Buet, K., Cherouat, A. (1997). Finite element simulations of textile composite formed including the biaxial fabric behavior, *Comp. B, Eng.*, 28, 453-464.
- Butcher, R. J., Rousseau, C. E., Tippur, H. V. (1999). A functionally graded particulate composite: Preparation, Measurements and failure analysis, *Acta Mater.* 47, 259-268.
- Chang, F. K., Scott, R. A., Springer, G. S. (1982). Strength of mechanically fastened composite joints, *J. Comp. Mat.* 16, 470-494.
- Change, F. K., Chang, K. Y. (1987). A progressive damage model for laminated composites containing stress concentrations, *J. Comp. Mat.*, 21, 834-855.
- Chen, X., Liu, Y. (2001). Multiple-cell modelling of fibre-reinforced composites with the presence of interphases using the boundary element method, *Comp. Mat. Sci.* 21, 86-94.
- Choo, H., Bourke, M. A. M., Daymond, M. R. (2001). A finite-element analysis of the inelastic relaxation of thermal residual stress in continuous-fiber-reinforced composites, *Comp. Sci. Tech.*, 61, 1757-1722.
- Davies, G. A. O., Zhang, X., Zhou, G., Watson, S. (1994). Numerical modelling of impact damage, *Composites*, 25, 342-350.
- Dehoff, P. H., Anusavice, W. J., Wang, Z. (1995). Three-dimensional finite element analysis of the shear bond test, *Dent. Mater.*, 11, 126-31.
- Diaz, J., Romera, L., Hernandez, S., Baldomir, A. (2010). Benchmarking of three-dimensional finite element models of CFRP single-lap bonded joints, *Int. J. Adhesion and Adhesives*, 30, 178-189.

- Dooms, D., Degrande, G., Roeck, G. D., Reynders, E. (2006). Finite element modelling of a silo based experimental nodal analysis, *Eng. Structures*, 28, 532-542.
- Fisher, F. T., Bradshaw, R. D., Brinson, L. C. (2003). Fiber waviness in nanotube-reinforced polymer composites – II: modeling via numerical approximation of the dilute strain concentration tensor, *Comp. Sci. Tech.*, 63, 1705-1722.
- Fisher, F. T., Brinson, L. C. (2001). Viscoelastic interphases in polymer-matrix composites: Theoretical models and finite element analysis, *Composites Sci. and Tech.* 61, 731-748.
- Gamble, K., Pilling, M., Wilson, A. (1995). An automated finite element analysis of the initiation and growth of damage in carbon-fibre composite materials, *Comp. Structures*, 32, 265-274.
- Guild, F. J., Young, R. J. (1989). A predictive model for particulate-filled composite materials, *J. Mat. Sci.*, 24, 298-306.
- Houshyar, S., Hodzic, A., Shanks, R. A. (2009). Modelling of polypropylene fibre-matrix composites using finite element analysis, *eXPRESS Polymer Letters*, 3, 2-12.
- Houshyar, S., Shanks, R. A. (2003). Morphology, thermal and mechanical properties of poly(propylene) fibre-matrix composites, *Macromol. Mat. Eng.*, 288, 599-606.
- Houshyar, S., Shanks, R. A. (2010). Interfacial properties of all-polypropylene composites, *e-Polymers*, 33, 1-13.
- Houshyar, S., Shanks, R. A., Hodzic, A. (2005). The effect of fiber concentration on mechanical and thermal properties of fiber-reinforced polypropylene composites, *J. Appl. Polym. Sci.*, 96, 2260-2272.
- Huang A., Bush M. B. (1997). Finite element analysis of mechanical properties in discontinuously reinforced metal matrix composites with ultra fine micro structure. *Mat. Sci. Eng. A*, 232, 63-72.
- Hutchinson, J. W., Jensen, H. M. (1990). Models of fiber debonding and pullout in brittle composites with friction, *Mecahnical of Materials*, 9,139-163.
- Hyung, Y. C., Wu, H. Y. T., Fu-Kuo, C. (1991). A new approach toward understanding damage mechanisms and mechanics of laminated composites due to low-velocity impact. II: analysis, *J. Comp. Mat.*, 25, 1012-1038.
- Isakson, G., Levy, A. (1971). Finite element analysis of interlaminar shear in fibrous composites, *J. Comp. Mat.*, 5, 273-276.
- Kang G. Z., Gao Q. (2002). Tensile properties of randomly oriented short Al_2O_3 fibre reinforced aluminium alloy composites: II. Finite element analysis for stress transfer, elastic modulus and stress-strain curve. *Composites A, Appl. Sci. Manufacturing*, 33, 657-667.
- Lamkanfi, E., Paeppegem, W. V., Degrieck, J., Ramault C. (2010). Strain distribution in cruciform specimens subjected to biaxial loading conditions. Part 1. Two-dimensional finite element model, *Polym. Testing*, 29, 7-13.
- Lee, S. H., Wang, S., Pharr, G. M., Xu, H. (2007). Evaluation of interphase properties in a cellulose fiber-reinforced polypropylene composite by nanoindentation and finite element analysis, *Composites A*, 38, 1517-1524.
- Lim, J. H., Ratnam, M. M., Abdul-Khalil H. P. S (2003). An experimental and finite element analysis of the static deformation of natural fiber-reinforced composite beam, *Polym. Testing*, 22, 169-177.

- Lin, D. X., Ni, R. G., Adams, R. D. (1984). Prediction and measurement of the vibrational damping parameters of carbon and glass fibre-reinforced plastics plates, *J. Comp. Mat.*, 18, 132-152.
- McCarthy, M. A., McCarthy, C. T. (2002). Finite element analysis of the effects of clearance on single-shear, composite bolted joints, *J. Plastics, Rubber Comp.*, 32, 1-11.
- Maligno A. R., Warrior N. A., Long A. C. (2008). Finite element investigations on the microstructure of fibre reinforced composites. *Express Polymer Letters*, 2, 665-676.
- Muheim, D., Griffen, O. H. (1990). 2-D to 3-D global/local finite element analysis of cross-ply composite laminates, *J. Reinforced Plastics Comp.*, 9, 492-502.
- Naik, N. K., Shembekar, P. S. (1992). Elastic behaviour of woven fabric composites: I – lamina analysis, *J. Comp. Mat.*, 26, 2196-2225.
- Nakamura, T., Imanishi, A., Kashima, H., Ohyama, T., Ishigaki, S. (2001), Stress analysis of metal-free polymer crowns using the three-dimensional finite element method, *Int. J. Prosthodont*, 14, 401-405.
- Nedele, M. R., Wisdom, M. R. (1994). Three-dimensional finite element analysis of the stress concentration at a single break, *Comp. Sci. Tech.*, 51, 517-524.
- Pal, P., Ray, C. (2002). Progressive failure analysis of laminated composite plates by finite element method, *J. Reinforced Plastics and Composites*, 21, 1505-1513.
- Pavier, M. J., Clarke, M. P. (1996). Finite element prediction of the post-impact compressive strength of fibre composites, *Comp. Structures*, 36, 141-153.
- Pegoretti, A., Fambri, L., Zappini, G., Bianchetti, M. (2002). Finite element analysis of a glass fibre reinforced composite endoplastic post, *Biomaterials*, 23, 2667-2682.
- Shati F. K., Esat I. I., Bahai H. (2001). FEA modelling of visco-plastic behaviour of metal matrix composites. *Finite Elements in Analysis and Design*, 37, 263-272.
- Song, Y. S., Youn, J. R. (2006). Evaluation of the effective thermal conductivity for carbon nanotube/polymer composites using control volume finite element method, *Carbon*, 44, 710-717.
- Soykasap, O. (2010). Finite element analysis of plain weave composites for flexural failure, *Polym. Composites*, 31, 581-586.
- Sun, C. T., Chen, J. K. (1985). On the impact of initially stressed composite laminates, *J. Comp. Mat.*, 19, 490-504.
- Tan, P., Tong, L., Steven, G. P. (1998). Modelling approaches for 3D orthogonal woven composites, *J. Reinforced Plastics Comp.*, 17, 545-577.
- Taylor, M., Verdonchot, N., Huiskes, R., Zioupos, P. (1999). A combined finite element method and continuum damage mechanics approach to simulate the in vitro behavior of human cortical bone, *J. Mat. Sci.: Mat. In Medicine*, 10, 841-846.
- Tserpes, K. I., Papanikos, P. (2005). Finite element modeling of single-walled carbon nanotubes, *Composites B. Eng.*, 36, 468-477.
- Tsui, C. P., Tang, C. Y., Lee, T. C. (2001). Finite element analysis of polymer composites filled by interphase coated particles, *J. Mat. Process. Tech.*, 117, 105-110.
- Vozkova, P. (2009). Elastic modulus FEA modelling of the layered woven composite material, in Petrone, G., Cammarata, G. (editors), *Modelling and Simulation*, SCIYO Publishing, ISBN 978-3-902613-25-7, p. 651.
- Wang, J., Kelly, D., Hillier, W. (2000). Finite element analysis of temperature induced stresses and deformations of polymer composite components, *J. Comp. Mat.*, 34, 1456-1471.

- Wang, Y., Sun, C., Sun, X., Hinkley, J., Odegard, G. M., Gates, T. S. (2003). 2-D nano-scale finite element analysis of a polymer field, *Comp. Sci. Tech.*, 63, 1581-1590.
- Whitcomb, J. D. (1981). Finite element analysis of instability related delamination growth, *J. Comp. Mat.*, 15, 403-426.
- Whitcomb, J., Srinengan, K. (1996). Effect of various approximations on predicated progressive failure in plane weave composites, *Comp. Structures*, 34, 13-20.
- Wisheart, M., Richardson, M. O. W. (1998). The finite element analysis of impact induced delamination in composite materials using a novel interface element, *Composites A*, 29A, 301-313.
- Wisnom, M. R. (2010). Modelling discrete failures in composites with interface elements, *Composites A*, doi:10.1016/j.compositesa.2010.02.011.
- Wong, S., Shanks, R. A., Hodzic, A. (2007). Effect of additives on the interfacial strength of poly(l-lactic acid) and poly(3- hydroxybutyric acid)-flax fibre composites, *Comp. Sci. Tech.*, 67, 2478-2484.

Finite Element Analysis of Nonlinear Isotropic Hyperelastic and Viscoelastic Materials for Thermoforming Applications

Erchiqui Fouad

Université du Québec an Abitibi-Témiscamingue

Rouyn-Noranda, Québec

Canada

1. Introduction

Thermoforming of cut sheets is extensively used in the industry for various commercial applications. The sheet is heated to a softened state and subsequently deformed in a mold using applied pressure, a vacuum, a moving plug, or a combination of these. The thermoforming market is expanding to include more complex geometries and a wider range of potential materials. Material must be selected early in the product development process, based on cost, service properties, and processability.

However, in the thermoplastic forming industry, it is a common practice to mostly rely on trial and error methods based on a high number of test specimens in order to assess the actual mass production planning and schedule, a procedure that involves costly operations. This situation also occurs when designing a new product or while improving existing processes. It is therefore appealing to develop more efficient alternatives based for example on reliable numerical predictive tool for the analysis and optimisation of these manufacturing processes. This is enabled by the fact that it is now possible to develop precise sophisticated models describing not only the thermal and mechanical phenomena taking place within the materials but also describing the interactions between the different components of the fabrication system. These mathematical models are then implemented into user friendly computer softwares that are validated through laboratory tests before their actual implementation in industrial product design and system manufacturing simulation.

In the finite element codes used to handle these problems, the application of loadings is in general defined in terms of pressure. For such kind of loading, quasi-static problems are largely discussed in the technical literature, among which one can mention the following works (DeLorenzi & Nied, 1991), (Laroche & Erchiqui, 2000) for hyperelastic materials and (Laroche & Erchiqui, 1998) for viscoelastic materials. For the case of dynamic problems, actually less considered than their quasi-static counterparts, we can mention references like (Verron & al, 2001), (Erchiqui & al, 2001) for thermoplastic materials. In these problems, strong mechanical and geometric non linearities induce instabilities during blowing and

thus require, for quasi-static problems, an accurate control of the loadings employed in the simulation (Khayat & al, 1992) and for dynamic problems, an adequate time step control (Lapidus & Pinder, 1982). These instabilities, associated with the descending segment of the load-deformation curve, correspond physically to regions in which the load (pressure) decreases while the expansion is still going on. In order to be able to follow the loading curve, we consider the dynamic problem of thermoforming of the viscoelastic and the hyperelastic sheet subjected to air flow loading (Erchiqui & al, 2005). The Inflation process of the membrane is controlled by the gas equation of state. The dynamic pressure is thus deduced from one of the three equations of state: the Redlich-Kwong, the van der Waals and the ideal equations.

For modeling purposes, we adopt the finite element method based on the Lagrangian formulation. Membrane theory and material incompressibility assumptions are considered. The membrane structure is discretized by plane finite elements (Zienkiewicz & Taylor, 1991) and time integration is performed via an explicit algorithm based on central difference method that is conditionally stable (Lapidus & Pinder, 1982). For constitutive behavior, the viscoelastic integral models of Lodge (Lodge, 1964) and Christensen (Christensen, 1980) type and the hyperelastic models of Mooney-Rivlin (Rivlin, 1948) and Ogden (Ogden, 1972) type are considered. Moreover, the influence of all these material constitutive models on the thickness distribution for thermoforming a part in ABS is analysed using a dynamic loading by air flow characteristics.

2. Kinematics

We consider a viscoelastic material medium, occupying a deformable continuous body Ω of surface Γ . Each material point P , belonging to the body Ω , can be identified by using three curvilinear co-ordinates $(\theta^1, \theta^2, \theta^3)$, which remain constant at the time of inflation. In the case of a thin viscoelastic structure, the number of these curvilinear co-ordinates can be reduced to two (θ^1, θ^2) defined on the membrane's mid surface. Consequently, the position of a point P belonging to this mid surface will be defined, in the undeformed reference configuration, by the vector $\mathbf{X}(\theta^1, \theta^2)$, and, in the deformed configuration, by the vector $\mathbf{x}(\theta^1, \theta^2)$. Also, at the mid surface point P , we associate, in the reference configuration, a thickness function $H_0(\theta^1, \theta^2)$ and local tangent vectors $(\mathbf{G}_1, \mathbf{G}_2)$, and in the deformed configuration, we associate a thickness $h(\theta^1, \theta^2)$ and local tangent vectors $(\mathbf{g}_1, \mathbf{g}_2)$. Thus, given the structure mid surface parameterizations defined by, $\mathbf{X}(\theta^1, \theta^2)$ and $\mathbf{x}(\theta^1, \theta^2)$, the introduced tangent vectors can be determined from the following classical formula:

$$\mathbf{G}_1 = \frac{\partial \mathbf{X}}{\partial \theta^1}, \quad \mathbf{G}_2 = \frac{\partial \mathbf{X}}{\partial \theta^2} \quad (1)$$

$$\mathbf{g}_1 = \frac{\partial \mathbf{x}}{\partial \theta^1}, \quad \mathbf{g}_2 = \frac{\partial \mathbf{x}}{\partial \theta^2} \quad (2)$$

However, in order to describe the deformation in the direction perpendicular to the thin structure's mid surface, the information on the tangent vectors $(\mathbf{G}_1, \mathbf{G}_2)$ and $(\mathbf{g}_1, \mathbf{g}_2)$ must be supplemented by out-of-plane information. To address this situation, we introduce two unit vectors, \mathbf{G}_3 and \mathbf{g}_3 , orthogonal to the surface covariant base vectors $(\mathbf{G}_1, \mathbf{G}_2)$ and $(\mathbf{g}_1, \mathbf{g}_2)$ defined respectively by:

$$\mathbf{G}_3 = \mathbf{N} = \frac{\mathbf{G}_1 \times \mathbf{G}_2}{\|\mathbf{G}_1 \times \mathbf{G}_2\|}, \quad \mathbf{g}_3 = \mathbf{n} = \frac{\mathbf{g}_1 \times \mathbf{g}_2}{\|\mathbf{g}_1 \times \mathbf{g}_2\|} \quad (3)$$

Vectors \mathbf{N} and \mathbf{n} are outward unit vectors normal to the tangent planes defined by vectors $(\mathbf{G}_1, \mathbf{G}_2)$ and $(\mathbf{g}_1, \mathbf{g}_2)$, respectively. The displacement field generated by the movement of point \mathbf{P} of the membrane's mid surface can be described by a function \mathbf{u} defined by (difference between the current position vector and the reference undeformed position vector):

$$\mathbf{u} = \mathbf{x} - \mathbf{X} \quad (4)$$

As a measure of the deformation of the membrane, we use the generalized Green-Lagrangian strain, whose components are:

$$\mathbf{E}_{\alpha\beta} = \frac{1}{2} (\mathbf{g}_\alpha^T \cdot \mathbf{g}_\beta - \mathbf{G}_\alpha^T \cdot \mathbf{G}_\beta), \quad \alpha, \beta = 1, 2 \quad (5)$$

The notation $*$ ^T denotes the transposition of the line vector associated with the column vector $*$. To describe the deformed covariant basis vector, \mathbf{g}_α , in terms of covariant undeformed basis vector \mathbf{G}_α , we use the transformation:

$$\mathbf{g}_\alpha = \underline{\underline{\mathbf{F}}} \cdot \mathbf{G}_\alpha \quad (6)$$

where $\underline{\underline{\mathbf{F}}}$ is the deformation gradient tensor at \mathbf{x} :

$$\underline{\underline{\mathbf{F}}} = \underline{\underline{\mathbf{I}}} + \nabla \mathbf{u} \quad (7)$$

$\underline{\underline{\mathbf{I}}}$ is the second-order identity tensor. Thus, relative to the reference configuration, the Green-Lagrange deformation tensor, given in general curvilinear co-ordinates by (5), becomes, in matrix form:

$$\underline{\underline{\mathbf{E}}} = \frac{1}{2} (\underline{\underline{\mathbf{C}}} - \underline{\underline{\mathbf{I}}}) \quad (8)$$

where $\underline{\underline{\mathbf{C}}}$ is the right Cauchy-Green deformation tensor defined by $\underline{\underline{\mathbf{C}}} = \underline{\underline{\mathbf{F}}}^T \cdot \underline{\underline{\mathbf{F}}}$. The virtual Green-Lagrange deformation tensor then becomes:

$$\delta \underline{\underline{\mathbf{E}}} = \frac{1}{2} (\delta \underline{\underline{\mathbf{F}}}^T \cdot \underline{\underline{\mathbf{F}}} + \underline{\underline{\mathbf{F}}}^T \cdot \delta \underline{\underline{\mathbf{F}}}) \quad (9)$$

3. Preliminary considerations

The problem of thin structure material inflation can be modelled by considering the large deformation of the body with finite strain. The continuum formulation using the updated Lagrangian description is presented using a 3-node isoparametric element. Although the

formulation is applicable for the 3D geometry, the finite element discretization is shown for the plane strain configuration only, for the sake of simplicity. The discussion is followed by a membrane formulation, and the solution method.

4. Dynamic force balance equation

We utilize the dynamic finite element method with both space and time discretization to simulate the inflation of the thermoplastic membrane. Thus, the principle of virtual work is expressed on the undeformed configuration for the inertial effects and internal work. Then, this weak form of the principle of virtual work is discretized and assembled by applying the finite elements methodology (Erchiqui & al, 2005). Because of the presence of the inertial force, a time discretization is required. This is handled here through the introduction of a centred finite difference technique that is conditionally stable (Lapidus & Pinder, 1982).

4.1 Virtual work expression

In the reference configuration, the expression of the principle of virtual work is given by the following formula (Erchiqui & al, 2005):

$$\delta W = \int_{\Omega_0} \delta E : \underline{S} d\Omega_0 + \int_{\Omega_0} \rho_0 \ddot{u}^T \cdot \delta u d\Omega_0 + \int_{\Omega_0} \underline{f}_0 \cdot \delta u d\Omega_0 - \int_{\Gamma} \delta u^T \cdot \underline{t} d\Gamma \quad (10)$$

where \underline{S} is the second Piola-Kirchhoff stress tensor, Γ is the surface of the deformed membrane, \underline{t} is the external surface force due to pressure and δu is a virtual displacement vector compatible with the displacement boundary conditions. The principle of virtual work expression, given by Eqn. (10), can be broken up into four terms:

$$\delta W(u, \delta u, t) = \delta W^{\text{int}} + \delta W^{\text{acc}} + \delta W^{\text{grav}} - \delta W^{\text{ext}} \quad (11)$$

where δW^{int} , δW^{acc} , δW^{ext} and δW^{grav} are the internal, inertial external and body work, respectively:

$$\delta W^{\text{int}} = \int_{\Omega_0} \delta E : \underline{S} d\Omega_0 \quad (12.a)$$

$$\delta W^{\text{ext}} = \int_{\Gamma} \delta u^T \cdot \underline{t} d\Gamma \quad (12.b)$$

$$\delta W^{\text{acc}} = \int_{\Omega_0} \rho_0 \ddot{u}^T \cdot \delta u d\Omega_0 \quad (12.c)$$

$$\delta W^{\text{grav}} = \int_{\Omega_0} \underline{f}_0 \cdot \delta u d\Omega_0 \quad (12.d)$$

By introducing the pressure force, $\underline{t} = \Delta p(t) \cdot \underline{n}$, which depends on the deformed geometry, we obtain for the external work:

$$\delta W^{\text{ext}} = \int_{\Gamma} \Delta p \delta u^T \cdot \underline{n} d\Gamma \quad \forall \delta u \quad (12.e)$$

where Δp represents the internal pressure.

4.2 Spatial approximation

After summation of all element contributions, the inflation problem is then reduced to the following discrete system of equations:

$$\mathbf{F}_{\text{acc}} + \mathbf{F}_{\text{grav}} + \mathbf{F}_{\text{int}} - \mathbf{F}_{\text{ext}} = 0 \quad (13)$$

where \mathbf{F}_{acc} , \mathbf{F}_{ext} , \mathbf{F}_{grav} and \mathbf{F}_{int} are the global nodal inertial, external, body and internal force vectors experienced by the thermoplastic membrane. By introducing the mass matrix \mathbf{M} associated with the inertial forces, the inflation problem can then be reduced to a system of second-order ordinary differential equations:

$$\mathbf{M} \cdot \ddot{\mathbf{u}}(\mathbf{t}) = \mathbf{F}_{\text{ext}} + \mathbf{F}_{\text{grav}} - \mathbf{F}_{\text{int}} \quad (14)$$

The system (14) is decoupled by applying the diagonalization method of (Lapidus & Pinder, 1982) to matrix \mathbf{M} , which transforms it into a diagonal mass matrix, \mathbf{M}^d .

4.3 Time discretization

Knowing the velocity and acceleration vectors at time t_n , it is possible, by using a time integration scheme, to calculate the displacement vector at a discrete time t_{n+1} . A centered finite difference scheme is adopted here (Erchiqui & al, 2005), and Eqn. (14) then becomes:

$$\mathbf{u}_i(\mathbf{t} + \Delta t) = \frac{\Delta t^2}{\mathbf{M}_{ii}^d} \left(\mathbf{F}_i^{\text{ext}}(\mathbf{t}) + \mathbf{F}_i^{\text{grav}}(\mathbf{t}) - \mathbf{F}_i^{\text{int}}(\mathbf{t}) \right) + 2\mathbf{u}_i(\mathbf{t}) - \mathbf{u}_i(\mathbf{t} - \Delta t) \quad (15)$$

\mathbf{M}_{ii}^d are the diagonal components of the matrix \mathbf{M}^d .

4.4 Stability limit of the dynamic explicit method

Advancing a time step of this scheme requires a lot of calculations. The price paid for this simplicity is *conditional stability* (Lapidus & Pinder, 1982). If the dynamic system includes phenomena which evolve more rapidly than the approximate solution itself, these will be incorrect for exponentially increasing components of the solution. This *instability* typically results in the blow-up of the solution. Conditional stability for mechanical problems is expressed through the Courant-Friedrichs-Lewy criterion: the time step must be smaller than the critical time steps Δt_{crit} :

$$\Delta t \leq \Delta t_{\text{crit}} = \varepsilon \frac{d}{c} \quad (16)$$

Here c is the wave speed in the medium and d the element size. The quantity d/c is the time that a wave needs to propagate across an element of size d . The proportionality constant, ε , depends on the integration scheme used.

5. Gas equations of state and pressure loading

In addition to the traditional approach, which considers loads in pressures for the numerical modeling of the free and confined inflation in the viscoelastic structures, there is a relatively recent approach which considers gas flows for modeling (Erchiqui, 2007). This is correct at normal temperatures and pressures. At low temperatures or high pressures, real gases deviate significantly from ideal gas behavior. Indeed, at normal pressures, the volume

occupied by the atoms or molecules is a negligibly small fraction of the total volume of the gas. But at high pressures, this is no longer true. As a result, real gases are not as compressible under high pressures as an ideal gas. The volume of a real gas is larger than the expected one from the ideal gas equation at high pressures. Moreover, the assumption of the absence of a force of attraction between ideal gas particles involves a violation of the physical observations such as these gases cannot condense to form liquids. In reality, there is a small force of attraction between gas molecules, which tends to hold them together. This force of attraction has two consequences: (1) gases condense to form liquids at low temperatures; and (2) the pressure of a real gas is sometimes smaller than expected for an ideal gas. To be in conformity with these observations, we are interested in this work by the laws of state which characterize real gases such as: the van de Waals gas and Redlich-Kwong state equations, in addition to the equation of ideal gases. These equations will enable determination of the acceptable dynamic pressure field distribution during free and confined inflation.

The assumptions used for the calculation of the dynamic pressure are:

- i) Gas temperature is assumed constant (T_{gas});
- ii) Pressure between the sheet and mould is assumed to be constant (p_0).

If we represent by V_0 an initial volume enclosing the membrane at the initial time t_0 and containing a number n_0 of gas moles (we assume that the forming process temperature is constant), then, for the cases of the ideal, Van Der Waals and the Redlich-Kwong equations of states of gas:

a) Ideal gas equation

The ideal gas equation of state is given by:

$$P_0 = \frac{n_0 R T_{\text{gas}}}{V_0} \quad (17)$$

P_0 is the initial pressure and R is the universal gas constant

b) Van der Waals gas equation

The van der Waals equation of state is given by (Landau & Lifshitz, 1984):

$$P_0 = \frac{n_0 R T_{\text{gas}}}{V_0 - n_0 b} - \frac{n_0^2 a}{V_0^2} \quad (18)$$

Parameters a and b are constants that change from gas to gas. The constant b is intended to account for the finite volume occupied by the molecules, the term n_0^2/V_0^2 account for the initial forces of attraction between molecules. Please note that when a and b are set at zero, the ideal gas equation of state results.

c) Redlich-Kwong gas equation

The Redlich-Kwong equation of state is given by (Redlich & Kwong, 1949):

$$P_0 = \frac{n_0 R T_{gas}}{V_0 - n_0 b} - \frac{n_0^2 a}{V_0 (V_0 + n_0 b) \sqrt{T_{gas}}} \quad (19)$$

where parameters a and b are calculated in terms of the values of pressure p_c and temperature T_c at the critical point of the gas under examination, i.e.:

$$a = 0.42748 \frac{\bar{R}^2 T_c^{2.5}}{p_c} \text{ and } b = 0.08664 \frac{\bar{R} T_c}{p_c} \quad (20)$$

With $\bar{R} = 8.314 \text{ kJ} / \text{kmol.K}$.

Now if we represent by $n(t)$ the additional number of moles of gas introduced for the inflation of the thermoplastic membrane, by $p(t)$ the internal pressure and by $V(t)$ the additional volume occupied by the membrane at time t , we get the following relations:

a) *Ideal gas equation*

$$\Delta P = P(t) - P_0 = \left[\frac{(n(t) + n_0) R T_{gas}}{(V(t) + V_0) - (n(t) + n_0) b} \right] - \left[\frac{n_0 R T_{gas}}{V_0 - n_0 b} \right] \quad (21)$$

b) *Van der Waals gas equation*

$$\Delta P = P(t) - P_0 = \left[\frac{(n(t) + n_0) R T_{gas}}{(V(t) + V_0) - (n(t) + n_0) b} - \frac{(n_0 + n(t))^2 a}{(V(t) + V_0)^2} \right] - \left[\frac{n_0 R T_{gas}}{V_0 - n_0 b} - \frac{n_0^2 a}{V_0^2} \right] \quad (22)$$

c) *Redlich-Kwong gas equation*

$$\Delta P(t) = P(t) - P_0 = \left[\frac{(n(t) + n_0) R T_{gas}}{(V(t) + V_0) - (n(t) + n_0) b} - \frac{(n(t) + n_0)^2 a}{(V(t) + V_0)((V(t) + V_0) + (n(t) + n_0) b) \sqrt{T_{gas}}} \right] - \left[\frac{n_0 R T_{gas}}{V_0 - n_0 b} - \frac{n_0^2 a}{V_0 (V_0 + n_0 b) \sqrt{T_{gas}}} \right] \quad (23)$$

Each equation, 21-23, represents the time evolution of pressure inside the thermoplastic membrane, closely related to the evolution of the internal volume of the membrane via the thermodynamic state equation. In these cases, the external virtual work becomes:

a) *ideal gas equation*

$$\delta W^{\text{ext.}} = \int_{\Gamma} \left(\left[\frac{(n(t) + n_0) RT_{\text{gas}}}{(V(t) + V_0)} \right] - \left[\frac{n_0 RT_{\text{gas}}}{V_0} \right] \right) \delta \mathbf{u}^T \cdot \mathbf{n} \, d\Gamma \quad \forall \quad \delta \mathbf{u} \quad (24)$$

b) *Van der Waals gas equation*

$$\delta W^{\text{ext.}} = \int_{\Gamma} \left(\left[\frac{(n(t) + n_0) RT_{\text{gas}}}{(V(t) + V_0) - (n(t) + n_0)b} - \frac{(n_0 + n(t))^2 a}{(V(t) + V_0)^2} \right] - \left[\frac{n_0 RT_{\text{gas}}}{V_0 - n_0 b} - \frac{n_0^2 a}{V_0^2} \right] \right) \delta \mathbf{u}^T \cdot \mathbf{n} \, d\Gamma \quad \forall \quad \delta \mathbf{u} \quad (25)$$

c) *Redlich-Kwong gas equation*

$$\delta W^{\text{ext.}} = \int_{\Gamma} \left(\left[\frac{(n(t) + n_0) RT_{\text{gas}}}{(V(t) + V_0) - (n(t) + n_0)b} - \frac{(n(t) + n_0)^2 a}{(V(t) + V_0)((V(t) + V_0) + (n(t) + n_0)b)\sqrt{T_{\text{gas}}}} \right] - \left[\frac{n_0 RT_{\text{gas}}}{V_0 - n_0 b} - \frac{n_0^2 a}{V_0(V_0 + n_0 b)\sqrt{T_{\text{gas}}}} \right] \right) \delta \mathbf{u}^T \cdot \mathbf{n} \, d\Gamma \quad \forall \quad \delta \mathbf{u} \quad (26)$$

The advantage of using a load expressed in term of gas flow instead of pressure loading is that it allows natural exploration of the load-deformation curve and nor worry about the instability phenomenon encountered when classical pressure loading is employed. The introduction of constant pressure as a loading force instead of gas flow velocity in the two used finite element formulations, quasi-static and dynamic, leads to a divergence of the computations for the values of pressure beyond the critical point (starting of the unstable segment of the load-deformation curve) (Erchiqui & al, 2001).

6. Implemented constitutive models

In this work, we consider the assumptions of the plane stresses and of the incompressibility of the thermoplastic material. It follows that the components of the Cauchy stress tensor have the following properties:

$$\sigma_{13} = \sigma_{23} = \sigma_{31} = \sigma_{32} = \sigma_{33} = 0 \quad (27)$$

In the thermoforming of cut sheet, the polymer is subsequently heated in an oven and thermoformed, usually under vacuum. The sheet forming operation is characterized by rapid deformation of the sheet. Modeling of this stage of the process can be done by using either hyperelastic or viscoelastic constitutive equation since the time dependence of the deformation is not critical.

Hyperelastic materials are defined by the existence of scalar function $W(\mathbf{F})$ called the strain energy function, from which stresses can be derived at each point. In order to satisfy the objectivity requirements, the strain energy function must be invariant under changes of the observer frame of reference. It is well known that the Cauchy-Green deformation tensor is invariant under changes of the observer frame of reference. Thus, if the strain energy function can be written as a function of \mathbf{C} , it automatically satisfies the objective principle.

The general stress-strain relationship is given by the formula:

$$S_{ij} = 2 \frac{\partial W}{\partial C_{ij}} \quad (28)$$

where \mathbf{S} is the Piola-Kirchhoff stress tensor.

The different models that exist in the literature define the strain energy as a function of the strain field. The Rivlin theory (Rivlin, 1948) for isotropic materials describes the energy as a function of the three Cauchy strain invariants I_1 , I_2 and I_3 . Assuming the incompressibility of the material, the stress can be obtained from the following Mooney-Rivlin form:

$$W = \sum_{i+j=1}^N (I_1 - 3)^i (I_2 - 3)^j \quad (29)$$

The use of two terms in the series is sufficient to describe the elastic modulus in both uniaxial and biaxial deformation modes.

Another popular strain energy formulation is the Ogden model (Ogden, 1972) that uses the principal values λ_i ($i=1, 2, 3$) from the right Green-Cauchy strain tensor:

$$W = \sum_{i=1}^N \frac{\alpha_i}{\mu_i} (\lambda_1^{\mu_i} + \lambda_2^{\mu_i} + \lambda_3^{\mu_i} - 3) \quad (30)$$

where α_i and μ_i are the material constants. The use of two terms in the series is usually sufficient to describe the material non-linear response under deformation.

Integral type viscoelastic models relate the true stress to the strain history. These models are more appropriate for representing polymers at the liquid (Lodge, 1964) or semi-solid state (Christensen, 1980), (Laroche & Erchiqui, 2000). In this work, to analyze the viscoelastic behavior and isotropic thermoplastic materials, we consider the Lodge model and Christensen model.

The Lodge model (Lodge, 1964) has been developed for representing the viscoelastic deformation of liquid materials. For this model, the Cauchy stress tensor σ , is related at time t to the history of the Finger deformation tensor \mathbf{B} by:

$$\sigma(t) = -p(t) \mathbf{I} + \int_{-\infty}^t \sum_k \frac{g_k}{\tau_k} e^{-(t-\tau)/\tau_k} \mathbf{B}(\tau, t) d\tau \quad (31)$$

where the Finger tensor \mathbf{B} is related to the right Cauchy Green deformation tensor \mathbf{C} by:

$$\mathbf{B} = \mathbf{C}^{-1} = [\mathbf{F}^T \mathbf{F}]^{-1} \quad (32)$$

\mathbf{F} is the deformation gradient tensor, \mathbf{I} is the identity tensor, p is the hydrostatic pressure, g_k and τ_k representing respectively the relaxation modulus and the relaxation time. The hydrostatic pressure p result from the plane stress state due to membrane hypothesis $\sigma_{zz}=0$.

The Christensen model (Christensen, 1980), has been developed for representing the viscoelastic deformation of solids. It can be used for the thermoforming of semi-solid materials. This model gives the second Piola-Kirchhoff stress tensor \mathbf{S} at time t as a function of the Lagrangian strain history \mathbf{E} :

$$\mathbf{S}(t) = -p(t) \mathbf{C}^{-1} + g_0 \mathbf{I} + \int_{-\infty}^t g_1(t-\tau) \frac{\partial \mathbf{E}(\tau, t)}{\partial \tau} d\tau \quad (33)$$

where g_0 is the hyperelastic modulus and g_1 is the material relaxation function given by equation:

$$g_l = \sum_k C_k e^{-(t-\tau)/\tau_k} \quad (34)$$

where C_k is the stiffness modulus. In a Lagrangian formulation, we consider, for our numerical modeling, the second Piola-Kirchhoff stress tensor \mathbf{S} defined by the following transformation:

$$\mathbf{S}(\mathbf{t}) = \mathbf{J}(\mathbf{t}) \mathbf{F}^{-1}(\mathbf{t}) \boldsymbol{\sigma}(\mathbf{t}) \mathbf{F}^{-T}(\mathbf{t}) \quad (35)$$

where the $\mathbf{J}(\mathbf{t})$ is the Jacobian of the transformation. In the case of incompressible thermoplastic materials, the jacobian is unity:

$$\mathbf{J}(\mathbf{t}) = \det(\mathbf{F}) = 1 \quad (36)$$

In our study, for the plane stress state, the deformation and stress matrices, $\mathbf{C}(\mathbf{t})$ and $\mathbf{S}(\mathbf{t})$, have respectively the following forms:

$$[\mathbf{C}(\mathbf{t})] = \begin{bmatrix} C_{xx}(\mathbf{t}) & C_{xy}(\mathbf{t}) & 0 \\ C_{xy}(\mathbf{t}) & C_{yy}(\mathbf{t}) & 0 \\ 0 & 0 & C_{zz}(\mathbf{t}) \end{bmatrix} \quad (37)$$

$$[\mathbf{S}(\mathbf{t})] = \begin{bmatrix} S_{xx}(\mathbf{t}) & S_{xy}(\mathbf{t}) & 0 \\ S_{xy}(\mathbf{t}) & S_{yy}(\mathbf{t}) & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (38)$$

The term $C_{zz}(\mathbf{t})$ in equation (37) can be directly computed from the other components of the deformation tensor:

$$C_{zz}(\mathbf{t}) = \lambda_3^2(\mathbf{t}) = \frac{1}{C_{xx}(\mathbf{t})C_{yy}(\mathbf{t}) - C_{xy}(\mathbf{t})C_{xy}(\mathbf{t})} \quad (39)$$

Where $\lambda_3(\mathbf{t})$ is the principal stretch ratio in thickness direction defined by:

$$\lambda_3(\mathbf{t}) = \frac{h(\mathbf{t})}{h_0} \quad (40)$$

where $h(\mathbf{t})$ and h_0 are membrane thicknesses respectively in the deformed and undeformed configurations.

The memory integral constitutive equation can be easily extended to incorporate temperature changes which may occur during the thermoforming history.

The adopted strategy for the computer implementation of the finite element formulation, developed above for the computation of the nodal displacements $\bar{\mathbf{u}}_{n+1}$ at time t_{n+1} , see the reference (Erchiqui & al, 2001).

7. Numerical validation: equation of state effect in free inflation of membrane

The dynamic finite element method outlined in the previous section was implemented in the general purpose finite element code ThermoForm, developed by the author. This code was developed in order to study the stresses and deformation arising in thermoforming sheet and stretch-blow molding problems. All computations were performed on a PC in single precision.

For numerical modeling, we considered for the outside pressure forces, induced by the airflow used for blowing the membrane, the three expressions associated respectively of the ideal gas equation of state (17), the van der Waals equation of state (18) and Redlich-Kwong equation of state (19). The rheological parameters used for the Lodge law of behaviour are found in Table 1, for HDPE-6200 material.

Lodge Model Constants at Reference Temperature	
g_k (MPa)	τ_k (second)
0.21177	0.007140
0.07837	0.049970
0.03165	0.349854
0.01122	2.448980
0.00324	17.14286
0.00117	120.0000
WLF parameters	
T_{ref} ($^{\circ}\text{C}$)	150.0
C_1	2.915
C_2 ($^{\circ}\text{C}$)	50.00

Table 1. Viscoelastic material constants for HDPE-6200

As a first application, we consider the free inflation of a square membrane (length = 15.0cm) made of HDPE material; with a load in terms of a non-linear air flow rate, similar to what is found in the practical applications (see Fig.1). The material temperature is supposed to be constant and equal to 140 $^{\circ}\text{C}$. Blowing time is fixed at 1.0s. The sheet is discretized by triangular membrane elements and its sides are fixed. The numerical analysis has been performed in two cases: i) initial sheet thickness, h_0 , was 1.50 mm and ii) initial sheet thickness was 3.00 mm.

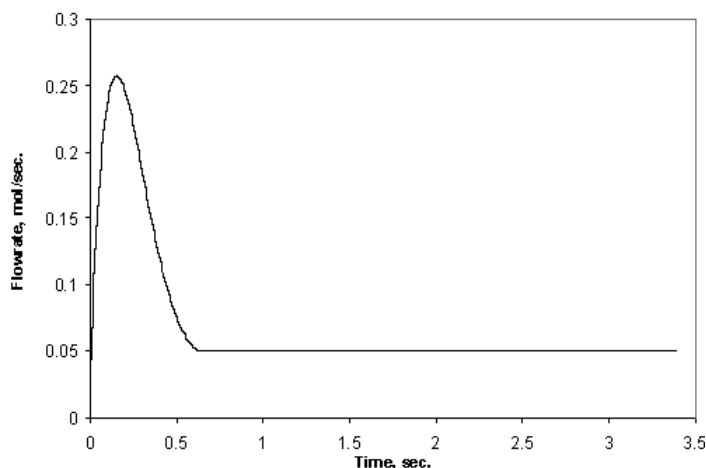


Fig. 1. Air flow loading function

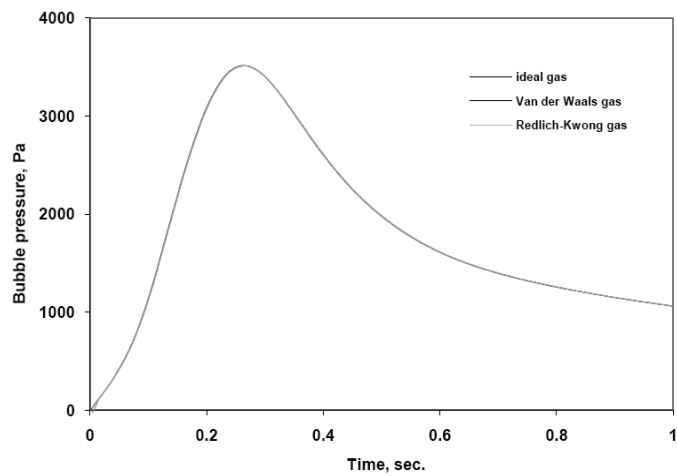


Fig. 2. Bubble pressure *vs* times, $h_0=1.5$ mm

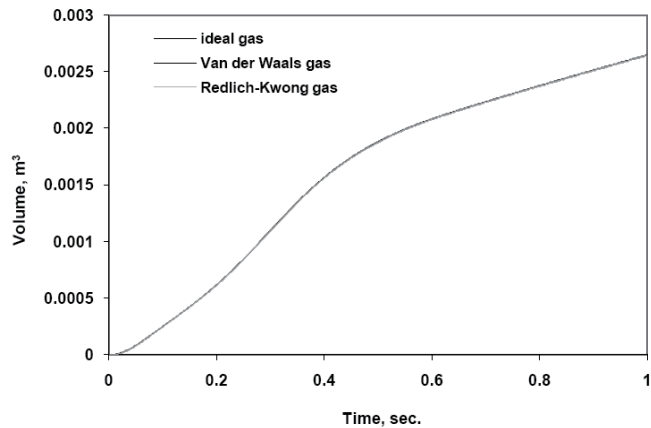


Fig. 3. Bubble volume *vs* times, $h_0=1.5$ mm

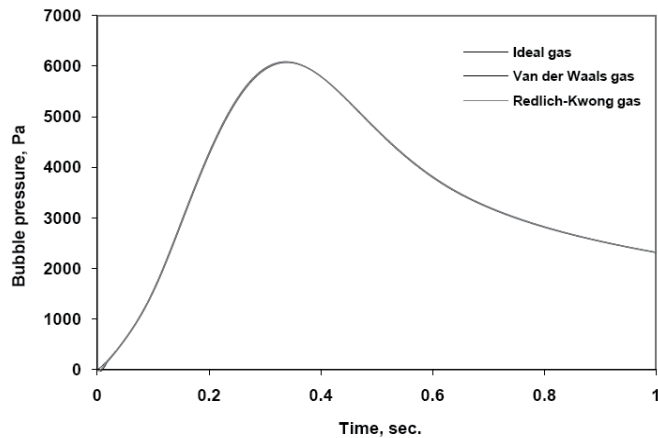


Fig. 4. Bubble pressure *vs* times, $h_0=3.0$ mm

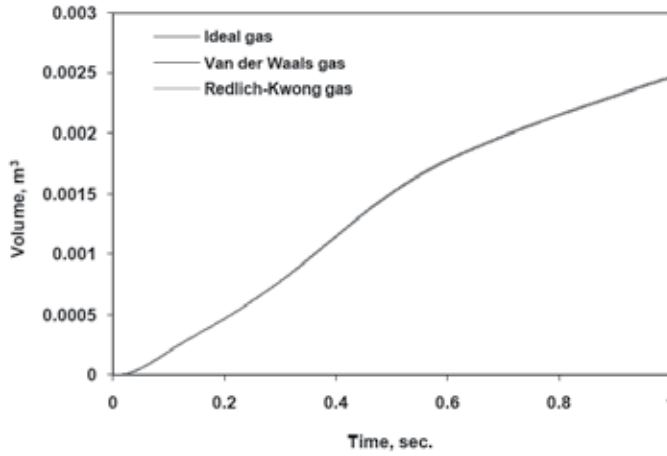


Fig. 5. Bubble volume *vs* times, $h_0=3.0$ mm

Figures 2-3 illustrate, respectively, the evolution of the internal thermoplastic pressure and volume *vs* time, obtained by the Thermoform finite element code, in the first case ($h_0 = 1.50$ mm). We notice that the three equations of states predict similar results. The absolute relative average error between the ideal gas equation of state and the two other models is lower than 0.15%. Even remarks in the second case where $h_0 = 3$ mm (see Figures 4-5).

8. Characterization

Deformation of a flat polymer sheet clamped around its edges into a 3D shape is the main feature of the thermoforming process. Generally the deformation is of a non-uniform multi-axial type and takes place at a forming temperature above the glass transition temperature. In this work, the bubble inflation technique (Joye & *al*, 1972) is used for the purpose of fitting the constitutive models used in the numerical simulation of the thermoforming process. The experimental set up has been described elsewhere (Derdouri & *al*, 2000).

The determination of hyperelastic and viscoelastic constants is performed with a biaxial characterization technique (Joye & *al*, 1972). This technique consists in a controlled free inflation of a sheet where the air flow rate is controlled while the forming pressure and sheet height are measured. The sheet is considered thin enough to be modeled by the following axisymmetric membrane inflation relationship:

$$\frac{d\bar{u}(R)}{dR} = \bar{F}(\bar{u}(R), R) \quad (41)$$

and

$$\bar{u}(R) = \{\lambda_1(R), \lambda_2(R), \theta(R), z(R), P(R)\} \quad (42)$$

where λ_1 and λ_2 are the two stretch ratios in meridian and circumferential directions respectively. θ , R and z represent the angular, circumferential and vertical position respectively. The internal pressure is given by P_e and is considered uniform on the membrane surface.

Theoretical material constants are obtained by minimizing the global absolute error, E , between the computed pressure, P_c , and the measured pressure, P_e , at each time step:

$$E = \sum_{i=1}^{N_{\text{exp}}} [P_c(t_i) - P_e(t_i)]^2 \quad (43)$$

where N_{exp} represent the number of experimental points represents. The problem of minimization of the error is solved by an algorithm of Levenberg-Marquardt (Marquardt, 1963), (Levenberg, 1944). The system describing the balance of the blowing membrane is solved by the finite differences method with differed corrections (Dennis & Schabel, 1983).

8.1 Numerical validation

As a validation of the formulation described above, we consider the problem of free blowing of a viscoelastic membrane. The material considered in this stage of work is acrylonitrile-butadiene-styrene (ABS). The initial sheet thickness was 1.57 mm. The exposed circular domain of radius $R_0=3.175$ cm is heated to the softening point inside a heating chamber using infrared heaters. When the temperature was quite uniform over the flat sheet, the inflation was started using compressed air at a controlled flow rate. In most inflation tests the experiment ended when the bubble burst. The bubble pressure, its height at the pole and time are recorded simultaneously using a video camera and a data acquisition system.

The measured pressure was interpolated by a polynomial function (Erchiqui & al, 2001) and used in the finite element program to compute the time evolution of the bubble height at the pole. The predictions obtained with Ogden Mooney-Rivlin and Lodge models give almost identical results and are shown against experimental data in Figure 6, at a setting temperature of 143°C. We notice that the the Mooney-Rivlin model, the Ogden model and the Lodge model all predict similar bubble height at the pole, which is close to the experimental measurements. However, the three models give different results, however close to the experimental measurements, when the bubble pressure level is close to the maximum pressure reached during experiments (Erchiqui & al, 2001).

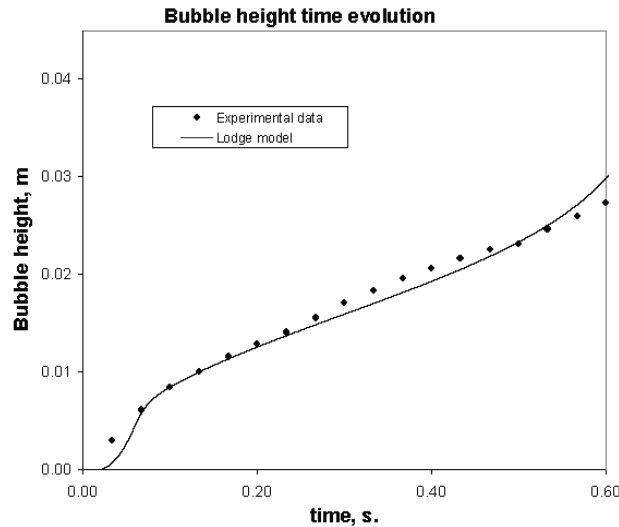


Fig. 6. Bubble height time evolution

8.2 Thermoforming application

Like the circular viscoelastic membrane blowing application of the previous section, we also use in this section, the dynamic approach with a load applied in term of linear air flow rate, to study the thermoforming of a container made of ABS material. The geometries of the mold and the parison, with a grid using the triangular elements membranes (936 nodes and 1756 elements), are presented on figure 7. The initial membrane configuration is a rectangular sheet with a length of 25.4 cm, a width of 15.24 cm and a uniform thickness of 0.16 cm and its edges are considered as fixed. In this example, only the influence of the Lodge and the Christensen constitutive models on the thickness, in the thermoforming sheet, is analyzed. However, in the paper, we present a comparative analysis on the thickness and the stress distribution in the thermoforming sheet for the viscoelastic behavior (Lodge, Christensen) and hyperelastic (Mooney-Rivlin, Ogden). The rheological parameters used for these materials can be found in (Erchiqui & *al*, 2001).

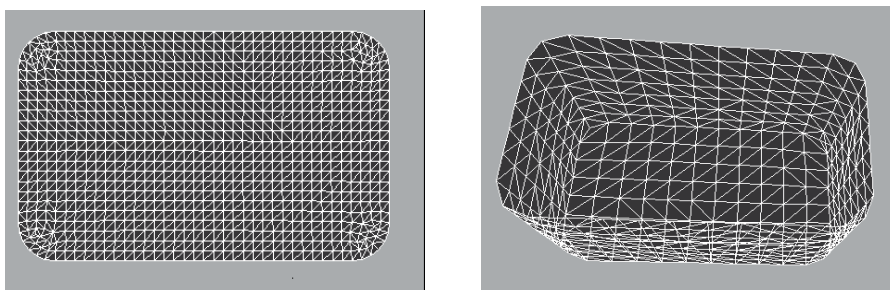


Fig. 7. Parison mesh and mesh mould

These parameters were obtained using the biaxial technique of material parameter identification (Derdouri & *al*, 2000), for a working temperature of 143°C. For the contact of the preform and the mold, we considered the assumption of sticking contact, because it is estimated that the polymer cools and stiffens quickly (at the time of the contact between the thermoplastic parison, which is very hot, and molds it, which is cold) and that the pressure of working is not sufficient to deform the part of the membrane which is in contact with mold (DeLorenzi & Nied, 1991). The time of blowing (time necessary so that the last node of the parison returns in contact with the mold), for the air flow used by our simulation, according to our numerical calculations, is about 1 second for the Lodge material, 1.15 second for the Christensen material, 0.76 second for the Ogden material and 0.72 second for the Mooney-Rivlin material.

In figure 8-9, we presented, for the models (Lodge, Christensen, Ogden Mooney Rivlin) the final thickness distribution $h=h_0\lambda_3$ (h_0 is the thickness in the undeformed configuration) on the half-planes of symmetry XZ and YZ in the thermoformed container.

A comparison study of the numerical results obtained, for the thermoforming of the ABS thermoplastic part, shows that there is a small difference in the results predicted by our modeling calculations for the four behavioral constitutive models. Indeed, the ABS thermoplastic membrane was characterized for the viscoelastic models of Lodge, Christensen, Mooney-Rivlin and Ogden. It thus follows, that theoretically the use of the rheological parameters obtained by biaxial identification (Erchiqui & *al*, 2001), in modeling, must, a priori, lead to the same final thickness distribution on the half-planes XZ and YZ of symmetry.

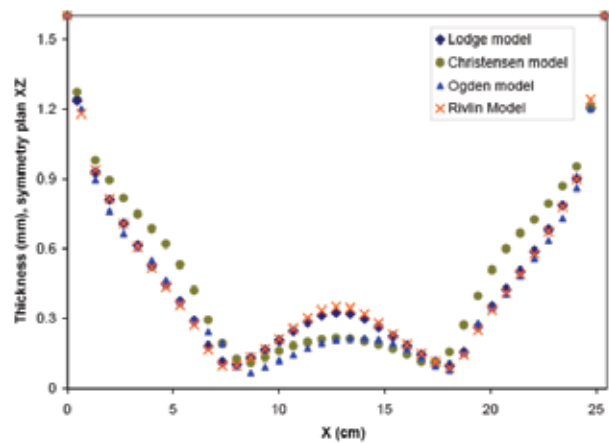


Fig. 8. Thickness distribution, symmetry plan XZ

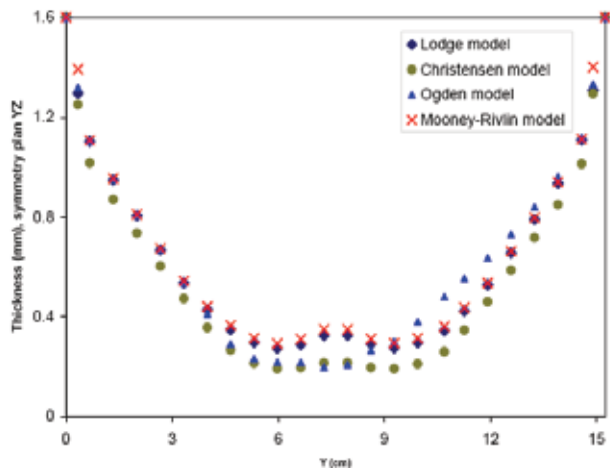


Fig. 9. Thickness distribution, symmetry plan YZ

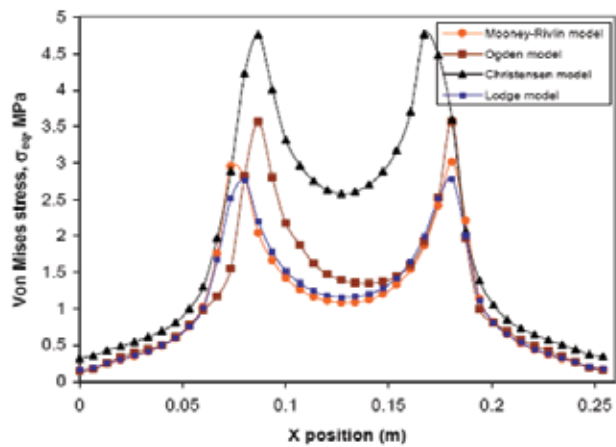


Fig. 10. Von Mises stress distribution, symmetry plan XZ

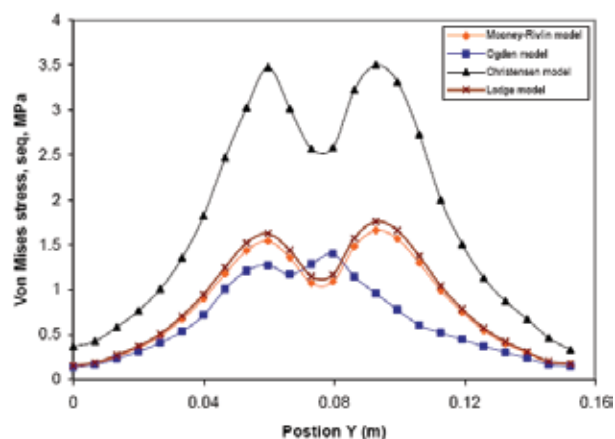


Fig. 11. Von Mises stress distribution, symmetry plan YZ

However numerical calculation shows that there is a difference, not very significant, in the thicknesses distribution (figures 8-9). However, this error can be improved perhaps, either by refining the grid, but at the detriment of a higher computing time, or by improving the technique of biaxial characterization, by the use of a higher number of air flows, as it was underlined in (Erchiqui & al, 2001).

In thermoforming numerical simulation, the thickness prediction is an important goal but the stress estimation is also helpful for part design. Indeed, the prediction of the residual stress and the shape stability of the part are strongly related to the estimated stress. In this section, the stress prediction obtained from the investigated constitutive models is discussed. The localized thinning effect of the deformed membrane is generally accompanied by the increase in the Cauchy stresses or the true stresses of the material. Figures 10-11 present the final von Mises stresses σ_{eq} distribution, predicted by using different constitutive models, on the XZ and YZ half-planes of symmetry in the thermoformed container. For clarity of presentation, the critical values of the von Mises stresses in side zone (A zone) are presented in Tables 2 and 3. The von Mises stress distribution σ_{eq} on the XZ half-plane exhibits the maximum in side zone and the minimum in center zone (B zone). These values are less significant in the half-plane YZ. A comparative study of the numerical results, obtained from the different constitutive models, shows that there is a significant difference between the Christensen model and the other models, for the von Mises stress. The maximum value is obtained for the Christensen model and the minimal value is obtained for the Mooney-Rivlin model. In the XZ half-plane, for zone A, the difference is 1.98 Mpa and for zone B, it is 1.50 Mpa. On the other hand, the difference between the Lodge and the Mooney-Rivlin models is very small. Finally, the Von Mises stress distribution and the localized thinning effect indicates that material failure due to large deformation induced by inflation is most likely to occur in zone A of the container.

Zone	Stress (MPa)	Christensen model	Lodge model	Ogden model	M. Rivlin model
A	σ_{eq}	4.76	2.78	3.57	2.78
B	σ_{eq}	2.58	1.16	1.35	1.08

Table 2. The critical values of the Von Mises stresses, symmetry half-plane XZ.

Zone	Stress (MPa)	Christensen model	Lodge model	Ogden model	M. Rivlin model
A	σ_{eq}	3.50	1.75	1.40	1.66
B	σ_{eq}	2.56	1.14	1.17	1.07

Table 3. The critical values of the Von Mises stresses, symmetry half-plane XZ.

The simple application example of thermoforming of a hollow thin part made of ABS material, shows the advantage of using the dynamic finite elements method based on a total Lagrangian approach, and of using a load in pressure derived from the thermodynamics law of perfect gases, to simulate the structural behavior of hyperelastic and viscoelastic materials.

9. Conclusion

In this work, we have presented the application of a dynamic finite element approach based on the total lagrangian formulation for simulating the response of isotropic, incompressible thermoplastic materials during thermoforming process. The forming load function is defined in terms of gas flow rate instead of static pressure. The validation of the developed finite element software is performed for an ABS material, in the case of free blowing of a circular membrane subjected to a pressure load distribution. Moreover, we have simulated the thermoforming of a rectangular container made of ABS material and studied the influence of hyperelastic (Ogden, Mooney-Rivlin) and viscoelastic (Lodge, Christensen) constitutive laws on the thickness distribution of this thin part, by varying the air flow loading distribution.

10. References

- Christensen, R. M. (1980), *Journal Applied Mechanics*. ASME, Vol.47, pp.762-768
- DeLorenzi, H. G. & Nied, H. F. (1991), *Progress in polymer processing*, Hanser Verlag, Vol.1, pp.117-171
- Dennis, J. E.; Schnabel, R. B. (1983), *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Prentice-Hall, Englewood Cliffs
- Derdouri, A.; Erchiqui, F.; Bendada, H & Verron, E. (2000), *Viscoelastic Behaviour of Polymer Membranes under Inflation*, *XIII International Congress on Rheology*, pp.394-396, Cambridge
- Erchiqui, F. (2008), *Journal of Reinforced plastics and composites*, Vol.27, No.5, pp.487-505
- Erchiqui, F.; Derdouri, A.; Gakwaya, A. & Verron, E. (2001), *Entropie*, 235/236, pp.118-125
- Joye, D. D.; Poehlein, G. W. & Denson, C. D. (1972), *Transactions of the Society of Rheology*, Vol.16, pp.421-445
- Khayat, R. E.; Derdouri, A. & Garcia-Réjon, A. (1992), *International Journal of Solids and Structures*, Vol.29, No.1, pp.69-87
- Landau, D. L. & Lifshitz, E. M. (1984). *Cours de Physique Théorique*, Tome V, *Physique Statistique*, Éditions MIR, Moscou
- Lapidus, L. & Pinder, G. F. (1982), *Numerical solution of partial differential equations in science and engineering*, John Wiley & Sons, New-Yorkk, NY

- Laroche, D. & Erchiqui, F. (2000), *Journal of Reinforced plastics and composites*, Vol.19, No.3, pp.231-239
- Laroche, D. & Erchiqui, F. (1998). *Simulation of materials processing: Theory, methods and applications*. Huétink & Baaijens (eds), Rotterdam: Balkema, ISBN 90 5410 970 X, pp.483-488
- Levenberg, K. (1944), An Algorithm for Least-Squares Estimation of Nonlinear Parameters, *Quarterly of Applied Mathematics*, Vol.2, pp.164-168
- Lodge, A. S. (1964), *Elastic liquids*. Academic Press, London
- Marquardt, S. (1963), An Algorithm for Least-Squares Estimation of Nonlinear Parameters, *SIAM Journal on applied Mathematics*, Vol.11, pp.431-441
- Ogden, R. W. (1972), Large Deformation Isotropic Elasticity - On the Correlation of Theory and Experiment for for Incompressible Rubberlike Solids, *Proceedings of the Royal Society of London A.*, A326, pp.565-565
- Rivlin, R. S. (1948), Some topics in finite elasticity I. Fundamental concepts, *Philosophical Transactions of the Royal Society A*, A240, pp.459-490
- Verron, E.; Marckmann, J & Peseux, B. (2001), *International Journal for Numerical Methods in Engineering*, Vol.50, No.5, pp. 1233-1251

Finite element analysis and Fracture in viscoelastic materials by M0v integral-Part I: crack initiation

Rostand Moutou Pitti¹, Frédéric Dubois² and Mustapha Taazount¹

¹*Clermont Université, Blaise Pascal University, LaMI, EA 3867, BP 206, 63000 Clermont
Ferrand, France*

²*Université de Limoges, GEMH, Centre Universitaire de Génie Civil, 19300 Egletons,
France*

1. Introduction

Since few years, the finite element method appears a good tool for study the mechanical behaviour of mechanical or civil engineering structures submitted to complex high level of loading and environmental effects. Also, timber elements, like notched beams or joints, are generally subject to complex crack kinetics principally due to loading modes and orthotropic characters. In order to predict the initiation and the crack growth process, many numerical tools have been developed providing the mechanical field characterisation in the crack neighbourhood. Among them, it shown that energy methods are based on the use of unvarying integrals providing the evaluation of the energy release rate in accordance with a thermodynamic approach. However, some of these tools are mathematically limited to simple or global fracture modes for isotropic or orthotropic media.

This chapter book deals with the conservative law method (Noether, 1918), based on a non-dependent path integral providing the mixed mode separation with an expensive finite element discretisation. In the literature, Bui et al., (1985) has proposed a generalization of the J-integral (Rice, 1968) by separating displacement fields into a symmetric and antisymmetric parts. This method is efficient but requires a symmetric mesh in the crack tip vicinity. Then, Chen & Shield, (1977) have developed the M-integral adapted to isotropic and elastic material, extended to orthotropic media by Moutou Pitti et al., (2007). This method allows the mixed mode fracture separation by using a virtual work principle introducing virtual fields in the integral definition. In order to introduce viscoelastic properties, the generalization of the M-integral for orthotropic material is investigated.

In order to develop a complete fracture mechanic algorithm, the first section reminds viscoelastic behaviour generalized for orthotropic configurations. The hereditary integral is transformed in an incremental formulation adapted for a finite element resolution. Since the fracture process is based on an energy balance, the numerical algorithm integrates the Helmholtz's free energy density concept.

The finite element implementation is presented in the second section. It is based on the incremental formulation treated by the virtual work principle. At each time increment, the subroutine enables us to access to the total mechanical field histories. An additional subroutine allows the time computation of the Helmholtz's free energy density requested to evaluate the energy release rate using invariant integrals.

The following part deals with a review of the thermodynamic formulation usually employed in the viscoelastic behaviour description and energetic balance. The main objective of this part is to recall different invariant integrals used in the energy release rate evaluation. More precisely, this section focus on the development of the $M\theta$ -integral concept for, firstly, orthotropic media and, in the other hand, for viscoelastic behaviour. An additional incremental formulation is proposed in order to compute, step by step the energy release rate evolution versus time by separating open and shear mode parts.

Validations are proposed in the last section for cracked orthotropic and viscoelastic media. Based on a Compact Tension Shear geometry the algorithm validation is separated in two parts. Firstly, it is demonstrated the non-dependence property of the $M\theta$ -integral. In a second time, numerical results are compared to analytic developments.

2. Viscoelastic formulation

The viscoelastic behavior is characterized by a time hereditary relationship between stresses and strains. A good understanding of the theory leads to present, firstly, a uniaxial development. Then, according to time evolutions of stress and strain scalar $\sigma(t)$ and $\varepsilon(t)$, respectively, the behavior law is described by a Boltzmann's integral :

$$\varepsilon(t) = \int_0^t J(t-\tau) \cdot \frac{\partial \sigma}{\partial \tau} d\tau \quad (1)$$

$J(t-\tau)$ is the time creep function in which t and τ are actual delayed times, respectively.

Since several years, this formulation is implemented in the finite element method allowing a mechanical field definition and energy interpretations. The finite element implementation, of the hereditary integral (1), requests to develop memorization techniques for mechanical field history. In this context, Zienkiewicz et al. (1968) have been the first author proposing a direct time integration . However, this method necessitates storing the complete past histories of strain and stress tensors step by step inducing a very quick computer memory saturation. An alternative technique has been proposed consisting to replace the complex past history by a simplified form considering a past creep loading according to an equivalent creep time. If this method is a good compromise for to reduce the calculus price, the equivalent time doesn't be calculated for several creep functions according to anisotropic or orthotropic symmetries. In order to overcome these technical limits, specific incremental formulations have been developed. For isotropic media, Mansuero (1993) has proposed a pseudo uni axial technique using, firstly, a time incremental formulation based to a Prony's series representation of creep property and, in the other hand, a decomposition of the tridimensional behavior in terms of spherical and deviatoric parts. This method leads to transform a tridimensional behavior into two uni axial behaviors according to linear assumptions and the superposition principle. This method is quite efficient by replacing all past mechanical history by an only thermodynamic variable updated at each time

increment. Dubois et al. (1999) have proposed a generalization of this approach for anisotropic and orthotropic behaviors. In the same time, the energy balance requests to define the free energy evolution versus time. By using a spectral decomposition of of creep functions, suggested by Mandel (1966), the method enables us the definition of specific internal variables allowing specific free energy potential definition. For a good understanding of the approach let us present the method for a uni axial behavior; the three-dimensional generalization being developed later.

2.1 Incremental formulation

The spectral decomposition method consists on a creep function representation by a Prony's seri. According to specific compliance constants $J^{(0)}$ and $J^{(m)}$, the creep tensor admits the following form :

$$J(t) = J^{(0)} + \sum_{m=1}^N J^{(m)} \cdot (1 - e^{-\lambda^{(m)} \cdot t}) \quad (2)$$

in which $\lambda^{(m)}$ are functions of specific relaxation time. This representation is directly compatible with a generalized Kelvin Voigt model shown in Figure 1 composed by springs and dashpots according to stiffnesses $k^{(0)}$, $k^{(m)}$ and viscosities $\eta^{(m)}$, respectively, such as

$$J^{(0)} = \frac{1}{k^{(0)}}, J^{(m)} = \frac{1}{k^{(m)}} \text{ and } \lambda^{(m)} = \frac{k^{(m)}}{\eta^{(m)}} \quad (3)$$

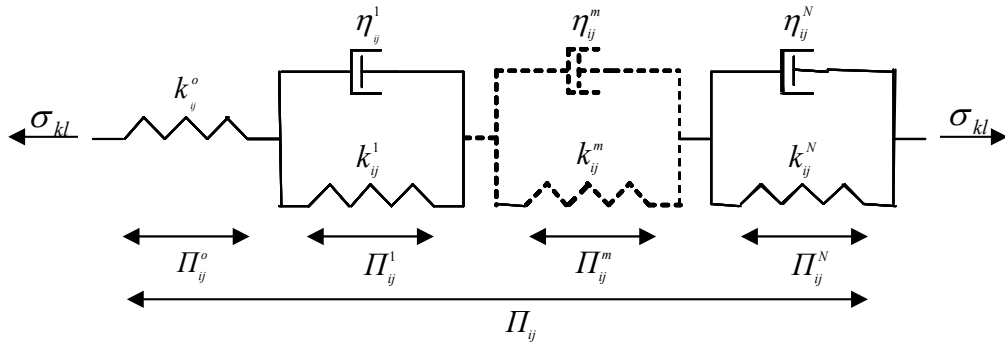


Fig. 1. Spectral decomposition of the strain tensor

Its form allows the strain separation such as

$$\varepsilon(t) = \varepsilon^{(0)} + \sum_{m=1}^N \varepsilon^{(m)} \quad (4)$$

$\varepsilon^{(0)}$ is the instantaneous and elastic strain. $\varepsilon^{(m)}$ is an differed strain part.

The time derivation of (4) is defined by

$$\frac{\partial \varepsilon(t)}{\partial t} = \frac{\partial \varepsilon^{(0)}}{\partial t} + \sum_{m=1}^N \frac{\partial \varepsilon^{(m)}}{\partial t} \quad (5)$$

The instantaneous strain rate is expressed by

$$\frac{\partial \varepsilon^{(0)}}{\partial t} = \frac{1}{k^{(0)}} \cdot \frac{\partial \sigma}{\partial t} \quad (6)$$

For the m^{th} series term, the relationship between stress σ and strain $\varepsilon^{(m)}$ is governed by the following differential equation

$$\sigma = k^{(m)} \cdot \varepsilon^{(m)} + \eta^{(m)} \cdot \frac{\partial \varepsilon^{(m)}}{\partial t} \quad (7)$$

The time integration of equations (6) and (7) request a time discretization the incremental formulation requests a time discretization. If I indicates the full number of the considered time increment ($n \in \{1; \dots; I\}$), a time function $\alpha(t)$ is employed by supposing its linearity during the time step Δt_n ($\Delta t_n = t_n - t_{n-1}$). With this restriction, we consider, at the time t_n , following notations

$$\Delta \alpha_n = \alpha_n - \alpha_{n-1}, \quad \frac{\partial \alpha_n}{\partial t} = \frac{\Delta \alpha_n}{\Delta t} \quad \text{and} \quad \alpha_n = \alpha(t_n) \quad (8)$$

With the precedent expressions, equation (5) becomes, according to relation (6) :

$$\Delta \varepsilon_n = \frac{1}{k^{(0)}} \cdot \Delta \sigma_n + \sum_{m=1}^N \Delta \varepsilon_n^{(m)} \quad (9)$$

The problem is finally reduced to the evaluation of $\Delta \varepsilon_n^{(m)}$. In this case, the differential equation with second member and constant coefficient (7) is resolved by the constant variation method. We have finally

$$\Delta \varepsilon_n^{(m)} = \left(e^{-\lambda^{(m)} \cdot \Delta t_n} - 1 \right) \cdot \varepsilon_{n-1}^{(m)} + \frac{1}{\eta^{(m)}} \cdot e^{-\lambda^{(m)} \cdot t_n} \cdot \int_{t_{n-1}}^{t_n} \sigma \cdot e^{-\lambda^{(m)} \cdot t} dt \quad (10)$$

However, supposing a linearity of σ between t_{n-1} and t_n :

$$\forall t \in \{t_{n-1} \dots t_n\}, \quad \sigma(t) = \sigma_{n-1} + \frac{t - t_{n-1}}{\Delta t_n} \cdot \Delta \sigma_n \quad (11)$$

By introducing the relation (11) in the formulation (10), we obtain

$$\begin{aligned} \Delta \varepsilon_n^{(m)} = & \left(e^{-\lambda^{(m)} \cdot \Delta t_n} - 1 \right) \cdot \varepsilon_{n-1}^{(m)} + \left(\frac{1}{k^{(m)}} \cdot \left(1 - e^{-\lambda^{(m)} \cdot \Delta t_n} \right) \cdot \sigma_{n-1} \right) \\ & + \frac{1}{k^{(m)}} \cdot \left(1 - \frac{1}{\lambda^{(m)} \cdot \Delta t_n} \cdot \left(1 - e^{-\lambda^{(m)} \cdot \Delta t_n} \right) \right) \cdot \Delta \sigma_n \end{aligned} \quad (12)$$

Finally, introducing expression (12) in (9), the strain increment can be written as follow

$$\Delta \varepsilon_n = \Psi_n \cdot \Delta \sigma_n + \tilde{\varepsilon}_{n-1} \quad (13)$$

Ψ_n represents the equivalent compliance traducing the effect of stress variation on the strain during the time increment Δt_n , defined by :

$$\Psi_n = \frac{1}{k^{(0)}} + \sum_{m=1}^N \frac{1}{k^{(m)}} \cdot \left(1 - \frac{1}{\lambda^{(m)} \cdot \Delta t_n} \cdot \left(1 - e^{-\lambda^{(m)} \cdot \Delta t_n} \right) \right) \quad (14)$$

$\tilde{\varepsilon}_{n-1}$ allows to update the material strain history. Actualized to each step of calculation, it's defined by

$$\tilde{\varepsilon}_{n-1} = \sum_{m=1}^N \left(\frac{1}{k^{(m)}} \cdot \left(1 - e^{-\lambda^{(m)} \cdot \Delta t_n} \right) \right) \cdot \sigma_{n-1} + \sum_{m=1}^M \left(e^{-\lambda^{(m)} \cdot \Delta t_n} - 1 \right) \cdot \varepsilon_{n-1}^{(m)} \quad (15)$$

2.2 3D generalizations

In order to generalize the uni axial incremental formulation, let us introduce the superposition principle considering pseudo several uni axial sollicitations in which only the stress component σ_{kl} is non zero. Hence, the correspondent strain tensor Π_{ijkl} is separated in different parts (Ghazlan et al., 1995) (Dubois & Petit, 2005)

$$\Pi_{ijkl} = \Pi_{ijkl}^{(0)} + \sum_{m=1}^N \Pi_{ijkl}^{(m)} \quad \text{avec } m = (1; \dots; N) \quad (16)$$

The behavior linearity allows to generalize the form (16) by employing the superposition principle such as

$$\varepsilon_{ij} = \sum_{k,l} \Pi_{ijkl}, \quad \varepsilon_{ij}^{(0)} = \sum_{k,l} \Pi_{ijkl}^{(0)} \quad \text{and} \quad \varepsilon_{ij}^{(m)} = \sum_{k,l} \Pi_{ijkl}^{(m)} \quad (17)$$

According to the generalized Kelvin Voigt model, the spectral decomposition of the creep tensor J_{ijkl} can be described by the association of springs $k_{ijkl}^{(p)}$ ($p \in \{0; 1; \dots; N\}$) and dashpots $\eta_{ijkl}^{(p)}$ ($m \in \{1; \dots; N\}$) such as

$$J_{ijkl}(t) = J_{ijkl}^{(0)} + \sum_{m=1}^N J_{ijkl}^{(m)} \cdot \left(1 - e^{-\lambda_{ijkl}^{(m)} \cdot t} \right) = \left[\frac{1}{k_{ijkl}^{(0)}} + \sum_{m=1}^N \frac{1}{k_{ijkl}^{(m)}} \cdot \left(1 - e^{-\lambda_{ijkl}^{(m)} \cdot t} \right) \right] \quad (18)$$

with $\lambda_{ijkl}^{(m)} = \frac{k_{ijkl}^{(m)}}{\eta_{ijkl}^{(m)}}$

In the time increment $\Delta t_n = t_n - t_{n-1}$ defined by the equation (8), by adding equation (13), the constitutive incremental law takes the following notation, without summation on the indices kl (Ghazlan et al. 1995)

$$\Delta \Pi_{ijkl}(t_n) = \Psi_{ijkl} \cdot \Delta \sigma_{kl}(t_n) + \tilde{\Pi}_{ijkl}(t_{n-1}) \quad (19)$$

$\Delta \Pi_{ijkl}(t_n)$ and $\Delta \sigma_{kl}(t_n)$ represent the increments of Π_{ijkl} and σ_{kl} respectively. $\tilde{\Pi}_{ijkl}(t_{n-1})$ is the pseudo stress at the time t_{n-1} witness of the influence of the strain past history in the various Kelvin Voigt cells. It takes the following notation:

$$\tilde{\Pi}_{ijkl}(t_{n-1}) = \sum_{m=1}^N \tilde{\Pi}_{ijkl}^{(m)}(t_{n-1})$$

$$\text{with } \tilde{\Pi}_{ijkl}^{(m)}(t_{n-1}) = \left(e^{-\lambda_{ijkl}^{(m)} \cdot \Delta t_n} - I \right) - \left[\frac{\sigma_{kl}(t_{n-1})}{k_{ijkl}^{(m)}} - \Pi_{ijkl}^{(m)}(t_{n-1}) \right] \quad (20)$$

Ψ_{ijkl} is the component of the viscoelastic four order tensor compliance Ψ computing according to is actual uni axial form in the interval time Δt_n , it is noted by

$$\Psi_{ijkl} = \frac{I}{k_{ijkl}^{(0)}} + \sum_{m=1}^N \frac{I}{k_{ijkl}^{(m)}} \cdot \left(I - \frac{I}{\lambda_{ijkl}^{(m)} \cdot \Delta t_n} \cdot \left(I - e^{-\lambda_{ijkl}^{(m)} \cdot \Delta t_n} \right) \right) \quad (21)$$

The superposition principle allows obtaining the three-dimensional incremental law

$$\Delta \varepsilon_{ij}(t_n) = \Psi_{ijkl} \cdot \Delta \sigma_{kl}(t_n) + \tilde{\varepsilon}_{ij}(t_{n-1}) \quad (22)$$

$$\text{with } \tilde{\varepsilon}_{ij}(t_{n-1}) = \sum_{k,l} \tilde{\Pi}_{ijkl}(t_{n-1}) \text{ et } \Delta \varepsilon_{ij}(t_n) = \sum_{k,l} \Delta \Pi_{ijkl}(t_n) \quad (23)$$

2.3 Helmholtz free energy density

The time fracture process in viscoelastic media is driven by energy approaches. More precisely, the energy balance puts in evidence two dissipation sources due to viscosity and crack lip separation. In these conditions, it's necessary to introduce an elastic released energy stored in the material which allows the justification of crack progression. According to specific energy definitions, this released energy is expressed as the Helmholtz free energy density F . According to notations introduced in (17) and (18), F takes the form defined as follow, (Moutou Pitti at al., 2007)

$$F = \frac{I}{2} \cdot k_{ijkl}^{(0)} \cdot \varepsilon_{ij}^{(0)} \cdot \varepsilon_{kl}^{(0)} + \sum_{m=1}^N \frac{I}{2} \cdot k_{ijkl}^{(m)} \cdot \varepsilon_{ij}^{(m)} \cdot \varepsilon_{kl}^{(m)} \quad (24)$$

$$\text{with } \varepsilon_{ij}^{(0)} = \frac{I}{k_{ijkl}^{(0)}} \cdot \sigma_{kl} \text{ et } \varepsilon_{ij}^{(m)} = \int_0^t \frac{I}{k_{ijkl}^{(m)}} \cdot \left(I - \exp^{-\lambda_{ijkl}^{(m)} \cdot (t-\tau)} \right) \cdot \frac{\partial \sigma_{kl}}{\partial \tau} d\tau \quad (25)$$

The evaluation of F , equation (24), during each time step requires the determination of the strain increments of $\varepsilon_{ij}^{(0)}$ and $\varepsilon_{ij}^{(m)}$. In this case, by coupling the relation (25) with the incremental form (22), we obtain

$$\Delta \varepsilon_{ij}^{(0)}(t_n) = \frac{I}{k_{ijkl}^{(0)}} \cdot \Delta \sigma_{kl}(t_n) \text{ et } \Delta \varepsilon_{ij}^{(m)}(t_n) = \sum_{k,l} \Delta \Pi_{ijkl}^{(m)}(t_n) \quad (26)$$

$$\begin{aligned} \Delta \Pi_{ijkl}^{(m)}(t_n) = & \left(e^{-\lambda_{ijkl}^{(m)} \cdot \Delta t_n} - I \right) \cdot \Pi_{ijkl}^{(m)}(t_{n-1}) + \left(\frac{I}{k_{ijkl}^{(m)}} \cdot \left(I - e^{-\lambda_{ijkl}^{(m)} \cdot \Delta t_n} \right) \cdot \sigma_{kl}(t_{n-1}) \right) \\ & + \frac{I}{k_{ijkl}^{(m)}} \cdot \left(I - \frac{I}{\lambda_{ijkl}^{(m)} \cdot \Delta t_n} \cdot \left(I - e^{-\lambda_{ijkl}^{(m)} \cdot \Delta t_n} \right) \right) \cdot \Delta \sigma_{kl}(t_n) \end{aligned} \quad (27)$$

3. Finite element implementation

3.1 Virtual work principle

The virtual work principle rests on an energy assessment with the free elastic energy density F and the work of the external efforts W_{ext} . In this case, supposing a virtual displacement kinematically admissible $\delta(\Delta \vec{u})$ and traducing a perturbation of a real field \vec{u} in a balance configuration at the time t_{n-1} , the principle is traduced by

$$\delta F = \delta W_{ext} \quad (28)$$

Being given that the strain and the stress fields are known at the time t_{n-1} , the problem is reduced to the determination of mechanical fields at the next time $t_n = t_{n-1} + \Delta t_n$. The virtual displacement field $\delta(\Delta \vec{u})$ of components $\delta(\Delta u_i)$, around the configuration $\Omega(t)$, induces a virtual strain $\delta(\Delta \varepsilon_{ij})$ given by the Cauchy's tensor strain :

$$\begin{aligned} \delta(\Delta \varepsilon_{ij}) &= \frac{1}{2} \cdot \left[\delta(\nabla(\Delta u)) + \delta(\nabla(\Delta u)^T) \right] \\ &= \frac{1}{2} \cdot \left[\delta(\Delta u_{i,j}) + \delta(\Delta u_{j,i}) \right], \text{ et } \delta(\Delta u_{i,j}) = \delta \left(\frac{\partial u_i}{\partial x_j} \right) \end{aligned} \quad (29)$$

The free energy density and the sum of the external work efforts are respectively traduced by

$$\delta F(t_n) = \int_{\Omega(t_n)} \sigma_{ij}(t_n) \cdot \delta(\Delta \varepsilon_{ij}) dV \quad (30)$$

$$\text{and } \delta W_{ext}(t_n) = \int_{\Omega(t_n)} f_{vi}(t_n) \cdot \delta(\Delta u_i) dV + \int_{\Omega_{fs}(t_n)} f_{si}(t_n) \cdot \delta(\Delta u_i) dS \quad (31)$$

f_{vi} are the components of the volume forces in the domain Ω at the time t_n . f_{si} are the components of the imposed surface forces on the frontier of Ω_f during the same time lapse. Now, considering the following variable change

$$\begin{aligned} \sigma_{ij}(t_n) &= \sigma_{ij}(t_{n-1}) + \Delta \sigma_{ij}^n \\ f_{vi}(t_n) &= f_{vi}(t_{n-1}) + \Delta f_{vi}^n \\ f_{si}(t_n) &= f_{si}(t_{n-1}) + \Delta f_{si}^n \end{aligned} \quad (32)$$

According to equations (28), (30), (31) and relations (32), we obtain

$$\begin{aligned} \int_{\Omega(t_n)} \left[\sigma_{ij}(t_{n-1}) + \Delta \sigma_{ij} \right] \cdot \delta(\Delta \varepsilon_{ij}) dV = \int_{\Omega(t_n)} \left[f_{vi}(t_{n-1}) + \Delta f_{vi} \right] \cdot \delta(\Delta u_i) dV \\ + \int_{\Omega_{f_s}(t_n)} \left[f_{si}(t_{n-1}) + \Delta f_{si} \right] \cdot \delta(\Delta u_i) dS \end{aligned} \quad (33)$$

The principle being checked at the time t_n and t_{n-1} , we obtain by recurrence the following expression

$$\int_{\Omega(t_n)} \Delta \sigma_{ij} \cdot \delta(\Delta \varepsilon_{ij}) dV = \int_{\Omega(t_n)} \Delta f_{vi} \cdot \delta(\Delta u_i) dV + \int_{\Omega_{f_s}(t_n)} \Delta f_{si} \cdot \delta(\Delta u_i) dS \quad (34)$$

3.2 Finite element method

The relation (34) is reconsidered in a finite element discretization domain. The nodal unknown factors $\{\Delta u\}(t_n)$ are the values of the displacement field variation at the nodes of each under field. These values are computed from the real displacements fields $\Delta u(t_n)$, according to shape functions $N(x_1, x_2, x_3)$ characterizing the employed discretization elements

$$\Delta u(t_n)(x_1, x_2, x_3) = N(x_1, x_2, x_3) \cdot \{\Delta u\}(t_n) \quad (35)$$

The variation of the strain field vector $\{\Delta \varepsilon\}(t_n)$, computing at the integration points, is determined with the nodal displacements variations and the Jacobean matrix $[B]$:

$$\{\varepsilon\}(t_n) = [B] \cdot \{\Delta u\}(t_n) \quad (36)$$

By introducing the equations (35) and (36) in the equation (34), the incremental formulation law (22) becomes

$$\begin{aligned} \int_{\Omega(t_n)} [B]^T \cdot A \cdot [B] \cdot \{\Delta u\}(t_n) dV + \int_{\Omega(t_n)} [B]^T \cdot \{\tilde{\sigma}\}(t_{n-1}) dV \\ = \int_{\Omega_{f_s}(t_n)} \{\Delta f_s\}(t_n) dS + \int_{\Omega(t_n)} \{\Delta f_v\}(t_n) dV \end{aligned} \quad (37)$$

$$\text{with } A = (\Psi_{ijkl})^{-1}(t_n) \text{ et } \tilde{\sigma}_{kl}(t_{n-1}) = -(\Psi_{ijkl})^{-1}(t_n) \cdot \tilde{\varepsilon}_{ijkl}(t_{n-1}) \quad (38)$$

$\{\tilde{\sigma}\}(t_{n-1})$ is the stress vector defined in the integration points of each calculated element starting from the stress field of components $\sigma_{kl}(t_{n-1})$. $\{\Delta f_v\}(t_n)$ and $\{\Delta f_s\}(t_n)$ are the volume and the surface nodal forces defined according to the forces of components $\Delta f_{vi}(t_n)$ and $\Delta f_{si}(t_n)$, respectively

$$\{\varepsilon\}(t_n) = [B] \cdot \{\Delta u\}(t_n) \quad (39)$$

$$\{\tilde{F}^p\}(t_{n-1}) = - \int_{\Omega} B^T \cdot \Psi^p \cdot \{\tilde{\varepsilon}^p\}(t_{n-1}) d\Omega \quad (40)$$

$$\{\Delta F_{ext}^p\}(t_n) = \int_{\Omega} \{\Delta f_s^p\}(t_n) dS + \int_{\Omega(t_n)} \{\Delta f_v^p\}(t_n) dV \quad (41)$$

K_T^P is the apparent rigidity tangent matrix in the time increment $\Delta t_n \cdot \{\tilde{F}^P\}(t_{n-1})$ is the supplementary viscous load vector which represents the complete mechanical past history until the time $t_{n-1} \cdot \{\Delta F_{ext}^P\}(t_n)$ is the increment of the external surface and volume nodal vector forces during the increment $\Delta t_n \cdot \{\tilde{\varepsilon}^P\}(t_{n-1})$ is the strain history. At the end, the introduction of equations (39), (40) and (41) in the equality (37) conduces at the following finite element balance equation allowing the calculation of the nodal displacement vector increment $\{\Delta u^P\}(t_n)$

$$K_T^P \cdot \{\Delta u^P\}(t_n) = \{\Delta F_{ext}^P\}(t_n) + \{\tilde{F}^P\}(t_{n-1}) \quad (42)$$

4. Viscoelastic fracture mechanic

The main purpose of this part is the generalization and the modeling of the static M-integral, initially proposed by Chen & Shield, (1977), to orthotropic viscoelastic behavior. In this case, the surrounding integrals given by the energetic processes and the local approaches must be recalled. The algorithm is resolve in the finite element software Castem coupling the previous incremental formulation for viscoelastic behavior and the M-integral. The main topic is the calculation of the energy release rate evolution versus time by operating a mixed mode separation in viscoelastic and orthotropic media.

4.1 Energetic method

These methods provide the evaluation of the fracture parameters far from the defect introduced by the crack tip where the mechanical fields are largely disturbed by a strong singularity. The development of these tools is resulting from the conservative laws (Noether, 1971) and non dependant integral (Bui 2007).

J integral

In linear elasticity, for plane configurations and static cracks, energy required to create new crack surfaces is defined by Rice and its J-integral, (1968)

$$J = \int_{\Gamma} \left[F \cdot n_I - \sigma_{ij} \cdot n_j \cdot \frac{\partial u_i}{\partial x_I} \right] d\Gamma \quad (43)$$

Γ is a curvilinear contour including the crack tip oriented by its normal vector \bar{n} of components n_j , Figure 2 (a).

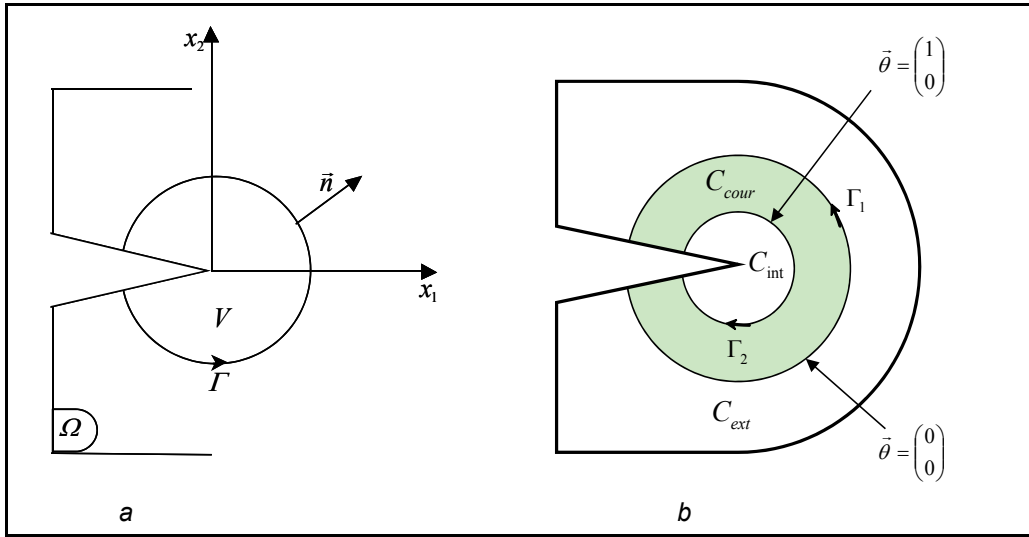


Fig. 2. a- Curvilinear line integral for the J-integral. b- θ field for the $G\theta$ -integral

$G\theta$ -integral

The J-integral, equation (43) is defined with a curvilinear line surrounding the crack tip. However, for the needs of implementation in a computer code and to ensure the fields continuity, it is preferable to define this expression on a surface integral in order to avoid the field's projections and numerical error sources. Destuynder et al. (1981, 1983) have defined a $\vec{\theta}$ vector field, continuous and differentiable allowing a virtual crown definition ($\theta_1 = 1$ inside the crown and $\theta_2 = 1$ outside it), Figure 2 (b). This vector field respects following proprieties:

- $\vec{\theta}$ is defined in the crack plan,
- $\vec{\theta}$ definition is in accordance with the crack growth direction,
- The support field $\vec{\theta}$ is concentrated in the crack tip neighborhood,

In fact, the contours Γ_1 and Γ_2 surrounding the crack tip can be defined. In we consider a crack growth in the x_1 direction, the domain is divided in three parts, Figure 3 b (Moutou Pitti et al, 2007):

- In C_{int} , $\vec{\theta}$ is constant and takes the unitary value of $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$;
- In C_{ext} , the field $\vec{\theta}$ is zero $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$;
- In C_{cour} , the field $\vec{\theta}$ vary continuously of $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ to $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ according to a constant gradient.

Hence, the energy release rate can be expressed in the following term:

$$G\theta = \int_V \left[F \cdot \theta_{k,k} + \sigma_{ij} \cdot u_{i,k} \cdot \theta_k \right] dV \quad k \in \{1, 2\} \quad (44)$$

If integrals J and $G\theta$ provide to determine an invariant leading at the mechanical state in the crack vicinity, they operate a global energy calculation independently to the mixed mode fractures. For this reason, these integrals are employed only for pure opening or pure shear fracture modes.

M0-integrals

In order to separate mixed mode fracture, Chen and Shield, (1977) have proposed the following invariant integral

$$M = \frac{I}{2} \cdot \int_{\Gamma} \left[\sigma_{ij,l}^{(v)} \cdot u_i - \sigma_{ij}^{(u)} \cdot v_{i,l} \right] \cdot n_j d\Gamma \quad (45)$$

$\sigma_{ij}^{(u)}$ and $\sigma_{ij}^{(v)}$ are the real and virtual stresses field. The particularity of the M-integral lies in the joint combination of the real u and virtual v displacements fields kinematically acceptable. As the J-integral, M is defined on a curvilinear integration domain. In accordance with the previous part, we'd rather prefer a surface contour. In this case, the $M\theta$ integral (Moutou Pitti, 2007b) is defined for plane problems and takes the following form:

$$M\theta = \frac{I}{2} \cdot \int_V \left[\sigma_{ij}^{(u)} \cdot v_{i,k} - \sigma_{ij,k}^{(v)} \cdot u_i \right] \cdot \theta_{k,j} dV \quad (46)$$

For the orthotropic materials, virtual displacements fields v are given by Sih's singular form (Sih, 1974) for each fracture mode

$$\begin{aligned} v_1 = & 2 \cdot K_I^{(\sigma)} \cdot \sqrt{\frac{r}{2 \cdot \pi}} \cdot \Re \left[\frac{I}{s_1 - s_2} \cdot (p_2 \cdot s_1 \cdot \sqrt{\rho_2} - p_1 \cdot s_2 \cdot \sqrt{\rho_1}) \right] \\ & + 2 \cdot K_2^{(\sigma)} \cdot \sqrt{\frac{r}{2 \cdot \pi}} \cdot \Re \left[\frac{I}{s_1 - s_2} \cdot (p_2 \cdot \sqrt{\rho_2} - p_1 \cdot \sqrt{\rho_1}) \right] \end{aligned} \quad (47)$$

$$\begin{aligned} v_2 = & 2 \cdot K_I^{(\sigma)} \cdot \sqrt{\frac{r}{2 \cdot \pi}} \cdot \Re \left[\frac{I}{s_1 - s_2} \cdot (q_2 \cdot s_1 \cdot \sqrt{\rho_2} - q_1 \cdot s_2 \cdot \sqrt{\rho_1}) \right] \\ & + 2 \cdot K_2^{(\sigma)} \cdot \sqrt{\frac{r}{2 \cdot \pi}} \cdot \Re \left[\frac{I}{s_1 - s_2} \cdot (q_2 \cdot \sqrt{\rho_2} - q_1 \cdot \sqrt{\rho_1}) \right] \\ \text{with } \rho_j = & \cos(\theta) + i \cdot s_j \cdot \sin(\theta) \text{ avec } j \in \{1; 2\} \end{aligned} \quad (48)$$

$$\text{and } p_j = S_{11} \cdot s_j^2 + S_{12} \text{ et } q_j = \frac{S_{22}}{s_j} + S_{12} \quad (49)$$

s_j are roots of the following characteristic equation:

$$S_{11} \cdot s_\beta^4 + (2 \cdot S_{12} + S_{33}) \cdot s_\beta^2 + S_{22} = 0 \quad (50)$$

S_{11} , S_{12} , S_{22} , S_{33} designate components of the compliance tensor in an orthotropic symmetry.

Physical interpretation of M0 integral

For real and virtual kinematically admissible displacement fields u and v , respectively, Dubois et al, 1999 have shown the following physical interpretation:

$$M(u, u) = J = G \quad (51)$$

In linear elasticity, the following physical interpretation can be easily demonstrated

$$M\theta(u, v) = C_I \cdot \frac{{}^u K_I \cdot {}^v K_I}{8} + C_2 \cdot \frac{{}^u K_{II} \cdot {}^v K_{II}}{8} \quad (52)$$

C_I and C_2 are the reduced elastic compliances in opening and shear mode allowing the estimation of elastic response between local stress field in the crack tip vicinity and the crack opening. For an orthotropic symmetry, their forms are defined by Valentin et al. (1989)

$$C_I = 4 \cdot \Re e \left[\frac{i \cdot (q_2 \cdot s_I - q_I \cdot s_2)}{s_I - s_2} \right] \text{ and } C_2 = 4 \cdot \Re e \left[\frac{i \cdot (p_2 - p_I)}{s_I - s_2} \right] \quad (53)$$

In order to express the real stress intensity factors ${}^u K_I$ and ${}^u K_{II}$, the perfect mixed mod separation is obtain by implementing two different calculations of the $M\theta(u, v)$ integral. In this case, judicious values for the virtual stress intensity factor ${}^v K_I$ and ${}^v K_{II}$ as chosen (Moutou Pitti 2008)

$${}^u K_I = 8 \cdot \frac{M({}^v K_I = I, {}^v K_{II} = 0)}{C_I} \text{ and } {}^u K_{II} = 8 \cdot \frac{M({}^v K_I = 0, {}^v K_{II} = I)}{C_{II}} \quad (54)$$

4.2 Generalization to viscoelastic behavior

The introduction of the M-integral in a viscoelastic behavior integrates the similitude between the generalized Kelvin Voigt model shown in Figure 1 and the Helmholtz free energy density which is a energy summation on different elastic elements, expression (24). In this case, the equation (45) has generalized as follow (Moutou Pitti et al 2007; Moutou Pitti, 2008)

$$M_v^{(k)} = \frac{I}{2} \cdot \int_{\Gamma} \left[\sigma_{ij,l}^{(k)}(v) \cdot u_i^{(k)} - \sigma_{ij}^{(k)}(u) \cdot v_{i,l}^{(k)} \right] \cdot n_j d\Gamma \text{ avec } k = (0, I, \dots, N) \quad (55)$$

$\sigma_{ij}^{(k)}(u)$ and $\sigma_{ij}^{(k)}(v)$ indicate the real and virtual stresses in the k^{th} spring, respectively. u_i^k and v_i^k are real and virtual displacements of this spring induced by real and virtual elastic stresses. According to the generalization of the expression (46), we obtain the M0v-integral for a viscoelastic field

$$M\theta_v^{(k)} = \frac{I}{2} \cdot \int_V \left[\sigma_{ij}^{(k)}(u) \cdot u_i^{(k)} - \sigma_{ij,k}^{(k)}(v) \cdot v_{i,k}^{(k)} \right] \cdot \theta_{k,j} dV \text{ avec } k = (0, I, \dots, N) \quad (56)$$

With the same way of equations (52) and (54), we obtain successively :

$$M\theta_v^{(k)} = C_I^{(k)} \cdot \frac{{}^u K_I^{(k)} \cdot {}^v K_I^{(k)}}{8} + C_2^{(k)} \cdot \frac{{}^u K_{II}^{(k)} \cdot {}^v K_{II}^{(k)}}{8} \quad (57)$$

$${}^u K_I^{(k)} = 8 \cdot \frac{M\theta_v^{(k)} \left({}^v K_I^{(k)} = I, {}^v K_{II}^{(k)} = 0 \right)}{C_I^{(k)}} \quad \text{and} \quad (58)$$

$${}^u K_{II}^{(k)} = 8 \cdot \frac{M\theta_v^{(k)} \left({}^v K_I^{(k)} = 0, {}^v K_{II}^{(k)} = I \right)}{C_2^{(k)}}$$

${}^v K_I^{(k)}$ and ${}^v K_{II}^{(k)}$ are virtual stress intensity factors characterizing the stress field induced respectively by $v_I^{(k)}$ and $v_2^{(k)}$ for the k^{th} spring. In the same case of equations (47), (48), (49) and (50), these virtual displacement are rewritten as

$$\begin{aligned} v_I^{(k)} = & 2 \cdot {}^v K_I^{(k)} \cdot \sqrt{\frac{r}{2 \cdot \pi}} \cdot \Re e \left[\frac{I}{s_I^{(k)} - s_2^{(k)}} \cdot \left(p_2^{(k)} \cdot s_I^{(k)} \cdot \sqrt{\rho_2^{(k)}} - p_I^{(k)} \cdot s_2^{(k)} \cdot \sqrt{\rho_I^{(k)}} \right) \right] \\ & + 2 \cdot {}^v K_{II}^{(k)} \cdot \sqrt{\frac{r}{2 \cdot \pi}} \cdot \Re e \left[\frac{I}{s_I^{(k)} - s_2^{(k)}} \cdot \left(p_2^{(k)} \cdot \sqrt{\rho_2^{(k)}} - p_I^{(k)} \cdot \sqrt{\rho_I^{(k)}} \right) \right] \end{aligned} \quad (59)$$

$$\begin{aligned} v_2^{(k)} = & 2 \cdot {}^v K_I^{(k)} \cdot \sqrt{\frac{r}{2 \cdot \pi}} \cdot \Re e \left[\frac{I}{s_I^{(k)} - s_2^{(k)}} \cdot \left(q_2^{(k)} \cdot s_I^{(k)} \cdot \sqrt{\rho_2^{(k)}} - q_I^{(k)} \cdot s_2^{(k)} \cdot \sqrt{\rho_I^{(k)}} \right) \right] \\ & + 2 \cdot {}^v K_{II}^{(k)} \cdot \sqrt{\frac{r}{2 \cdot \pi}} \cdot \Re e \left[\frac{I}{s_I^{(k)} - s_2^{(k)}} \cdot \left(q_2^{(k)} \cdot \sqrt{\rho_2^{(k)}} - q_I^{(k)} \cdot \sqrt{\rho_I^{(k)}} \right) \right] \\ \rho_j^{(k)} = & \cos(\theta) + i \cdot s_j^{(k)} \cdot \sin(\theta) \quad \text{with } j \in \{1; 2\} \quad \text{and} \\ \text{with } p_j^{(k)} = & S_{11}^{(k)} \cdot \left(s_j^{(k)} \right)^2 + \left(S_{12}^{(k)} \right)^2 \quad \text{and} \quad q_j^{(k)} = \frac{S_{22}^{(k)}}{s_j^{(k)}} + S_{12}^{(k)} \cdot s_j^{(k)} \end{aligned} \quad (60)$$

$$S_{11}^{(k)} \cdot \left(s_j^{(k)} \right)^4 + \left(2 \cdot S_{12}^{(k)} + S_{33}^{(k)} \right) \cdot \left(s_j^{(k)} \right)^2 + S_{22}^{(k)} = 0$$

By combining equations (57) and (58), we obtain the following viscoelastic energy release rate

$$G_v^{(k)} = {}^I G_v^{(k)} + {}^2 G_v^{(k)} = C_I^{(k)} \cdot \frac{\left({}^u K_I^{(k)} \right)^2}{8} + C_2^{(k)} \cdot \frac{\left({}^u K_{II}^{(k)} \right)^2}{8} \quad (61)$$

5. Local mechanical fields

In order to define the mechanical fields at the crack tip, Chazal & Dubois, (2001) and Dubois et al, (2002) have proposed, for plane problems, two viscoelastic stress intensity factors ${}^u K_\alpha^{(\sigma)}$ ($\alpha = \{1; 2\}$) and two viscoelastic opening displacement intensity factors such as, (Figure 3)

$$\sigma_{ij} = \frac{I}{\sqrt{2 \cdot \pi \cdot r}} \cdot {}^u K_\alpha^{(\sigma)} \cdot f_{ij\alpha}(\theta) \quad (62)$$

$$[u_\alpha] = \sqrt{\frac{r}{2 \cdot \pi}} \cdot {}^u K_\alpha^{(\varepsilon)} \quad (63)$$

$f_{ij\alpha}(\theta)$ is a function which depends on the local properties of material (Irwin, 1957). $[u_\alpha]$ are the components of the crack opening displacement which designates the relative displacement vector of crack lips. Considering the Boltzmann integral (1) into expressions (62) and (63), the relationship between stress and crack opening intensity factors takes the following form (Dubois et al. 1999)

$${}^u K_\alpha^{(\varepsilon)} = \int_0^t C_\alpha(t-\tau) \cdot \frac{\partial {}^u K_\alpha^{(\sigma)}}{\partial \tau} d\tau \quad (64)$$

C_α is the viscoelastic compliance function for α mode and takes a similar creep function form in accord with a generalized Kelvin Voigt model, Figure 3 :

$$C_\alpha(t) = \frac{I}{k_\alpha^{(0)}} + \sum_{m=1}^N \frac{I}{k_\alpha^{(0)}} \cdot \left(1 - \exp^{-\lambda_\alpha^{(m)} \cdot t} \right) \text{ avec } \lambda_\alpha^{(m)} = \frac{k_\alpha^{(m)}}{\eta_\alpha^{(m)}} \quad (65)$$

$k_\alpha^{(p)}$ ($p \in \{0; 1; \dots; N\}$) and $\eta_\alpha^{(m)}$ ($m \in \{1; \dots; N\}$) are the contribution of tensor components $k_{ijkl}^{(p)}$ and $\eta_{ijkl}^{(p)}$ respectively. $\frac{I}{k_\alpha^{(0)}}$ are the reduced elastic compliance defined

by the equation (53). With a strain analogy, equation (4) allows a partition of the crack opening intensity factor traduced by

$${}^u K_\alpha^{(\varepsilon)} = {}^u K_\alpha^{(0)} + \sum_{m=1}^N {}^u K_\alpha^{(m)} \quad (66)$$

$${}^u K_\alpha^{(0)} = \frac{I}{k_\alpha^{(0)}} \cdot {}^u K_\alpha^{(\sigma)} \text{ et}$$

$$\text{with} \quad \text{and } {}^u K_\alpha^{(p)} = \int_0^t \frac{I}{k_\alpha^{(p)}} \cdot \left(1 - \exp^{-\lambda_\alpha^{(p)} \cdot (t-\tau)} \right) \cdot \frac{\partial {}^u K_\alpha^{(\sigma)}}{\partial \tau} d\tau \quad (67)$$

If we take into account the definition of the Helmholtz energy density of introduced (Staverman & Schwarzl 1952), the viscoelastic energy release rate, traduced by the local proprieties at the crack tip, is (Moutou Pitti et al, 2007)

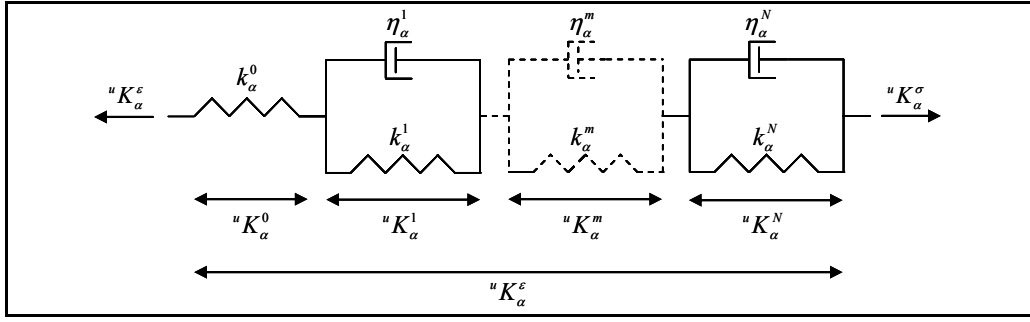


Fig. 3. Local generalized Kelvin Voigt model

$${}^{\alpha}G_v(t) = \frac{1}{8} \cdot \int_0^t \int_0^t [2 \cdot C_{\alpha}(t-\tau) - C_{\alpha}(2 \cdot t - \tau - \beta)] \cdot \frac{\partial {}^u K_{\alpha}^{(\sigma)}}{\partial \tau} \frac{\partial {}^u K_{\alpha}^{(\sigma)}}{\partial \beta} d\tau d\beta \quad (68)$$

By combining Equations (61), (67) and (68), the partition of viscoelastic energy release rate ${}^{\alpha}G_v$ is given by

$${}^{\alpha}G_v = {}^{\alpha}G_v^{(0)} + \sum_{m=1}^N {}^{\alpha}G_v^{(m)} \quad (69)$$

$$\text{with } G_v^{(p)} = \frac{1}{8} \cdot \frac{1}{k_{\alpha}^{(p)}} \cdot [{}^u K_{\alpha}^{(p)}]^2 \quad \text{and} \quad {}^{\alpha}G_v^{(0)} = \frac{1}{8} \cdot \frac{1}{k_{\alpha}^{(0)}} \cdot [{}^u K_{\alpha}^{(0)}]^2 \quad (70)$$

5.2 Numerical algorithm

This section deals with the numerical procedure implemented in the finite software Castem. The uncoupling between the viscoelastic incremental formulation and the fracture mode process is proposed. In the general subroutine, the algorithm computing the virtual mechanical fields is added (Moutou Pitti et al. 2007). In order to explain this algorithm, we suppose that mechanical fields are known at time t_{n-1} and we have fixed the time increment Δt_n . All properties of viscoelastic material are experimentally defined (Dubois et al. 2001). The different steps of the algorithm are defined as follow, Figure 4

- The tensor Ψ , expression (21) is computed, and the global tangent rigidity matrix K_T is deduced.
- The supplementary viscoelastic load field $\{F^p\}(t_{n-1})$ is determined with expression (40).
- By introducing the exterior vector force $\{\Delta F_{ext}\}(t_n)$, equation (41), in the equation (42), the nodal displacement incremental $\{\Delta u\}(t_n)$ and the different mechanical fields $\{u\}(t_n)$, $\{\varepsilon\}(t_n)$ et $\{\sigma\}(t_n)$ are obtained. The compliance tensor Ψ^P is

introduced according to the proprieties k_{ijkl}^p of the material. After, the strain tensor $\{\varepsilon^p\}(t_n)$ is calculated and the viscoelastic tangent matrix K_T^p is obtained.

- The elastic stress tensor $\{\sigma^{(p)}\}(t_n)$ and the nodal force vector $\{F^{(p)}\}(t_n)$ are calculated. For each model spring, nodal displacement vector $\{u^{(p)}\}$ are deduced of the following finite element balance equation $K_T^p \cdot \{u^{(p)}\}(t_n) = \{F^{(p)}\}(t_n)$
- Finally, the stress intensity factor $^u K_\alpha^{(p)}$ and the energy release rate $^\alpha G \theta_v^{(p)}$ evaluation necessitate the virtual displacement $v_\alpha^{(p)}$, expression (59), the viscoelastic compliance $C_\alpha^{(p)}$, and virtual stress tensor $^\alpha \sigma_{virt}^{(p)}$. $^u K_\alpha^{(p)}$ and $^\alpha G \theta_v^{(p)}$ are given by the $M\theta$ subroutine. At the end, a final summation on $^\alpha G \theta_v^{(p)}$ gives the global energy release rate for each fracture mode $^\alpha G_v$:

$$^\alpha G_v = ^\alpha G \theta_v^{(0)} + \sum_{m=1}^N ^\alpha G \theta_v^{(m)} = \sum_{m=1}^N ^\alpha G \theta_v^{(p)} \quad (71)$$

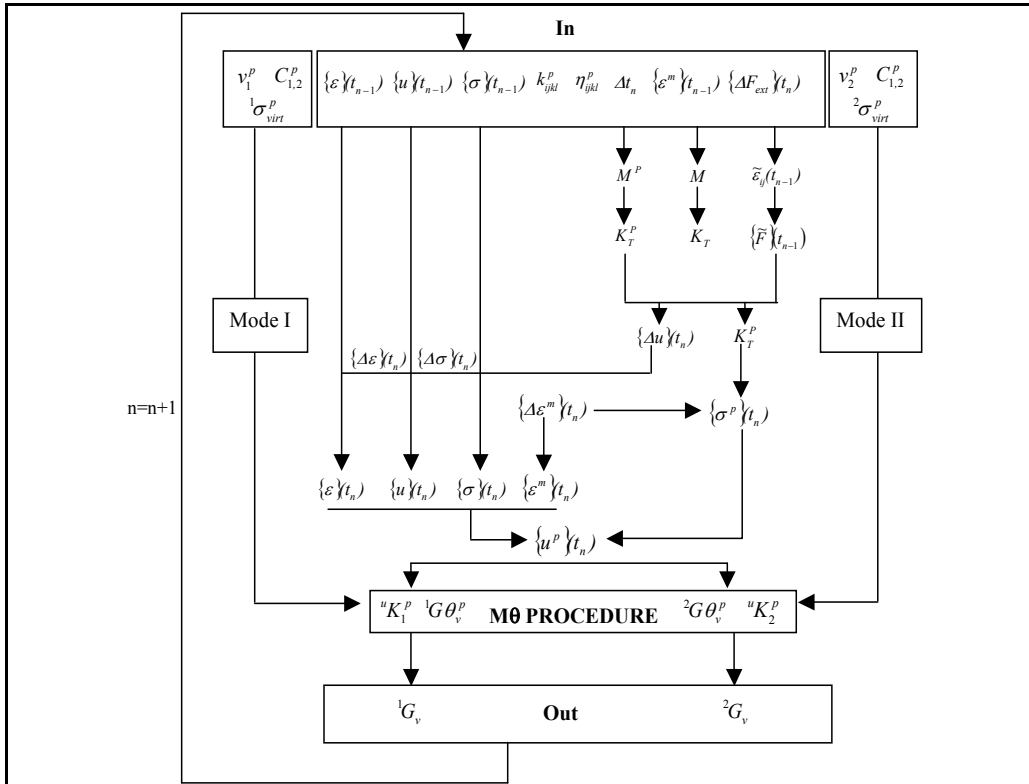


Fig. 4. Numerical subroutine

6. Numerical validation

6.1 Compact Tension Shear specimen

The CTS geometry has been initially developed by Richard, (1981) in order to separate fracture modes in isotropic materials. Valentin & Caumes (1989) have adapted this specimen to orthotropic material as wood. On Figure 5, the initial crack length chosen is 25 mm. The external load is a unitary loading applied to a perfect rigid steel arm (which presents a large crack growth zone), Figure 6. Points A_α and B_α with $\alpha \in (1...7)$ are holes where forces can be applied with the angle β oriented according to the trigonometrically direction for different mixed mode ratios. The simple opening mode is obtained by applying opposite forces in A_1 and B_1 with $\beta = 0^\circ$. The loading $\beta = 90^\circ$, in A_7 and B_7 corresponds to a simple shear mode configuration. Intermediary positions induce different mixed mode ratios.

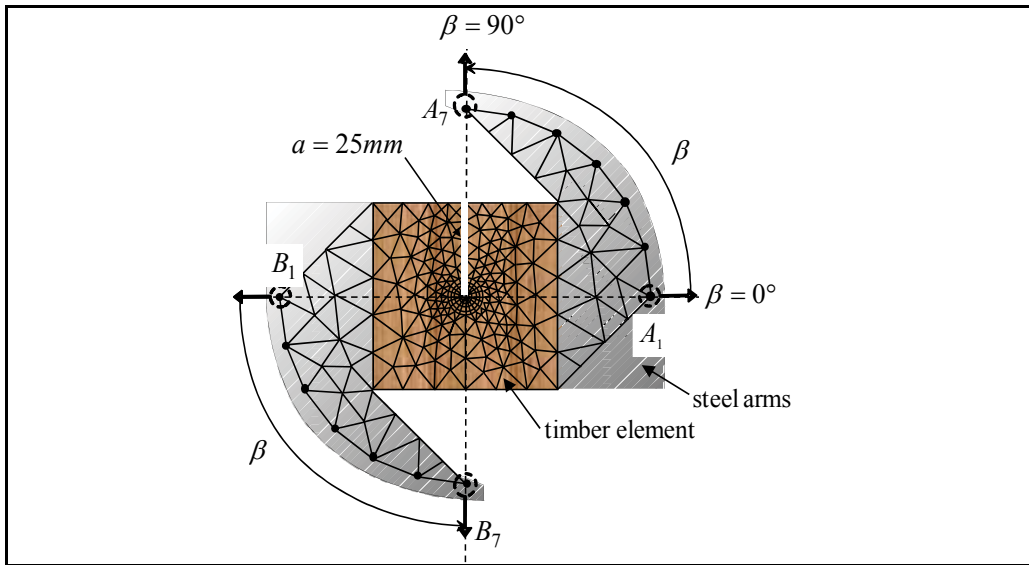


Fig. 5. CTS Specimen

In order to simplify the analytical development in the temporal field, the creep tensor is chosen as a proportional time function such as

$$[J](t) = \frac{I}{E_L(t)} \cdot [C_0] \quad (72)$$

C_0 is a constant and unit compliance tensor composed by a unity elastic modulus and a constant Poisson coefficient of 0,4. $E_L(t)$ designates the tangent modulus for the longitudinal direction. In this context, the creep properties are given in terms of creep function by interpolating $\frac{I}{E_L(t)}$ with six Kelvin Voigt cells, Figure 6.

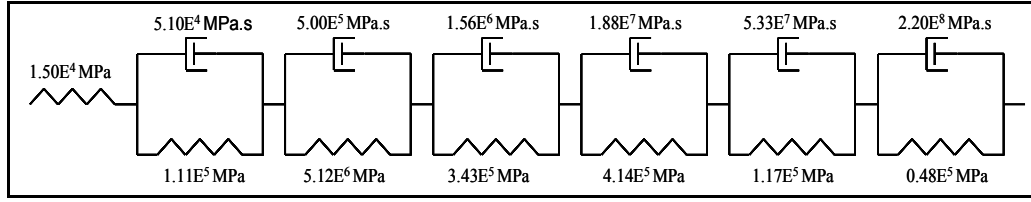


Fig. 6. Spectral decomposition of $1/E_L(t)$ (Moutou Pitti, 2008)

$1/E_L(t)$ takes the following analytic form :

$$\frac{1}{E_L(t)} = \frac{1}{E_L(0)} \cdot \left[1 + \frac{1}{74,3} \cdot \left(1 - \exp\left(-\frac{74,3}{3,37}t\right) \right) + \frac{1}{74,4} \cdot \left(1 - \exp\left(-\frac{74,4}{33,37}t\right) \right) \right. \\ \left. + \frac{1}{22,9} \cdot \left(1 - \exp\left(-\frac{22,9}{104,09}t\right) \right) + \frac{1}{27,6} \cdot \left(1 - \exp\left(-\frac{27,6}{1251}t\right) \right) \right. \\ \left. + \frac{1}{7,83} \cdot \left(1 - \exp\left(-\frac{7,83}{3554}t\right) \right) + \frac{1}{3,23} \cdot \left(1 - \exp\left(-\frac{3,23}{14660}t\right) \right) \right] \quad (73)$$

in which $E_L(0) = 15000 \text{ MPa}$ is the elastic longitudinal Young modulus for longitudinal direction. C_0 admits the definition for plane configurations

$$C_0 = \begin{bmatrix} 1 & -\nu & 0 \\ -\nu & \frac{E_L(t)}{E_R} & 0 \\ 0 & 0 & \frac{E_L(t)}{G_{LR}} \end{bmatrix} \quad (74)$$

$E_R = 600 \text{ MPa}$ and $G_{LR} = 700 \text{ MPa}$ are the transverse and shear modulus, respectively.

6.2 Displacement fields and meshes

The linear triangular elements with 3 nodes were used. If the thickness of the specimen is very low compared to two other dimensions, the modeling in plan stress is used. Like boundary conditions, the crack tip displacement is blocked. It is the same for the lower part of the wood specimen. On the line of cracking, displacements along the axis x are prevented. In order to have stable results, a radiant mesh is used around the crack tip. Figure 7 illustrates the virtual finite element deformation in opening and shear modes.

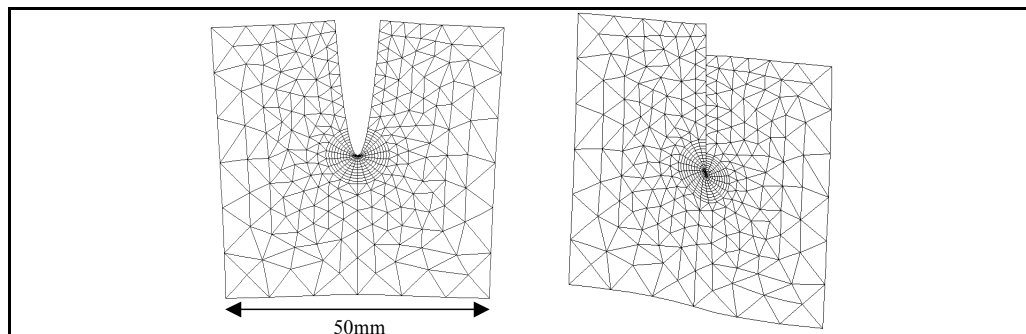


Fig. 7. Virtual displacements in opening and shear modes

Figure 8 (a) presents the detail of the radiant mesh around the crack tip on which the component θ_L of the field $\vec{\theta}$ is visualized, Figure 8 (b).

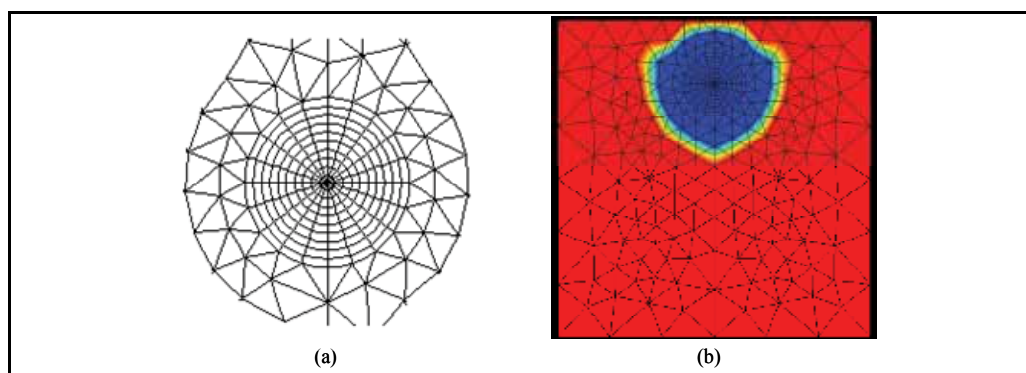


Fig. 8. Radiating circular mesh (a) and θ field (b) around the crack tip (Moutou Pitti, 2008)

6.3 Path independence domain

The independence path integral is checked by representing the various variations of the energy release rate versus each crown illustrating the size of the field $\vec{\theta}$. Five crowns noted and numbered C0 to C8, Figure 9, were tested.

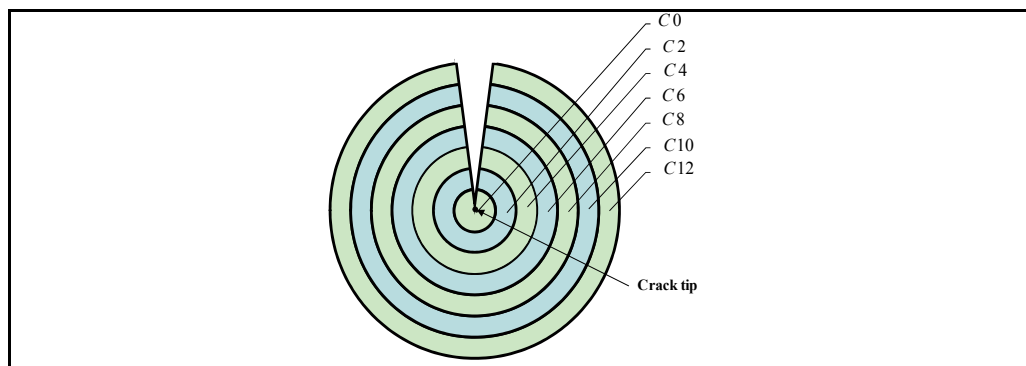


Fig. 9. Integration crowns

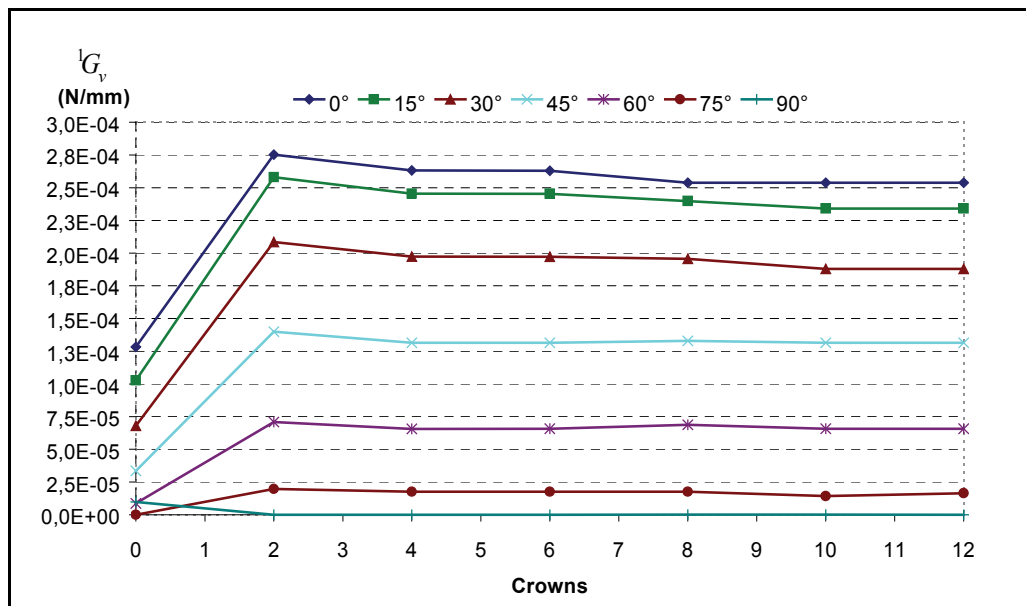


Fig. 10. Path independence domain 1G_v (opening mode) versus orientation angle β

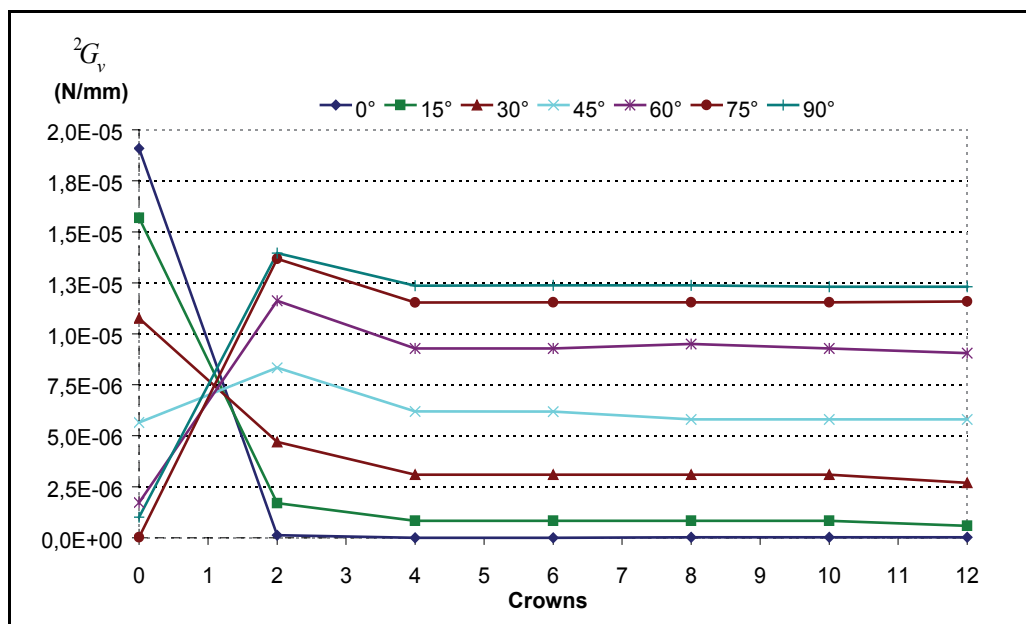


Fig. 11. Path independence domain 2G_v (shear mode) versus angle orientation β

Figures 9 and 10 show the energy release rate evolution in mode I and mode II versus the loading orientation and the various integration crowns, for a unit loading. Results were obtained after a creep time of 720 seconds. The constancy of the viscoelastic energy release rate (less of fluctuation) versus crowns in mode I, mode II and in mixed modes (share of

mode I and leaves mode II, respectively) is noted. These observations validate the path independence domain and the stability of results ensured by the model. However, the singularity of the mechanical fields at the crack tip causes field disturbance integration around the crack area (first two crowns), and this, for the two fracture modes.

6.3 Comparison of analytical and numerical solution

In order to validate the viscoelastic procedure, Results given by the numerical solution are compared with the analytical calculus resulting from the isothermal Helmholtz free energy density, equations (68). From relation (65), the viscoelastic creep tensor in mode I and mode II, respectively, takes the following form:

$$C_I(t) = C_I^{(0)} \cdot f(t) = 7,35 \cdot 10^{-3} \cdot f(t) \quad (75)$$

$$C_2(t) = C_2^{(0)} \cdot f(t) = 1,47 \cdot 10^{-3} \cdot f(t) \quad (76)$$

$C_I^{(0)}$ and $C_2^{(0)}$ represent the reduced elastic compliances, equations (53). The time creep function $f(t)$ admits an similarly form of (73) such as

$$f(t) = \left[\begin{aligned} &1 + \frac{1}{74,3} \cdot \left(1 - \exp\left(-\frac{74,3}{3,37}t\right) \right) + \frac{1}{74,4} \cdot \left(1 - \exp\left(-\frac{74,4}{33,37}t\right) \right) \\ &+ \frac{1}{22,9} \cdot \left(1 - \exp\left(-\frac{22,9}{104,09}t\right) \right) + \frac{1}{27,6} \cdot \left(1 - \exp\left(-\frac{27,6}{1251}t\right) \right) \\ &+ \frac{1}{7,83} \cdot \left(1 - \exp\left(-\frac{7,83}{3554}t\right) \right) + \frac{1}{3,23} \cdot \left(1 - \exp\left(-\frac{3,23}{14660}t\right) \right) \end{aligned} \right] \quad (77)$$

In plane configuration, the energy release rate in each mode is translated analytically by the expression (70)

$$G_I(t) = \frac{1}{8} \cdot [2 \cdot C_I(t) - C_I(2t)] \cdot \left({}^u K_I^{(0)} \right)^2 \quad (78)$$

$$G_2(t) = \frac{1}{8} \cdot [2 \cdot C_2(t) - C_2(2t)] \cdot \left({}^u K_2^{(0)} \right)^2 \quad (79)$$

${}^u K_I^{(0)}$ and ${}^u K_2^{(0)}$ are the instantaneous stress intensity factors in opening and shear mode, respectively, given by an initial finite element calculus. Figure 12 and 13 present the comparison of numerical results of viscoelastic energy release rate given by the $M\theta$ procedure and analytical results resulting of equations (78) and (79) in pure opening mode ($\beta = 0^\circ$) and pure shear mode ($\beta = 90^\circ$). Figure 14 and 15 show the same comparison of

the energy release rate in mixed mode $\beta = 45^\circ$. (1G_v part of opening mode, 2G_v part of shear mode). The progression of the energy release rate is given versus time. The results are

calculated by using the crown integration *C6*. We observe a perfect agreement between numerical and analytical results. We note also, in the case of mode I, mode II and mixed, the average error is definitely lower than 1%.

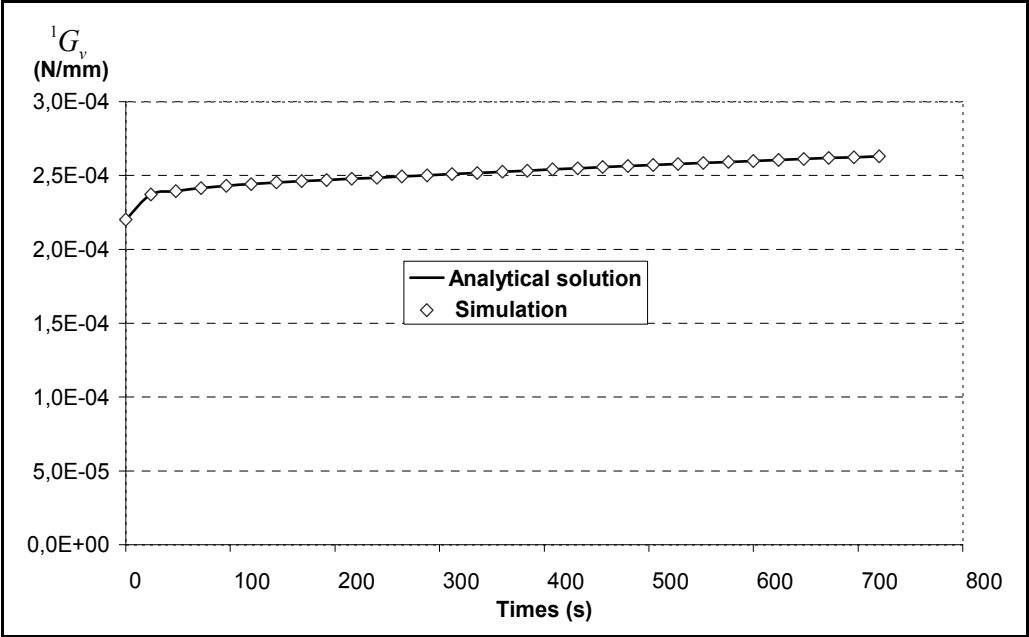


Fig. 12. Analytical and numerical solution in opening mode

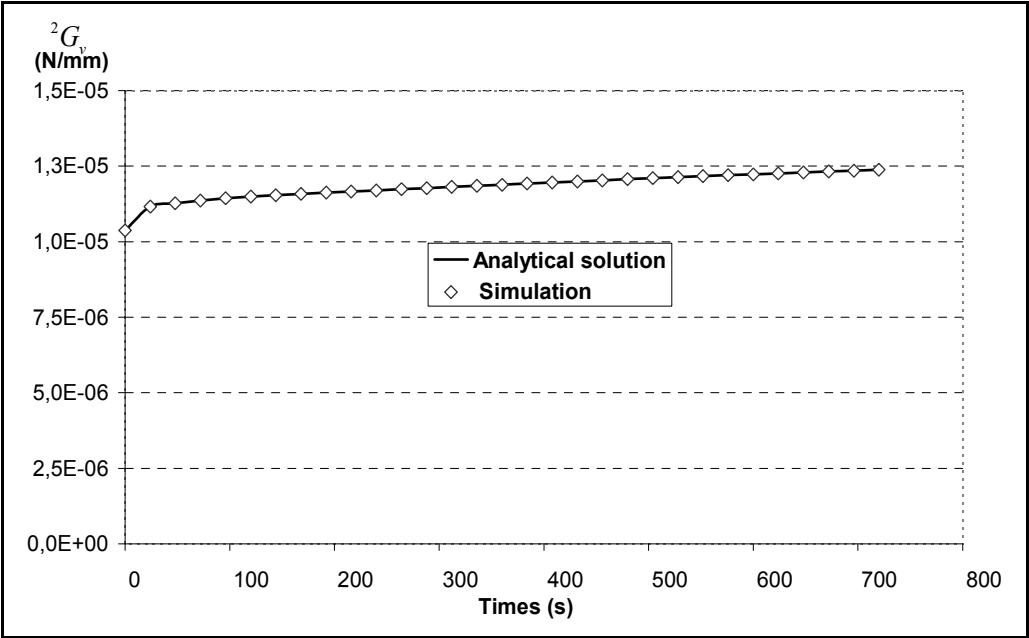


Fig. 13. Analytical and numerical solution in shear mode

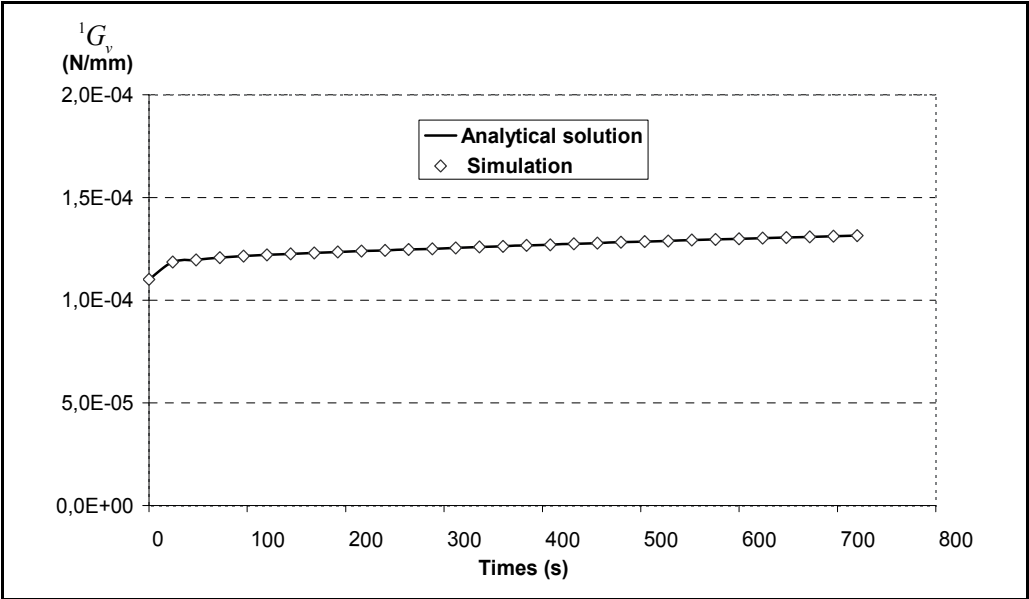


Fig. 14. Analytical and numerical solution in mixed mode (part of opening mode 45°)

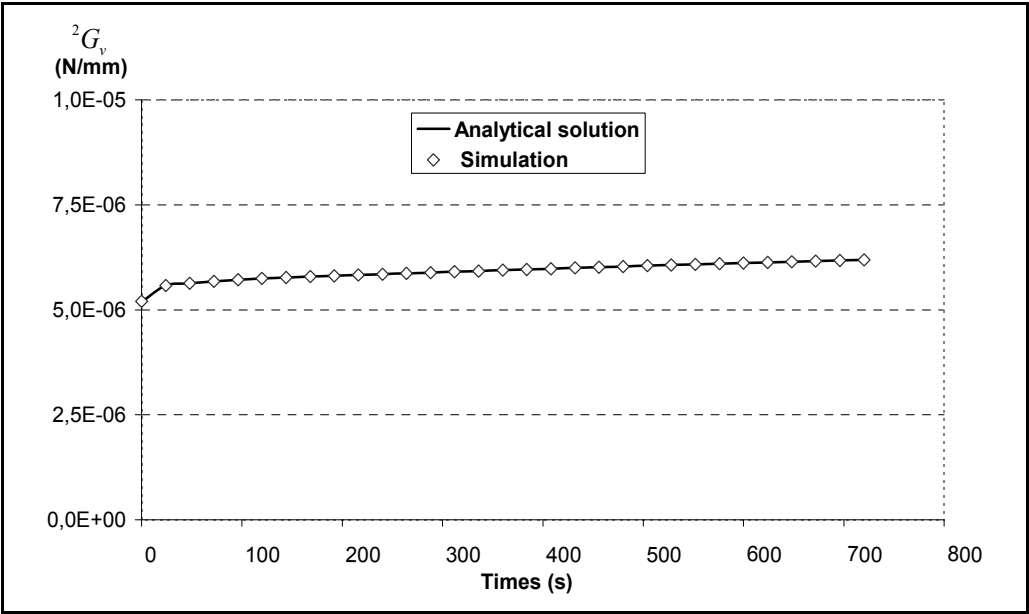


Fig. 15. Analytical and numerical solution in mixed mode (part of shear mode 45°)

7. Conclusion

This chapter has treated the complex problem of the fracture mechanic process in an orthotropic and viscoelastic media. The global algorithm, implemented in the finite element method, is a coupling of viscoelastic subroutine and fracture mechanic tools. According to a

stationary crack, the $M\theta$ -integral is employed in order to compute open and shear parts of energy release rate versus time by taking into account, via the Helmholtz's free energy potential, dissipated and released energy induced by viscoelastic properties.

However, if the crack growth initiation is a high important problem in terms of timber structure design, the problematic of the crack process remains important in the structure live approach taking into account long term behaviours. In this condition, this work leads to be completed in order to integrate the time crack growth process in elements loading by constant or variable loadings.

Finally, timber structures placed in outdoor conditions are subject to climatic variations. In this case, this work must be generalized by introducing moisture variation effect in the crack tip vicinity in the speed increase of crack growth initiation and crack growth propagation.

8. References

- Bui, H.D. & Proix, J.M. (1985). Découplage des modes mixtes de rupture en thermo-élasticité par des intégrales indépendantes du contour. *Actes du Troisième Colloque Tendances Actuelles en Calcul de Structure*, Bastia pp. 631-643.
- Bui, H.D (2007) Conservation laws, duality and symmetry loss in solid mechanics. *International Journal of Fracture*, Vol. 147, 163-172.
- Chen, F.M.K. and Shield, R.T. (1977). Conservation laws in elasticity of the J-integral type. *Journal of Applied Mechanics and Physics*, Vol. 28, No 1, 1-22.
- Chazal, C., & Dubois, F. (2001). A new incremental formulation in the time domain of crack initiation in an orthotropic linearly viscoelastic solid. *Mechanics of Time Dependent Materials*, Vol. 5, 2001, 3-21.
- Destuynder, Ph., Djaoua, M. and Lescure, S. (1983). Quelques remarques sur la mécanique de la rupture élastique. *Journal de Mécanique Théorique et Appliquée*, Vol. 2, No 1, 113-135.
- Dubois, F. Petit, C. (2005). Modeling of the crack growth initiation in viscoelastic media by the $G\theta_v$ -integral. *Engineering Fracture Mechanics*, Vol. 72, 2821-2836.
- Dubois, F., Chazal, C., Petit, C. (1999) Modeling of crack growth initiation in a linear viscoelastic material. *Journal of Theoretical and Applied Mechanics*, Vol. 37, No 2, 207-222.
- Dubois, F., Chazal, C., Petit, C. (2002) Viscoelastic crack growth process in wood timbers: An approach by the finite element method for mode I fracture. *International Journal of Fracture*, Vol. 113, No 4, 367-388.
- Ghazlan, G., Caperaa, S. and Petit, C. (1995). An Incremental formulation for the linear analysis of thin viscoelastic structures using generalized variables. *International Journal of Numeric Methods Engineering*, Vol., No 38: 3315-33.
- Irwin, G.R. (1957) Analysis of stresses and strains near the end of a crack traversing a plate. *Journal Applied Mechanics*, Vol. 24, 361-385.
- Moutou Pitti, R., Dubois, F., Petit, C., Sauvat N (2007) Mixed mode fracture separation in viscoelastic orthotropic media: numerical and analytical approach by the $M\theta_v$ -integral. *International Journal of Fracture*, Vol. 145, No 3, 181-19, ISSN 0376-9429 (Print) 1573-2673 (Online).

- Moutou Pitti, R. (2008). Mixed mode fracture separation in viscoelastic orthotropic materials : modeling and experimentation. Ph.D.thesis, Limoges University . <http://www.unilim.fr/theses/2008/sciences/2008limo4025/notice.htm>. Accessed 23 Jan 2009.
- Noether, E. (1971). Invariant variations problem. *Transport Theory and Statistical Physics*, Vol. 1, No 3, 183-207.
- Richard, HA. (1981) A new compact shear specimen. *International Journal of Fracture*, Vol. 17, No 5, R105-R107.
- Rice, J.R. (1968). A path independent integral and the approximate analysis of strain concentrations by notches and cracks. *Journal of Applied Mechanics*, Vol. 35, 379-386.
- Masuro, J.R. & Creus, G.J. (1993) Finite elements analysis of viscoelastic fracture. *International Journal of Fatigue*, Vol. 60, 267-282.
- Staverman, A. J., & Schwarzl, P. (1952) Thermodynamics of viscoelastic behavior. *Proceeding Academic Science*, Vol. 55, 474-492.
- Valentin, G. & Caumes, P. (1989) Crack propagation in mixed mode in wood: a new specimen. *Wood Science and Technology*, Vol. 23, No 1, 43-53.
- Zienkiewicz, O.C., Watson, M., King, I.P. (1968) A numerical method of viscoelastic stress analysis. *Int. J. Mech. Sci.*, Vol. 10, 807-827.

Finite Element modelling of Elastic-Plastic Contact of Rough Surfaces

Jamil Abdo, M. Danish Haneef and Abdullah M. Al-Shabibi
Sultan Qaboos University
Oman

1. Introduction

An improved mathematical elastic-plastic model for the contact of rough surfaces that is based on an accurate finite elements solution of a deformable single asperity and a rigid flat surface is developed to provide dimensionless expressions for the contact area and contact load. This model differs from the existing models, in that it accounts for the level of interference beyond expected failure. The finite element solution is used to define the limits at which failure occurs. The derivation of the contact model is facilitated through the definition of the ultimate-stress asperities that are assumed to be embedded at a critical depth within the actual surface asperities. This model considers a realistic picture of elastic-plastic deformation where elastic, plastic and failure behaviors can occur simultaneously for an asperity. Subsequent comparison of the results for estimating contact area and load using the present model and the earlier methods shows identical results for pure elastic contacts with plasticity index values at about 0.5 but substantial difference for the net elastic-plastic contacts having plasticity index values above 0.8. When plasticity index reaches 6 and beyond the three models predicts similar total contact area and load values and that the contact is purely plastic.

2. Background

The role of surface roughness in contact mechanics is relevant to processes ranging from adhesion to friction, wear and lubrication¹. It also promises to have a deep impact on applied science, including coatings technology and design of micro electro-mechanical systems. Despite the considerable results achieved by indentation experiments⁴, particularly in the measurement of bulk hardness on small scales, the contact behavior of realistic surfaces, showing random multi-scale roughness, remains largely unknown. In many engineering applications, frictional contact occurs between machine parts and the characterization of contact behavior becomes an important subject in solving tribological problems such as friction induced vibration (Brockley, 1970; Ibrahim & Rivin, Tzou et al.,1998; Abdo, 2006), wear (Mulhearn & Samuels, 1962; Samuels, 1978; Archard, 1953; Halling et al.,1975; Abdo & Yahmadi, 2004), issues related to mechanical sealing, performance and life of machine elements, and thermal to name few. The pioneering work

of Greenwood and Williamson (Greenwood & Williamson, 1966) (GW model) has been utilized by many researchers (Greenwood & Tripp, 1970; Chang et al., 1987; Zhao et al., 2000; Abdo & Farhang, 2005) as a basic for further extension to obtain contact models for general or specific contact problems for mainly elastic contact. On the other hand, the work of Pullan and Williamson (Pullen & Williamson, 1972) utilized as a basic model for pure plastic contact. In an attempt to bridge the gap between the pure elastic and pure plastic contact, Chang et al. (CEB model) developed a wide intermediate range of interest where elastic-plastic contact triumph. The existing probabilistic models of contact may be viewed with respect to the premise of elastic and plastic contact. The elastic models primarily rely on the Hertz theory of contact between two elastic bodies (Greenwood & Williamson, 1966; Greenwood & Tripp, 1967; Greenwood & Tripp, 1970; Hisakado, 1974; Bush et al., 1975; McCool, 1986). These models differ in their assumptions related to surface and asperity geometry and material properties. These extensions have included, for instance, the inclusion of the surface curvature effects (Greenwood & Tripp, 1967), allowance for non-uniform curvature of asperity summits (Hisakado, 1974) and presumption of average elliptic paraboloidal representation of asperity (Bush et al., 1975). Whereas the elastic and plastic models are seen to be advantageous for extreme cases of loading, in a large number of engineering applications, contact loads may fall within ranges that do not warrant adequate representation by either elastic or plastic model. This fact has led researchers to consider what is referred to as elastic-plastic models (Chang et al., 1987; Ishigaki et al., 1987). The Zhao model [Zhao et al., 2000] suggests that the contact interference at the inception of fully plastic deformation would be at least 54 times that of the initiating yielding. However, the exact inception of fully plastic deformation was still not given. All of the initial models do not provide a solution due to the basic problem of lacking accuracy in the elastic-plastic contact regime. Such accurate solution calls for the use of a Finite Element Method (FEM)

3. General Approach to solve contact problems by FEM

Due to the complex contact mechanics the scientific community has become more and more interested in using the finite element method to solve contact problems. Analytical solutions for the elastic-plastic asperities interaction are limited and constrained by a number of assumptions and approximations. Finite element analysis is a powerful tool when analytical solutions are difficult to obtain. The idea of finite element is based on the discretization of the continuous domain or geometry into sub-domains or elements. An approximate solution is then sought for each element and integrated with solution of the other elements. The accuracy of the approximated solution depends solely on the number of elements used. To determine the appropriate number of elements, a series of run needs to be performed where the convergence of the finite element solution is monitored. Finite element analysis has been extensively used to study the elastic-plastic interaction between asperities and rigid flat surface. (Kogut & Etsion, 2002) presented a finite element solution of a single asperity contact. Their model have provided more accurate results for the contact parameters such as separation, contact area and contact pressure as functions of the plasticity index and contact load. (Jackson & Green, 2004) presented a finite element study of an elastic-plastic hemisphere in contact with a rigid flat surface. In order to improve the efficiency of computation, Jackson and Green (2005) developed an axisymmetric 2-D model (Fig.1). They used commercially available ANSYS software to for the analysis and produced more

refined results than Kogut & Etsion (2002). The contact region was meshed by 100 contact elements. The meshed contact area also controlled to ensure that at least 30 contact elements are in contact for each applied interference. These are in essence very stiff springs attached between surface nodes and they activate only when penetration onset into the rigid flat is detected.

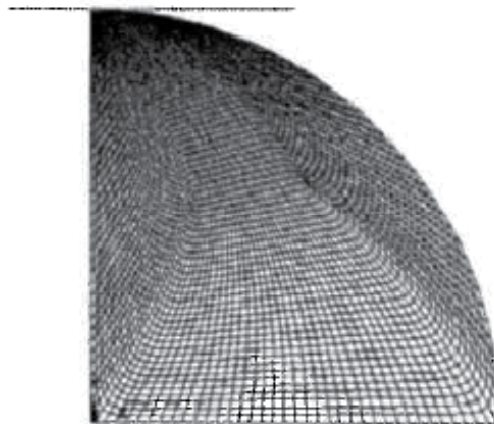


Fig. 1. Finite element mesh of a sphere generated by ANSYS

They found that the fully plastic average contact pressure is not constant and varies with the deformed contact geometry. They also extended the range of ω/ω_c beyond the 110 considered by Kogut and Etsion. (Eid & Adams, 2007) used finite element analysis to study the interaction of two-hemispherical asperities with a flat rigid surface. In their study they have considered asperities of different heights and separations. They presented dimensionless results for the contact force and contact area versus interference. Generally though, the differences are small enough that the FEM solutions practically confirms to the Hertzian solution at interferences below critical (and even slightly above). In this presented work the finite element method was used to solve the elastic-plastic contact of a single asperity. The contact problem and the elasto-plastic material property make the analysis highly nonlinear and difficult to converge. An iterative scheme is used to solve for the solution, and many load steps are used to enhance solution convergence. Initially, a small interference is set of the total interference and then it is incremented after the load step converges. ANSYS internally controls the load stepping to obtain a converged solution by using the bisection method. This continues until a converged solution is found for the desired interference. This work presents an elastic-plastic contact model of flat on rough surface. It is based on the work developed by (Abdo & Farhang, 2005). A fictitious ultimate stress asperity is introduced to facilitate the derivation of the contact model. The finite element method is utilized to define the limit of the expected failure.

4. Overview of the Contact Model

The contact between smooth and rough surface is considered. The contact model is based on the presumption that a surface can, in effect, be represented by a distribution of asperities (Fig. 1). As two surfaces are brought into contact, the macroscopic contact characteristic in

question is a cumulative effect of localized interactions of the smooth surface and the asperities on the rough surface. This approach has required the statistical formulation of a surface and statistical summation of microscopic contact effects to obtain probabilistic macroscopic expectation of the contact characteristic (contact area, load, and stiffness). Many of the contact models predict extreme situations; contact is purely elastic or purely plastic.

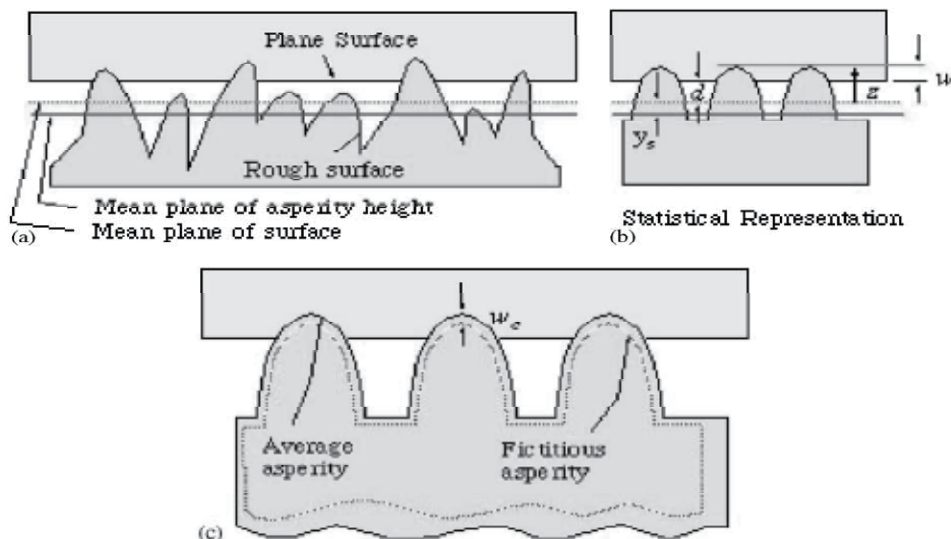


Fig. 2. Contact between a flat and a rough surface

However, the contact is better described, for moderate load ranges, when it includes both elastic and plastic contacts. The proposed model approximates the behavior of the intermediate load ranges, which is referred to as elastic-plastic contact model. In the derivation of the equations the contact between one asperity on a rough surface and a plane is considered. The behavior of the asperity is initially elastic. As the load is increased the elastic behavior continues to describe the deformation until a critical interference is reached. At this critical load and beyond, the asperity deforms as a purely plastic body. Hence, for every asperity there are two types of interactions. The first is the elastic contact between the plane and the surface asperity. If the interference (w) exceeds the critical interference (w_c), then the interaction also includes plastic contact. The shaded volume representing the interference of the plastic asperities and the plane contribute to the plastic portion of contact whereas the remaining volume of interference contributes to the elastic contact (Fig. 2).

5. The Mathematical Model

In a recent pioneering work by Abdo and Farhang (AF model) presented the plastic asperity concept for modeling the elastic-plastic contact of rough surfaces. The salient feature of their approach is outlined here in brief in order to set a scene for the present analysis. Considering the contact between one single asperity on rough surface and a rigid plane, the behavior of the asperity is initially elastic. As the load is increased the elastic behavior continues to describe the deformation until a critical interference is reached. At this critical load and beyond, the asperity deforms as a purely plastic body. Hence for every asperity

there are two types of interactions. The first is the elastic contact between the plane and the asperity. If the interference (ω) exceeds the critical interference (ω_c), then the interaction also includes plastic contact. Abdo and Farhang considered elastic-plastic contact through the introduction of a fictitious asperity that can only deform plastically. The characteristics of contact (area, or load) Q , in AF model is defined as

$$Q = Q_{e1} - Q_{e2} + Q_{p2} \quad (1)$$

Where (Q_{e1}) is the characteristic of contact corresponding to the contribution due to the elastic interference between the plane and surface asperity, (Q_{e2}) is the characteristic of contact corresponds to the contribution due to the elastic interference between the plane and plastic asperity and must be subtracted from (Q_{e1}) to obtain the net elastic contribution. Next the contribution from plastic interaction due to the plastic interference of the plane and plastic asperity (Q_{p2}) must be added to the result to obtain the net elastic-plastic characteristics of contact.

In this work, the elastic-plastic contact model (AF model) of rough surfaces is developed further to account for the level of interference beyond which failure is expected. That is if the interference in AF model is increased further until it equals or exceeds the critical interference corresponds to the ultimate stress of the material (ω_u) at which failure to occur, then the interaction results in failure of the surface asperity and the interference volume will be separated from the asperity.

To account for the level of interference beyond expected failure, Fig. 3 that describes schematically the geometry of the three types of interactions of contacting of rough surfaces is introduced. Therefore, the characteristics of contact, Q , may be obtained by appropriately accounting for the aforementioned interactions. That is

$$Q = Q_{e1} - Q_{e2} + Q_{p2} - Q_{p3} \quad (2)$$

As it is shown in Fig. 3, the plastic portion of the interference may also include the interference with the ultimate-stress asperity (Q_{p3}), which is to be subtracted from (Q_{p2}) to give the net plastic interference. The critical interferences (ω_c) and (ω_u) in Fig. 3 are used to define the limits of the plastic and ultimate-stress asperities, respectively. The surfaces of the plastic and the ultimate-stress asperities are obtained by displacement of every point on the physical asperity by (ω_c) and (ω_u) along the direction normal to the original surface asperity. Using mathematical mapping of a point on the physical asperity to a corresponding point on the ultimate-stress asperity and similar to the derivation of the summit radius of curvature of the plastic asperity (R_p) in (Abdo & Farhang, 2005) an expression for the summit radius of curvature for the ultimate-stress asperity (R_u) can be presented as

$$R_u = R - \omega_u \quad (3)$$

where R is the radius of curvature for the original surface asperity and (R_p) is derived by Abdo & Farhang as (See appendix A)

$$R_p = R - \omega_p \quad (4)$$

The ultimate-stress asperity is assumed to be embedded at a critical depth of (ω_u) within the actual surface asperities. In **Fig. 3**, the shaded volume representing the interference of the plastic asperities and the plane contribute to the plastic portion of contact. The interference beyond this plastic portion of contact will be separated from the asperity. The model consider the interaction of a rough surface covered with a number of asperities that have a spherical shape at their summits with uniform asperity radius of curvature, R , asperity height, z , and separation of the surfaces, d , and a nominally flat surface. z and d are measured from the reference plane define by the mean of the original surface heights. h represents the separation based on surface heights and the standard deviations σ_s and σ correspond to the asperity and surface height, respectively. The difference between h and d is denoted as y_s .

The expected number of asperities in contact is

$$N_c = N \int_{d^*}^{\infty} \phi^*(z^*) dz^* \quad (5)$$

where the total number of asperities, N , the density of asperities, η , and the nominal area are related according to

$$N = \eta A_n \quad (6)$$

where all length dimensions are normalized by σ and the dimensionless values are denoted by $*$ and the dimensionless asperity height probability density function is

$$\phi^*(z^*) = (2\pi)^{-0.5} (\sigma/\sigma_s) \exp\left[-0.5(\sigma/\sigma_s)^2 (z^*)^2\right] \quad (7)$$

For this type of contact, the dimensionless interference is define as

$$\omega^* = z^* - d^* \quad (8)$$

The contact of rough surfaces can be modeled by a flat and a smooth surface in contact with a rough surface. The assumptions are along with the ones given by GW model and CEB model. Since the plastic asperity only deform plastically and by imposing the conversation of volume, Abdo and Farhang derived the equations describing the elastic contact area A_e , the elastic contact load P_e , the plastic contact area A_p and the plastic contact pressure P_p as:

$$A_e = \pi R \omega \quad (9)$$

$$P_e = \frac{4}{3} ER^{1/2} \omega^{3/2} \quad (10)$$

$$A_p = 2\pi R_p \omega \quad (11)$$

$$P_p = 2\pi KHR_p \omega \quad (12)$$

The limit of elastic interference ω_c at the inception of elastic-plastic is defined in GW model and AF model:

$$\omega_c = \left(\frac{\pi KH}{2E} \right)^2 R \quad (13)$$

where H is the hardness of the softer material of the contacting surfaces and K , the hardness coefficient, is related to the Poisson ratio of the softer material by (Lin & Lin, 2005)

$$K = 0.4645 + 0.3141\nu + 0.1943\nu^2 \quad (14)$$

E is the Hertz elastic modulus and given by

$$\frac{1}{E} = \frac{1-\nu_1^2}{E_1} + \frac{1-\nu_2^2}{E_2} \quad (15)$$

E_1 , E_2 and ν_1 , ν_2 are Young's moduli and Poisson's ratios of the contacting surfaces, respectively.

From the end of the fully elastic deformation to the inception of fully plastic deformation, the contact of the two surfaces goes through an elastic-plastic deformation. At the inception of fully plastic, the maximum contact pressure P_{\max} of an asperity changes from KH to H . Therefore, the plastic contact pressure at the inception of fully plastic deformation can be written as:

$$P_p = 2\pi HR_p \omega \quad (16)$$

6. The Finite-Element Model

While an accurate measurement of ω_u is not easy to obtain, the minimum value at which failure is expected to occur may be estimated based on a finite element analysis. Using a finite element method (Kogut & Estion, 2002) found that the entire elastic-plastic contact regime of a single asperity extends over the range $1 \leq \omega/\omega_c < 110$, with a transition at $\omega/\omega_c = 6$ that divides it into sub-regions. In the present work, an axisymmetric two-dimensional finite element model was constructed to model and solve the asperity contact problem, using commercial software ANSYS 8.1. The asperity was modeled by a quarter of circle. The Von Mises yielding criterion was used to detect the transition from elastic to

elastic-plastic deformation. The rigid flat was modeled by a line and the material of the sphere was assumed elastic-perfectly plastic with identical behavior in tension and compression. The input parameters were normalized with respect to ω_c . The normalization of the output parameters was done with respect to the yield strength Y of the sphere material. As in Kogut and Etsion (2002) the validity of this normalization was tested by solving the problem for different material properties $100 < E/Y < 1000$, $\nu = 0.3$ and sphere radii ($0.1\text{ mm} < R < 10\text{ mm}$). The dimensionless results of P/P_c , A/A_c and p/Y versus the dimensionless interference, ω/ω_c were always the same regardless of the selection of material properties and sphere radius.

The mesh was composed of 1792 eight-node quadrilateral axisymmetric elements comprising a total of 5537 nodes. For a good representation of the contact geometry, higher order elements are selected to better fit the curvature of the sphere. A zero displacement in all directions is specified for the nodes at the bottom and symmetry boundary conditions are applied for the nodes along the centerline. The finite element model was verified with results obtained from measurements and simulations of Kogut and Etsion.

To accurately compare to their work, the region was divided into two mesh densities. Zone I within a $0.1R$ distance from the sphere tip has a higher mesh density consisting of 43% of the nodes and 44% of elements, to better capture the high stress concentration and achieve an accurate discrimination for detection of the contact area radius. In this region the mesh size was $0.001a_c$ where $a_c = (R\omega_c)^{1/2}$. Zone II outside the $0.1R$ distance, had a coarser mesh compared to that of Zone I. The model also contained a single 2-D target element laying on the flat surface and 32 two dimensional surface-to-surface contact elements on the sphere surface in Zone I. A series of runs were conducted and for each run an incremental displacement was applied to the rigid body. The contact forces are obtained from the nodal forces of the active contact elements. The length of the contact zone is determined from the number of the active contact elements. For each run the Von Mises stress is checked to determine the extension of the plastic zone. Fig. 4 shows the evolution of plastic region when $\omega/\omega_c < 11$. Up to $\omega/\omega_c = 6$, the plastic region is completely surrounded by elastic material. At $\omega/\omega_c = 6$, the plastic region first reaches the sphere surface at a radius of about $3a_c$. At this point an elastic core remains locked between the plastic region and the sphere surface. Therefore, it is anticipated that this critical interference (ω_u) corresponds to the interaction between the flat surface and the ultimate-stress asperity. The interference beyond this critical interference will be separated from the asperity.

Utilizing Eqn. (2), the derivation presented in the previous section and the dimensionless length definition one can obtain the dimensionless contact area between rough surfaces:

$$A(d) = A_{e1} - A_{e2} + A_{p2} - A_{p3} \quad (17)$$

where,

$$A_{e1}(d) = \pi \beta \int_{\frac{d}{\omega_c}}^{\infty} \omega \phi(z) dz \quad (18)$$

$$A_{e2}^*(d) = \pi \beta_p^* \int_{d+\omega_c^*}^{\infty} \omega_p^* \phi^*(z) dz \quad (19)$$

$$A_{p2}^*(d) = 2\pi \beta_p^* \int_{d+\omega_c^*}^{\infty} \omega_p^* \phi^*(z) dz \quad (20)$$

$$A_{p3}^*(d) = 2\pi \beta_u^* \int_{d+\omega_u^*}^{\infty} \omega_{pu}^* \phi^*(z) dz \quad (21)$$

where β, β_p, β_u and y_s^* are defined as follows:

$$\beta = \eta R \sigma, \quad \beta_p = \eta(R - \omega_c), \quad \beta_u = \eta(R - \omega_u), \quad y_s^* = h - d = \frac{1.5}{\sqrt{108\pi\beta}},$$

$$\omega^* = z^* - d^*, \quad \omega_c^* = \omega_c / \sigma, \quad \omega_p^* = \omega^* - \omega_c^*, \quad \omega_{pu}^* = \omega^* - \omega_u^*,$$

In similar manner, the contact load may be written as;

$$P(d) = P_{e1}^* - P_{e2}^* + P_{p2}^* - P_{p3}^* \quad (22)$$

where the dimensionless contact load, P^* , is obtained by dividing the nominal contact pressure, P/A_n , by the hardness H , i.e. $P^* = P/(A_n H)$.

$$P_{e1}^*(d) = \frac{2}{3} \frac{\pi K \beta^*}{(\omega_c^*)^{0.5}} \int_{d+\omega_c^*}^{\infty} \omega^{3/2} \phi^*(d) dz \quad (23)$$

$$P_{e2}^*(d) = \frac{2}{3} \frac{\pi K \beta_p^*}{(\omega_c^*)^{0.5}} \int_{d+\omega_c^*}^{\infty} \omega_p^{3/2} \phi^*(d) dz \quad (24)$$

$$P_{p2}^*(d) = 2\pi K \beta_p^* \int_{d+\omega_c^*}^{\infty} \omega_p^* \phi^*(d) dz \quad (25)$$

$$P_{p3}^*(d) = 2\pi K \beta_u^* \int_{d+\omega_u^*}^{\infty} \omega_{pu}^* \phi^*\left(\frac{d}{z}\right) dz \quad (26)$$

7. Results and Discussions

For the purpose of comparing the results evaluated by KE (Kogut & Etsion, 2002), LL (Lin & Lin, 2005) models and the present model the equations established in these models and in the previous section are evaluated numerically. The area-separation and load-separation are investigated for typical values of plasticity index Ψ . The values of the plasticity index determine the nature of contact and are used to analyze the effect surface roughness and material properties on the contact of rough surfaces. For $\Psi < 0.6$, the contact is predominantly elastic and for $\Psi > 1.0$, the contact is predominantly plastic. Thus plasticity, Ψ , values are considered in the range of 0.5–2.5 in order to consider the whole range of deformation from predominantly elastic to predominantly plastic including elastic-plastic. The values of β are selected according to AF model (Abdo & Farhang, 2004) and the maximum contact pressure, K , are calculated from equations (14).

Fig. 5 (a) (b) (c) illustrates the dimensionless mean separation, h^* , vs. the dimensionless contact load, P^* , for various values of plasticity index Ψ . While the three models predict similar contact loads for low plasticity index, $\Psi = 0.5$, the present model predicts a lower values for materials of higher plasticity index.

Fig. 6 (a) (b) (c) depicts the dimensionless contact area vs. dimensionless separation for different values of plasticity index as predicted by KE, LL models and the present model. In the case of $\Psi = 0.5$, the total contact area curves of the three models almost grouped together over a wide range of dimensionless separations. If the plasticity index Ψ rises, the variations of the dimensionless total contact area for the three models are shown. The total contact load and area curves for plasticity index Ψ at 0.8 and 2.5 predicted by the present model are lower than the KE and LL curves. The variation came from the subtraction of the characteristic of contact corresponds to the contribution due to the elastic interference between the plane and the plastic asperity and the subtraction of the plastic portion of the interference that include the interference with the ultimate-stress asperity beyond expected failure. That is subtracting A_{e2}^* and A_{p3}^* integral forms to obtain the net contact area and

subtracting P_{e2}^* and P_{p3}^* integral forms to obtain the net contact load. It is noteworthy to mention that in these integrals, the lower limit of the integration are shifted by apparent critical interferences ω_c^* and ω_u^* . Furthermore, the asperity curvature corresponding to plastic asperities as given in eqn. (4) and the asperity curvature corresponding to ultimate-stress asperity in eqn. (3) need to be used in the corresponding integrals in eqns. (21 and 26) and eqns. (19, 20, 24, and 25), respectively. Furthermore, when the plasticity index reaches 6, transition value of Ψ corresponds to $\omega/\omega_c \cong 110$, and beyond the three models predicts similar total contact area and load values and that the contact is approximately plastic.

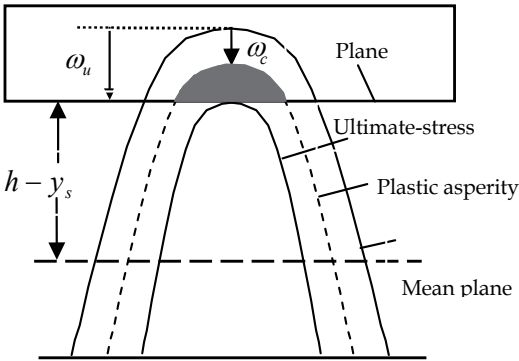


Fig. 3. Stages of interaction of flat surface on rough asperity

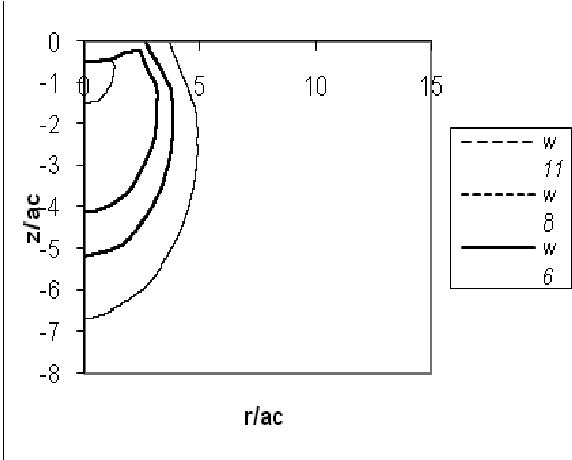
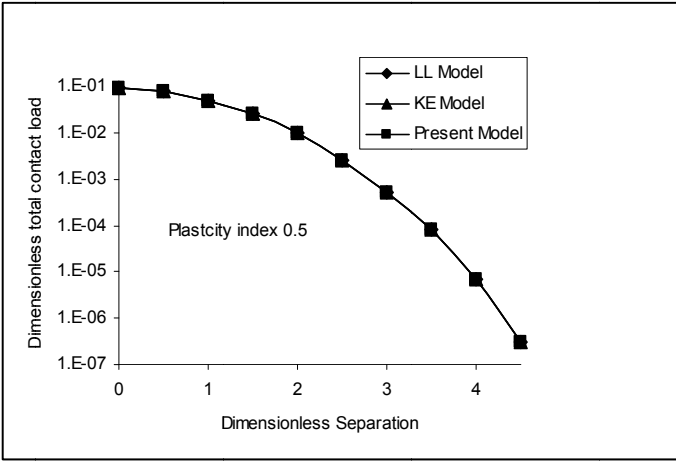
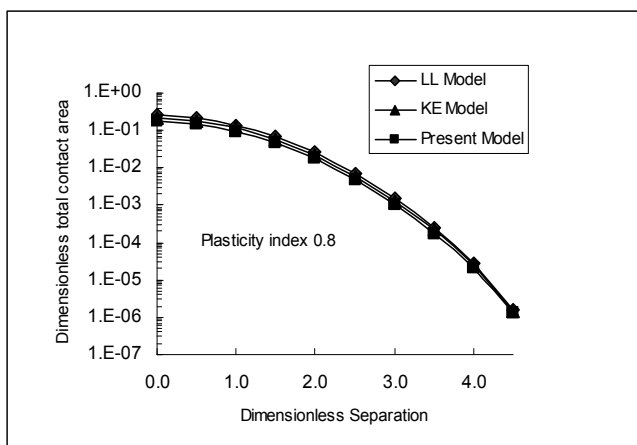


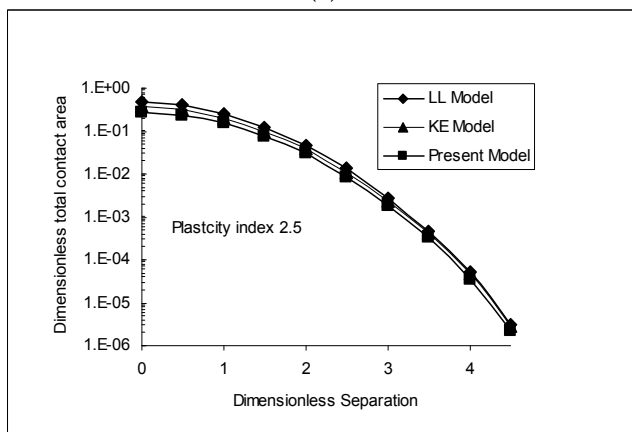
Fig. 4. Evolution of the plastic region in the asperity tip for a range (2-11) of ω/ω_c



(a)

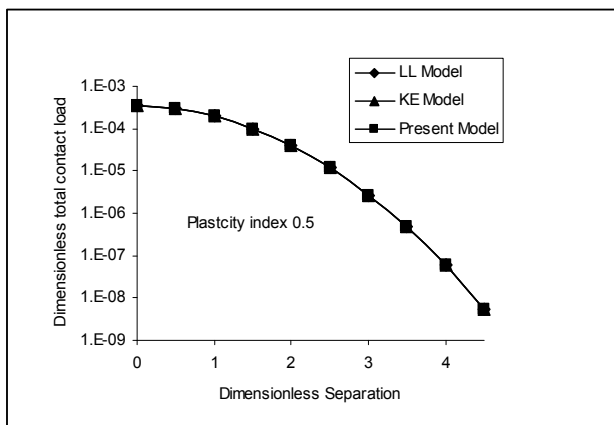


(b)

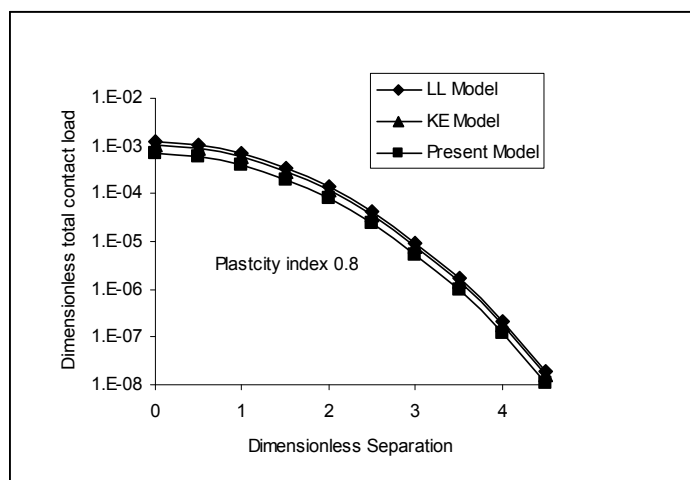


(c)

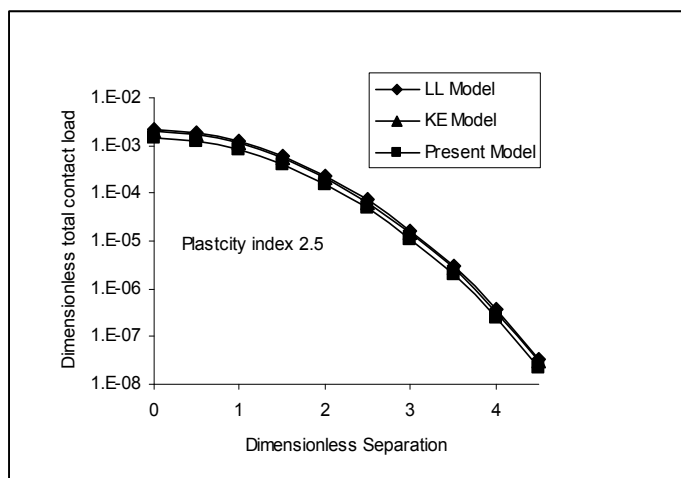
Fig. 5. Variation of the dimensionless total contact load with the dimensionless separation; (a) $\Psi = 0.5$, (b) $\Psi = 0.8$, (c) $\Psi = 2.5$



(a)



(b)



(c)

Fig. 6. Variation of the dimensionless total contact area with the dimensionless separation; (a) $\Psi = 0.5$, (b) $\Psi = 0.8$, (c) $\Psi = 2.5$

8. Conclusions

The contact area and contact load of an elastic-plastic micro-contact was calculated. The ultimate stress asperity is embedded at a critical depth within the actual surface asperities. The finite element solution is used to define the limit at which failure is to occur. The present model is more accurate than the previous models since it accounts for the net elastic-plastic by subtracting the plastic portion that reached the ultimate-stress asperity limit. Comparisons of the present model with the existing models for elastic-plastic contact have been performed. Results show that the previous elastic-plastic model underestimates the loading force particularly at elastic-plastic. Either total contact area or total contact load of

the rough surfaces depend is dependent upon the model developed for the elastic-plastic regime and the plasticity index. At a small plasticity index, these two contact parameters predicted by the three models are quite close. However, substantial differences in the contact parameters among these three models are apparent as the plasticity becomes sufficiently large. When the plasticity index reaches 6 and beyond the three models predicts similar total contact area and load values and that the contact is approximately plastic.

Nomenclature

Q = The characteristics of contact (area, or load),

Q_{e1} = characteristic of contact corresponds to the contribution due to the elastic interference between the plane and surface asperity

Q_{e2} = characteristic of contact corresponds to the contribution due to the elastic interference between the plane and plastic asperity

Q_{p2} = contribution from plastic interaction due to the plastic interference of the plane and plastic asperity

Q_{p3} = ultimate-stress asperity

ω_u = ultimate stress of the material

R_p = radius of curvature of the plastic asperity

R_u = radius of curvature for the ultimate-stress asperity

R = radius of curvature for the original surface asperity

Z = asperity height

d = separation of the surfaces

h = separation based on surface heights

σ_s = Standard deviation corresponding to asperity

σ = Standard deviation corresponding to Surface height

y_s = The difference between h and d

N = total number of asperities

η = density of asperities

A_n = nominal area

ω^* = dimensionless interference

A_e = elastic contact area

P_e = elastic contact load

A_p = plastic contact area

P_p = plastic contact pressure

H = hardness of the softer material of the contacting surfaces

K = the hardness coefficient

E = Hertz elastic modulus

ν = Poisson ratio of the softer material

E_1, E_2 = Young's moduli of the contacting surfaces

ν_1, ν_2 = Poisson's ratios of the contacting surfaces

P_{\max} = maximum contact pressure

P^* = dimensionless contact load

Ψ = plasticity index

h^* = dimensionless mean separation

Appendix A

As illustrated in Fig. A.1, to obtain the mathematical description of the plastic asperity, the mapping of a point A on the surface to a point B on the plastic asperity must be considered. It is also noted that an asperity is described (Fig. A.1) in terms of a frame of reference whose origin is at the asperity peak and ordinate points towards the mean plane (Greenwood & Tripp, 1970). Therefore, ρ -y frame is used to describe the original asperity whereas xy -frame is employed for the plastic asperity.

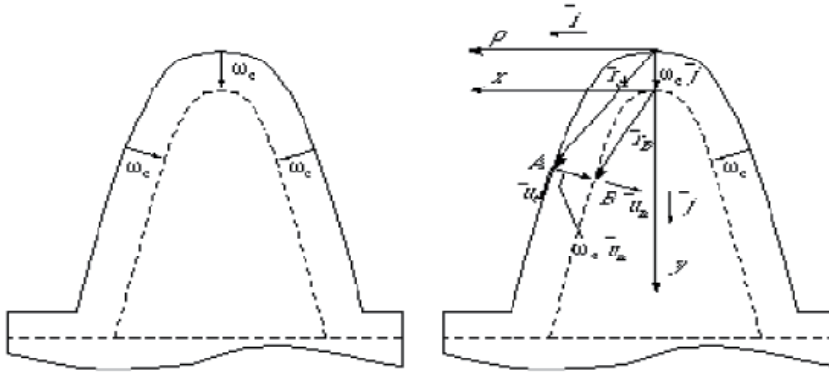


Fig A.1 - Plastic (fictitious) asperity shape

The respective positions of points A and B are denoted by \bar{r}_A and \bar{r}_B , as depicted in Fig. A.1. The position of point B on the fictitious asperity is

$$\bar{r}_B = -\omega_c \bar{j} + \bar{r}_A + \omega_c \bar{u}_n \quad (\text{A.1})$$

where, \bar{u}_n is the unit normal vector to the original asperity at point A. As usual the unit vectors \bar{i} and \bar{j} are defined along x and y axes, respectively. Employing the notation of Greenwood and Tripp

$$\bar{r}_A = \rho \bar{i} + \frac{1}{2R} \rho^2 \bar{j} \quad (\text{A.2})$$

Eqs. (A.1) and (A.2) may be used to obtain the equation describing the plastic asperity as follows:

Consider an asperity and assume that its shape is quadratic as proposed by Greenwood and Tripp. The equation of the surface asperity is given by

$$y = \frac{1}{2R} \rho^2 \quad (\text{A.3})$$

as shown in Fig. A.1. The figure also illustrates fictitious plastic asperity whose shape is obtained by a displacement of ω_c along the normal to the quadratic curve. Let \bar{u}_t and \bar{u}_n represent the tangential and normal unit vector to the quadratic at point A. The position of point A is given by vector \bar{r}_A as

$$\bar{r}_A = \rho \bar{i} + y \bar{j} = \rho \bar{i} + \frac{\rho^2}{2R} \bar{j} \quad (\text{A.4})$$

The unit tangential vector is obtained as

$$\bar{u}_t = \frac{d\bar{r}_A}{|d\bar{r}_A|} = \frac{1}{\sqrt{1 + \frac{\rho^2}{R^2}}} \left(\bar{i} + \frac{\rho}{R} \bar{j} \right) \quad (\text{A.5})$$

Hence the unit normal vector is:

$$\bar{u}_n = \frac{1}{\sqrt{1 + \frac{\rho^2}{R^2}}} \left(-\frac{\rho}{R} \bar{i} + \bar{j} \right) \quad (\text{A.6})$$

Then the description of the plastic asperity is obtained by

$$\bar{r}_B = \bar{r}_A + \omega_c \bar{u}_n - \omega_c \bar{j} \quad (\text{A.7})$$

Or

$$\bar{r}_B = \rho \left(1 - \frac{\omega_c / R}{\sqrt{1 + \frac{\rho^2}{R^2}}} \right) \bar{i} + \left(\frac{\rho^2}{2R^2} + \frac{\omega_c}{\sqrt{1 + \frac{\rho^2}{R^2}}} - \omega_c \right) \bar{j} \quad (\text{A.7})$$

For small ω_c / R and ρ / R , \bar{r}_B maybe approximated by,

$$\bar{r}_B = \rho \left(1 - \frac{\omega_c}{R} \right) \bar{i} + \frac{\rho^2}{2R} \left(1 - \frac{\omega_c}{R} \right) \bar{j} \quad (\text{A.8})$$

Let

$$x = \rho \left(1 - \frac{\omega_c}{R} \right)$$

Then

$$\bar{r}_B = x \bar{i} + \frac{x^2}{2(R - \omega_c)} \bar{j} \quad (\text{A.9})$$

Therefore the shape of the plastic asperity is given by:

$$y = \frac{x^2}{2(R - \omega_c)} \quad (\text{A.10})$$

Using the critical interference and additionally by imposing the conservation of volume, (Chang et al, 1986). derived the modified equations describing the contact area and load on an asperity:

$$A_0 = \frac{\pi a^2}{4} = \pi R \omega \left(2 - \frac{\omega_c}{\omega} \right) \quad (\text{A.11})$$

And

$$P_0 = \pi R \omega \left(2 - \frac{\omega_c}{\omega} \right) KH \quad (\text{A.12})$$

Since the plastic asperities only deform plastically, their introduction allows the reduction of equations (A.11) and (A.12) to the well known forms:

$$A_0 = 2\pi R_p \omega \quad (\text{A.13})$$

$$P_0 = 2\pi KHR_p \omega \quad (\text{A.14})$$

where, R_p represents the summit radius of curvature of the plastic asperity. Based on Eqn. (19), this radius of curvature is:

$$R_p = (R - \omega_c) \quad (\text{A.15})$$

9. References

- Abdo J. and AL-Yahmadi A. (2004), A wear model for rough surfaces based on the ultimate stress asperity concept *International Journal of Applied Mechanics and Engineering*, 9, pp. 11-19.
- Abdo J. and Farhang K. (2005), Elastic-plastic contact model for rough surfaces based on plastic asperity concept, *Int. J. Non-Linear Mech.*, 2005; 40(4), pp. 495
- Abdo J. (2006), Investigation of contact stiffness and its relation to friction-induced noise and vibration, *International Journal of Modeling and Simulation*; 26/4, pp. 295.
- Archard J. F. (1953), Contact and rubbing of flat surfaces. *J. Appl. Phys.* **24**, pp. 981-988
- Brockley C. A. and Ko P. L. (1970), The measurement of friction and friction-induced vibration. *ASME J. Lubrication Technology* **92**, 2007, pp. 543-549.

- Bush A.W., Gibson R.D., Thomas T.R. (1975), The elastic contact of a rough surface, *Wear*, V 35 (1975), pp. 87–111.
- Chang W.R., Etsion, I., and Bogy, D.B. (1986), Bogy, An elastic–plastic model for the contact of rough surfaces, *J. Tribol.* 109, (1986), pp. 257–263.
- Eid H., George G.A (2007), An elastic–plastic finite element analysis of interacting asperities in contact with a rigid flat, *J. Phys. D: Appl. Phys.* 40, 2007, 7432.
- Greenwood J.A., Williamson J.B.P. (1966), Contact of nominally flat surfaces, *Proc. Roy. Soc. (London)*, A 295 (1966), pp. 300–319.
- Greenwood J.A., Tripp J.H. (1967), The elastic contact of rough sphere, *ASME J. Appl. Mech.* 34 (1967), pp. 153–159.
- Greenwood, J. A., and Tripp, J. H (1970), The Contact of Two Rough Nominally Flat Rough Surfaces, *Proc. Instn. Mech. Engrs.*, V. 185, pp. 625–633.
- Halling J., Nuri K. A., Eds. De Pater and Kalker. (1975), The normal approach between rough flat surfaces in contact, *Wear*, Volume 32, Issue 1, pp. 81–93
- Hisakado T. (1974), Effect of surface roughness on contact between solid surfaces, *Wear*, V 28 (1974), pp. 217–234.
- Ibrahim, R. A. and Rivin, E. (1994-I), Friction-Induced Vibration, Chatter, Squeal, and Chaos; Part I: Mechanics of Contact and Friction; eds., *Applied Mechanics Reviews*, Vol. 47, 227, pp. 209–253.
- Ibrahim, R. A. and Rivin, E. (1994-II), Friction-Induced Vibration, Chatter, Squeal, and Chaos; Part II: Mechanics of Contact and Friction; eds., *Applied Mechanics Reviews*, Vol. 47, pp. 209–253.
- Ishigaki h., Kawaguchi I., Mizuta S. (1979), A simple estimation of the elastic–plastic deformation of contacting asperities, *Wear*, V 54 (1979), pp. 157–164.
- Jackson R.L., Green I., Quicksall J.J. (2004), Elasto-plastic hemispherical contact models for various mechanical properties, *Proc. Instn Mech. Engrs: J. Engineering Tribology*, Vol. 218 Part J, pp. 13–322.
- Karl, E. R., Komvopoulos, K., and Bogy, D. B. (1993), Elastic-Plastic Finite Element Analysis of Repeated Indentation of Half-Space by Rigid Sphere, *ASME J. Appl. Mech.*, 1993; 60, pp. 829.
- Kogut, L. and Etsion, I. (2002), A finite element based elastic-plastic model for the contact of rough surfaces *ASME Jour. of Appl. Mech.*, 2002; 69, pp. 657
- Kogut, L. and Etsion, I. (2003), A Finite Element Based Elastic-Plastic Model for the Contact of Rough Surfaces, *Trib. Trans.* 2003; 46, pp. 383–390.
- Kucharski, S., Klimczak, T., Polijaniuk, A. and Kaczmarek, J. (1994), Finite-elements model for the contact of rough surfaces, *Wear*, 1994; 177, 1.
- Lin L. and Lin J. (2005), An Elastoplastic Micro-asperity Contact Model for Metallic Materials, *ASME J. Tribol.*, 2005; 127, pp. 666
- Liu G. Zhu, J., Wang, Q. J. (2001), Elasto-Plastic Contact of Rough Surfaces, *Tribol. Trans.*, 44; 2001, pp. 437.
- McCool J.I. (1986), Predicting microfracture in ceramics via a microcontact model, *ASME J. Tribol.* 108 (1986), pp. 380–386.
- Mulhearn T. O. and Samuels L. E., The abrasion of metals: A model of the process, *wear*, Volume 5, Issue 6, November–December 1962, pp. 478–498
- Pullen, J., Williamson (1972), On the plastic contact of rough surfaces, *Proc. Roy. Soc. (London)*, A 327 (1972), pp.159–173.

- Samuels L. E. (1978), The mechanisms of abrasive machining, *AATA Number: 16-905*, Volume Number: 239, pp. 132-152.
- Tzou, k., Wickert, J., Akay, A (1998), In-plane vibration modes of arbitrarily thick disks. *ASME Journal of Vibration and Acoustics* **120** (1998a), pp. 384-391.
- Zhao, Y., Maietta, D. M., and Chang, L. (2000), An Asperity micro-contact model incorporating the transition from elastic deformation to fully plastic flow, *ASME J. Tribol.*, 2000; 122(1), pp. 86.

Numerical study of backward extrusion process using finite element method

K.Abrinia and S.Orangi

*School of Mechanical Engineering, College of Engineering,
University of Tehran, I.R.Iran*

1. Introduction

One of the important bulk forming processes which has a lot of applications is the backward extrusion that has been the subject of study of researchers for many years. This process is very useful for making shaped sections with closed ends or integrated parts which are hard to produce by means of other processes. Numerical analyses of bulk forming processes are difficult to perform due to the large deformations occurring during the material flow which causes heavy mesh distortions. However for those bulk forming processes which have a non steady state nature like the backward extrusion of shaped sections this difficulty increases. The amount of previous works on the numerical study of backward extrusion is therefore much less as compared with analytical research carried in the same area. Here are given a review of a few numerical analyses for the backward extrusion and then some of the analytical works would be examined as well.

Im et al (2004) gave an account of a FEM analysis for a test which was used to determine the friction factor for the bulk forming processes. This was called a tip test. In their study, a set of tip tests was carried out to determine the friction condition for the backward extrusion process of aluminum alloy. It had been experimentally observed that the radial distance from the tip to the side wall of the specimen increased with higher level of frictions and thus could be used as an effective measure of the friction condition. It was also found that the radial tip position was more sensitive to the change of friction condition at the punch than the die and that the global average friction level at the punch interface was higher than that at the die interface. Finite element simulation results also clearly confirmed that the relationships among the radial tip distance, forming load and shear friction factor were linear. Shin et al (2005) presented a new approach to process optimal design in non-isothermal, non-steady metal forming process. They investigated the development of a finite element-based approach for process optimization in forming of a difficult-to-form materials, such as titanium alloys. Due to strong dependence of the flow stress of these materials on temperature, interaction between deformation and heat transfer in the workpiece and contact heat transfer between the tool and the workpiece were rigorously reflected in the proposed approach. Also reflected was the effect of remeshing, which was essential for the treatment of a complex-shaped product. Consequently, the main merit of the proposed approach was that the diverse process parameters that could be treated as design variables,

and the applicability was not affected by the shape complexity of the product. Saboori et al (2006) presented a finite element solution for the backward extrusion process using ABAQUS commercial software. They investigated the effect of the die radius on the extrusion pressure and obtained an optimum design. In a similar paper to above Bakhshi-Jooybari et al (2006), based on the slab method and an iterative algorithm, obtained the optimum die profile in backward rod extrusion of lead. Furthermore, by using the finite element software, ABAQUS, the optimum die angle for conical die was determined. Uyyuru and Valberg (2006) studied the material flow of aluminum-alloy slug over the punch head in backward cup extrusion process by physical modeling technique combined with FE-simulations. Specially designed intrinsic tube/circle pattern made of contrast material was used for the purpose. New physical pattern technique was very useful as it allowed calculation of the extension of surface all over the inner wall of the cup with ease. Degree of surface extension over the punch head was observed to vary along the cup length. Close to the base of the cup, extension was very high compared with extension at top of the cup. Long (2006) investigated the effects of elastic-plastic deformation and temperature variations of the workpiece and tools on dimensional errors of a cold backward extruded cup. He developed a finite element analysis procedure to predict component dimensional deviations during different stages of the cold extrusion process. In an entire cold forming cycle, the material plastic deformation, heat generation, temperature distribution and tool elastic deformation in backward cup extrusion were obtained, which enabled a quantitative evaluation of the dimensional errors of the cold formed component with a view to the effects of forming stages on dimensional errors of the extruded cup. Abrinia and Gharibi (2008) investigated the backward extrusion of thin wall cups using FEM. In this investigation the extrusion pressure, the forces acting on the punch head and the thickness of the walls were studied while the shape of the punch head was varied using different profiled curves. It was shown that smaller and more uniform can wall thickness could be achieved by using a proper punch head profile.

As could be seen from the literature review of the previous work the finite element analyses of backward extrusion has been centered on products such as cups which are fairly simple geometrical shapes with axisymmetric nature. However considering the FEM analyses of the backward extrusion of shaped sections, with its three dimensional material flow, barely any relevant works could be seen in the literature and that is why the rest of this section deals with the review of the works involving upper bound and other non numerical methods.

Bae and Yang (1992) gave an account of an upper-bound method to determine the final-stage extrusion load and the deformed configuration for the three-dimensional backward extrusion of internally elliptic-shaped tubes from round billets. In another article, Bae and Yang (1993-a) presented a simple kinematically admissible velocity field for the backward extrusion of internally circular shaped tubes from arbitrarily shaped billets. They carried out experiments with full annealed aluminum alloy billets at room temperature using four circular-shaped punches. Another new kinematically admissible velocity field was proposed by Bae and Yang (1993-b) to determine the final-stage extrusion load and the average extruded height in the backward extrusion of internally non axisymmetric tubes from round billets. Lee and Kwan (1996) presented a modified kinematically admissible velocity field for the backward extrusion of internally circular-shaped tubes from arbitrarily shaped billets. From the proposed velocity field, the upper-bound extrusion load and average extruded

height for regular polygonal shaped billets were determined with respect to the chosen parameters. A new upper-bound elemental method was proposed by Lin and Wang (1997) to improve the ineffectiveness of the upper bound elemental technique (UBET) for solving forging problems that were geometrically complex or needed a forming simulation for predicting the profile of the free boundary. An upper-bound formula was developed to analyze the backward extrusion forging of regular polygon cup-shaped components in an article by Moshksar and Ebrahimi (1998). Guo et al (2000) analyzed two- and one-way axisymmetric hot backward extrusion problems by a combined finite element method. A finite element simulation for the backward extrusion of internally hollow circular sections from polygonal billets was performed by Abrinia and Orangi (2007). In this article, investigation of process parameters for the backward extrusion of arbitrary-shaped tubes from round billets was carried out using finite element approach.

The purpose of this work is to present a detailed study of the backward extrusion of shaped sections including circular and non-circular billets using finite element method. The work presented here is based on the publications of the authors given in the references and mainly on the work by Abrinia and Orangi, 2009.

2. Theory and modeling of the Process

The theory of finite element analysis for the backward extrusion process is presented here and subjects like analysis step, interaction, boundary condition and meshing are dealt with.

2.1 Analysis steps

A sequence of one or more analysis steps can be defined within a model. The step sequence makes a way to change the loading and boundary conditions of the model, changes in the way parts of the model interact with each other and the addition or removal of parts. In addition, steps allow changing the analysis procedure, the data output, and various controls. Steps may be used to define linear perturbation analyses about nonlinear base states.

The initial step allows one to define boundary conditions, predefined fields, and interactions that are applicable at the beginning of the analysis.

The initial step is followed by one or more analysis steps. Each analysis step is related with a specific procedure that defines the type of analysis to be performed during the step, such as a static stress analysis or a transient heat transfer analysis. One can change the analysis procedure from step to step in any desired way, so, there will be great flexibility in performing the analyses. Considering the fact that the state of the model (stresses, strains, temperatures, etc.) is updated throughout all general analysis steps, the effects of previous history are always included in the response for each new analysis step.

Between two main nonlinear methods of solutions that are often used in cold forming process simulations; namely the explicit dynamic and the implicit Newton incremental methods, the explicit dynamic method was chosen for all the simulations carried out in this work.

The explicit dynamic analysis method is based on the implementation of an explicit integration rule (Fig. 1) by using the diagonal or lumped element mass matrices. The basic motion equation of the body is:

$$M\ddot{u} + F(u) = 0$$

Now integrating and using the explicit central difference integration rule, we have:

$$\dot{u}^{(i+\frac{1}{2})} = \dot{u}^{(i-\frac{1}{2})} + \frac{\Delta t^{(i+1)} + \Delta t^{(i)}}{2} \ddot{u}^{(i)},$$

$$u^{(i+1)} = u^{(i)} + \Delta t^{(i+1)} \dot{u}^{(i+\frac{1}{2})},$$

where \dot{u} is the velocity and \ddot{u} is the acceleration. The superscript (i) refers to the increment number and $i - \frac{1}{2}$ and $i + \frac{1}{2}$ are middle increment values.

The use of the diagonal element mass matrices is the computational efficiency key to the explicit dynamics procedure because the inversion of the mass matrix that is used in the computation for the accelerations at the beginning of the increment is trivial.

$$\ddot{u}^{(i)} = M^{-1} \cdot (F^{(i)} - I^{(i)})$$

where M is the diagonal lumped mass matrix, F is the applied load vector and I is the internal force vector.

Iterations and tangent stiffness matrix construction are not required in this procedure. The explicit procedure integrates through time by employing small time increments. The central difference operator is conditionally stable, and the stability limit for the operator (with no damping) is given by the maximum eigenvalue in the system as:

$$\Delta t \leq \frac{2}{\lambda_{\max}}$$

Its advantage is its simplicity and it is straightforward to implement. However, the disadvantages are: firstly that it is conditionally stable; i.e. it may become unstable. Secondly, the accuracy of the integration depends, of course, on the time step size, Δt . In comparison with other analysis such as static and implicit methods, the time and cost of explicit method is lower than that of the others. The reason is that the explicit method determines the solution without iterating and it doesn't require large numbers of time increments such as the case for the static analysis. These are the most important advantages when dealing with large deformations such as the one for the backward extrusion process.

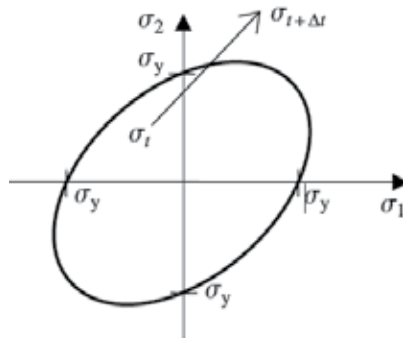


Fig. 1. Schematic representations of explicit integration using the radial return method, of Von Mises plasticity equations (Wriggers, 2008).

2.2 Interaction (friction)

Cold forming process systems involve contact between the work piece and the tools. Contact itself is a typical type of nonlinear boundary conditions encountered in finite element analysis. Contact conditions are a special class of discontinuous constraints allowing forces to be transmitted from one part of the model to another. This constraint is discontinuous because it is applied only when two surfaces are in contact. When the two surfaces are separated, no constraint is applied.

The interaction between contacting surfaces consists of two components: one normal to the surfaces and the other tangential to the surfaces. The tangential component consists of the relative motion (sliding) of the contacting surfaces and associated frictional shear stresses.

In addition to determining whether contact has occurred at a particular point, the analysis also must calculate the relative sliding of the two surfaces of the contact pair and their tangential interaction. There exist two sets of formulations in contact modeling: small sliding and finite sliding formulations. Small sliding formulation is used when the magnitude of sliding is small. In contrast, finite sliding formulations are used when the magnitude of sliding is large and finite. It is much less expensive computationally to model problems where the sliding between the surfaces is small. Finite sliding is the most general and allows arbitrary motion of the surfaces forming the contact pair. In cold forming process simulation, since usually there is large relative motion between the work piece and tools due to the punch and tool movement or because of deformation and shape changes of the work piece, the finite sliding model is often used between them. Small sliding model can be used between two tools where there is no or little relative motion.

Coulomb friction is the friction model used to describe the interaction of contacting surfaces. In the basic form of the Coulomb friction model, two contacting surfaces can carry shear stresses up to a certain magnitude across their interface before they start sliding relative to one another; this state is known as sticking. The Coulomb friction model defines this critical shear stress, τ_{crit} , at which sliding of the surfaces starts as a fraction of the contact pressure,

p , between the surfaces ($\tau_{crit} = \mu p$). The stick/slip calculations determine when point transition from sticking to slipping or from slipping to sticking occurs. The fraction, μ , is known as the coefficient of friction. The basic friction model assumes that μ is the same in all directions (isotropic friction). For a three-dimensional simulation there are two local orthogonal components of shear stresses, τ_1 and τ_2 , along the interface between the two bodies. We may combine the two shear stress components into an equivalent shear stress, $\bar{\tau}$, for the stick/slip calculations, where $\bar{\tau} = \sqrt{\tau_1^2 + \tau_2^2}$. In addition, the two slip

velocity components can be combined into an equivalent slip rate, $\sqrt{\dot{\gamma}_1^2 + \dot{\gamma}_2^2}$.

There are three contact algorithms used in ABAQUS. The first one is an algorithms used in the standard analysis and built around the Newton-Raphson technique. The second one is a kinematic contact algorithm and the third one is a penalty contact algorithm. The second and third algorithms mentioned above are used in ABAQUS/explicit. ABAQUS/explicit uses a kinematic contact formulation that achieves precise compliance with the contact conditions using a predictor/corrector algorithm. We applied penalty contact algorithm in our simulations because it is more sensitive to element distortion than the kinematic

algorithms. In ABAQUS software, contact is determined by defining two surfaces namely the slave and the master surfaces. Contact is applied on slave nodes (101,102,103) and master nodes (1,2,3). In the case of work piece and die, work piece is considered as slave and die as master (Fig. (2)).

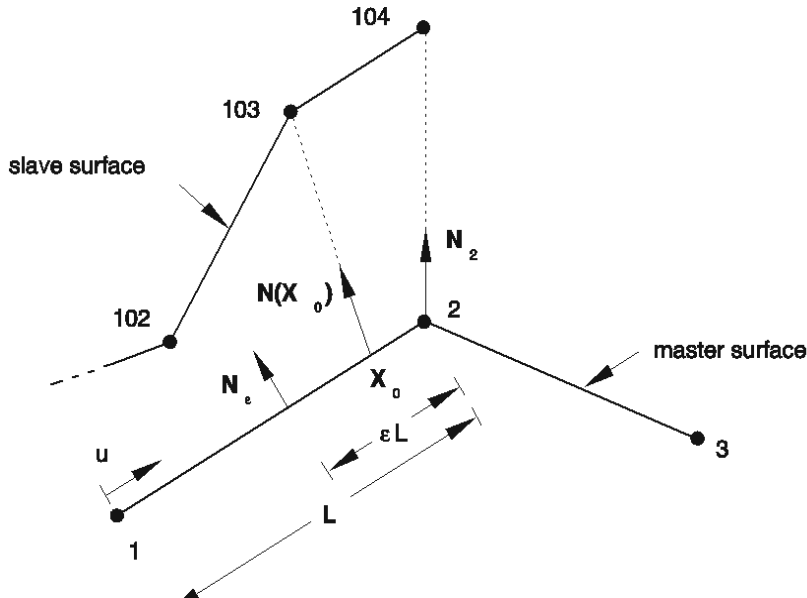


Fig. 2. Slave and master surfaces

The constitutive behavior in tangential direction is very complex. It depends on many factors such as surface roughness, magnitude of the normal pressure, tangential relative velocities, contaminants or humidity, etc. This complexity can be shown by the constitutive equations which then depend on many material parameters. Since it is not easy to determine such parameters, the simplest constitutive relation, the so-called Coulomb law is employed in many engineering applications which only depend upon one material parameter, the coefficient of friction. However, depending on the material pairing, there exist many variants of this constitutive relation for the tangential stresses.

2.3 Load (boundary condition):

Boundary condition can be used to specify the values of all basic solution variables (displacements, rotations, warping amplitude, fluid pressures, pore pressures, temperatures, electrical potentials, normalized concentrations, acoustic pressures, or connector material flow) at nodes;

There are various ways to define boundary conditions in finite element. These are based on the followings:

- Symmetry/Anti-symmetry
- Displacement/Rotation
- Velocity/Angular velocity
- Acceleration/Angular acceleration
- Connector displacement or velocity or acceleration

We have implemented velocity/ angular velocity as the boundary condition. The definition is given in terms of the angular velocity instead of the total rotation if velocity-type boundary conditions are used to prescribe rotations.

2.4 Meshing

Behavior of an element is characterized by five aspects, namely: family, degrees of freedom (directly related to the element family), number of nodes, formulation and integration. The fundamental variables calculated during the analysis are the degrees of freedom. For a stress/displacement simulation the degrees of freedom are the translations.

2.5 Three dimensional elements

Three-dimensional elements are defined in the global X, Y, and Z space. These elements are utilized when the geometry and or the applied loading are too complex for any other element type with fewer spatial dimensions.

2.5.1 Stress/displacement elements

Stress/displacement elements are used in the modeling of linear or complex nonlinear mechanical analyses that may involve contact, plasticity, and/or large deformations. Stress/displacement elements can be used in the following analysis types: static and quasi-static analysis, implicit transient dynamic, explicit transient dynamic, modal dynamic, and steady-state dynamic analysis, so, we chose stress/displacement elements for the analysis. Hence, we used stress/ displacement elements in our modeling.

In explicit analysis, triangular, tetrahedral, and wedge elements can be used. Solid elements can include first- and second-order triangles, tetrahedral, and wedge elements for planar, axisymmetric, and three-dimensional analysis.

Triangular and tetrahedral elements are utilized in automatic meshing algorithms. These elements ease meshing process of complicated shapes and they are common in general applications. The disadvantage of these elements is that they are less sensitive in distortion. In contrast, hexahedral and quadrilateral elements have high accuracy computationally. Besides these elements have good performance and accuracy because of their sensitivity to distortion. In addition, hexahedral and quadrilateral elements have good convergence in comparison with tetrahedral and triangular ones. This is the reason that in the present analysis hexagonal elements have been used to analyse the backward extrusion of aluminum alloys.

2.5.2 Adaptive meshing

Backward extrusion process involves large amounts of permanent deformation. This high degree of deformation may cause problems during the finite element simulation. The analysis may easily end as a result of severe mesh distortion. Therefore, the special meshing techniques are required to analyse these problems. Adaptive meshing and remeshing or rezoning are two solutions to these problems.

Remeshing or rezoning: when the mesh is distorted severely one can pause the process of analysis and then by creating new mesh and mapping the solution onto the new mesh, the analysis will continue. For instance in mesh control one can change the amount of distortion

in terms of the length ratio (i.e. the ratio of the length of the initial mesh to that of the distorted mesh) that is one can control distortion during the process.

Adaptive meshing is another powerful tool used to maintain high-mesh quality throughout the process of analysis. Without changing the mesh topology, adaptive meshing moves the nodes only. For cases of large deformation, the mesh is allowed to move independently of the material and thus keeps away from excessive distortions. In addition, in regions where large gradients are more probable to occur, adaptive meshing allows the mesh to be refined automatically in those regions. The refining algorithm is as given below:

- Select an appropriate initial mesh which approximates the geometry of the problem accurately.
- Solve the discrete problem.
- Compute error estimators or error indicators.
- Test, whether the global error lies within the given tolerance.
- If yes, the computation is finished.
- If no, a new mesh has to be created.
- Throughout this process the former computed deformations and internal variables have to be projected onto the new mesh.

3. Modelling the Material Behaviour

This section presents the models for the behavior of the material and its definition in the FEM environment. Other FEM parameters which influence the modeling behavior will also be discussed. The following topics are covered:

- ✓ The part module
- ✓ The property module
- ✓ The Assemble module
- ✓ The step module
- ✓ The interaction module
- ✓ The load module
- ✓ The mesh module
- ✓ The job module.

3.1 The part module

In the part module, one can model geometrically the parts such as the container, die, work piece and the punch or may import parts modeled in other modeling software. Punches, containers and billets were modeled in various shapes. In all cases, billets were modeled deformable, 3D and extrusion option was used for this modeling. For elliptical, rectangular, and hexagonal tubes, the diameter of circular billet was 25 mm with a length of 25 mm. Besides, billet dimensions for the AA1100-O were 22 mm width from side to side and 20 mm length.

3.2 The property module

The property module performs the following tasks:

- Defines the materials.
- Defines the beam section profiles.

- Defines the sections.
- Assigns sections, orientations, normals, and tangents to the parts.
- Defines the composite layups.
- Defines skin reinforcement.
- Defines inertia (point mass, rotary inertia, and heat capacitance) on a part.
- Defines springs and dashpots between two points or between a point and ground.

Material definition: in all models, material was considered as Von Misses rigid-plastic material with isotropic hardening. Aluminum AA2024-O and AA1100-O were chosen as the working material. The material properties for the simulated models were assumed as follows:

For AA2024-O [2]: $\bar{\sigma} = 292.77 \times \bar{\epsilon}^{0.15} (Mpa)$

$$\rho = 2780 (kg / m^3)$$

$$E = 73.1 (Gpa)$$

$$\nu = 0.33$$

$$\sigma_y = 78.5 (Mpa)$$

And for AA1100-O:

$$\bar{\sigma} = 136 \times \bar{\epsilon}^{-0.254} (MPa)$$

$$\rho = 2715 (Kg / m^3)$$

$$E = 68.9 (GPa)$$

$$\nu = 0.33$$

$$\sigma_y = 34.5 (MPa)$$

The billet was modeled as a von Mises rigid-plastic material with isotropic hardening material property and the extrusion container and die were modeled as perfectly rigid.

3.3 The assembly module

After geometrical modeling and assigning properties to parts, if it is necessary to assemble them, assembling is performed in the assembly module. For instance, the gap between container and billet and the distance between punch and billet for the beginning of the process were adjusted here. These dimensions were assumed 0.2 mm and 0.5 mm respectively.

3.4 The step module

Step module tasks are:

- Creating analysis steps
- Specifying output requests
- Specifying adaptive meshing
- Specifying analysis controls
- Creating analysis steps

The time of the process, type of analysis; as an example, stress/displacement were taken in this part.

3.5 Interaction module

The Interaction module is used to define and manage the following objects:

- Mechanical and thermal interactions between regions of a model or between a region of a model and its surroundings.
- Analysis constraints for the regions of a model.
- Assembly-level wire features, connector sections, and connector section assignments to model connectors.
- Inertia (point mass, rotary inertia, and heat capacitance) on regions of the model.
- Cracks on regions of the model.
- Springs and dashpots between two points of a model or between a point of a model and ground.

In our model, contact was chosen as mechanical interaction, tangential behavior with penalty formulation. The friction factors, 0.1 and 0.2 are used for all models.

3.6 The load module

The load module is used to define and manage the following prescribed conditions:

- Loads
- Boundary conditions
- Predefined fields

In our model, the blank was constrained along its base in the z-direction and at the axis of symmetry in the r-direction. Radial expansion was prevented by contact between the blank and the die. The punch and the die were fully constrained, with the exception of the prescribed vertical motion of the punch.

3.7 The Mesh module

The Mesh module contains tools that allow us generating meshes on parts and assemblies created within Abaqus/CAE. Besides, there are functions that verify an existing mesh. Various levels of automation and control are available here. As with creating parts and assemblies, the process of assigning mesh attributes to the model—such as seeds, mesh techniques and element types—is feature based. As a result we can modify the parameters that describe a part or an assembly, and the mesh attributes that we specified within the mesh module are regenerated automatically.

The Mesh module ensures the following features:

- Tools for prescribing mesh density at local and global levels.
- Model coloring that shows the meshing technique assigned to each region in the model.
- Mesh controls, such as element shape, meshing technique, meshing algorithm, adaptive remeshing rule

- A tool for assigning Abaqus/Standard and Abaqus/Explicit element types to mesh elements. The elements can belong either to a model that you created or to an orphan mesh.
- A tool for verifying mesh quality.
- Tools for refining the mesh and for improving the mesh quality.
- A tool for saving the meshed assembly or selected part instances as an orphan mesh part

Element failure criteria for all models in our simulation: the optimum element size was obtained by trial and error until the appropriate aspect ratio (i.e.3) was achieved

- Type of element for the container modeled as discrete rigid: R3D8R
- Type of element for the punch modeled as analytical rigid: R3D4, A 4-node 3D bilinear rigid quadrilateral.
- Element type for the billet modeled as deformable: C3D8R, An 8-node linear brick element, reduced integration with hourglass control.
- Element Library: Explicit, 3D stress.
- Partitioning

Element shape: Hexahedral with sweep technique

In some cases during the analysis, distortion control was applied during the process which was considered as distortion length=0.5.

3.8 The Job module

Creating and managing analysis jobs are possible by the Job module and also to view a basic plot of the analysis results. We can use the Job module to create and manage adaptivity analyses as well.

Once all of the tasks involved in defining a model have been finished (such as defining the geometry of the model, assigning section properties, and defining contact), Job module is used to analyze our model. This module allows us to create a job, to submit it to Abaqus/Standard or Abaqus/Explicit for analysis, and to monitor its progress. If desired, creating multiple models and jobs and running and monitoring the jobs simultaneously is possible.

In backward extrusion, the area reduction is defined as follows:

$$RA\% = \frac{a}{A}(100)$$

where a is the cross sectional area of the punch and A the cross sectional area of the billet. The gap between the billet and the container is neglected because the accuracy of calculations was 0.1mm. Effect of this gap in area reduction is about 0.025-0.05 mm²; thus it had no considerable influence on the area reduction and was omitted in this equation.

In addition, friction factors for all simulations regarding AA1100-O were taken as 0.2 and for the AA2024-O, as 0.1 obtained from experimental data in references [2] and [4].

4. Analysis of the backward extrusion

Backward extrusion of shaped sections has been investigated in here using FEM simulation. A general schematic view of the backward extrusion process has been shown in figure 3.

Different cases of shaped extrusions simulated in this work have also been illustrated in figure 4.

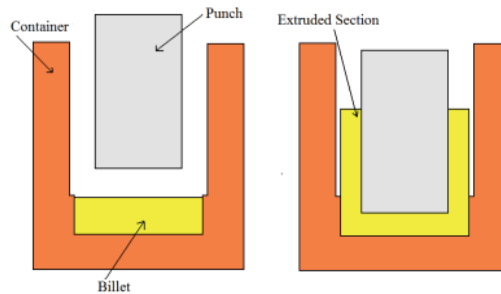


Fig. 3. A schematic view of the backward extrusion process

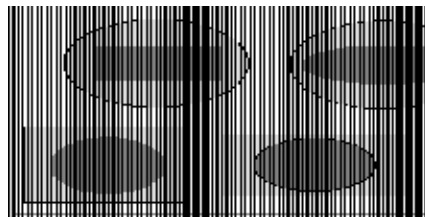


Fig. 4. Different cross sections simulated for the backward extrusion of shaped sections

ABAQUS\Explicit CAE commercial software was used for all simulations. The results are presented here according to the billet-punch shape (figure (4)).

4.1 Circular Billet-Elliptic Punch

4.1.1 Extrusion Load

Fig. 5 shows the effect of area reduction on the extrusion load for backward extrusion of internally elliptical-shaped tubes from round billets. The extrusion load increases with increasing reduction of area for a fixed aspect ratio and the given friction factor. FEM simulation results were compared with the experimental and theoretical values (Bae and Yang, 1992) and good agreements were found.

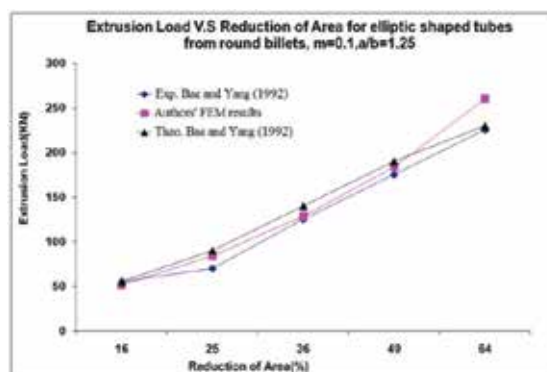


Fig. 5. Effect of area reduction on the extrusion load for the backward extrusion of elliptical section from round billet

For reduction of area of 36-49%, the extrusion loads were in good agreement with the experimental loads. The maximum difference between the value of the load obtained from the simulation and the experimental value was seen to be for the 64% area reduction.

The present simulation results for the extrusion load versus various aspect ratios in Fig. 6 are in good agreement with the experimental values and theoretical predictions (Bae and Yang, 1992). For aspect ratios of 1 to 1.3, the simulation results are closer to experimental ones. Fig. 7 shows the simulation results for the effect of friction factor on the extrusion load. For a given reduction of area and fixed aspect ratio, the extrusion load increases with the increase in the friction factor.

4.1.2 Distribution of Velocity and Configuration of the Free Surface of the Extruded Billet

The simulation results for the velocity distribution of the extruded billet are shown in Fig. 8-1 to Fig. 8-3 for different area reductions, aspect ratios, and friction factors, respectively. The effect of area reduction on the configuration of the free surface and velocity field contours of extruded billet for a given aspect ratio and the given friction factor are shown in Fig. 8-1. Fig. 8-2 shows the effect of aspect ratio on the configuration of the free surface and velocity field contribution of extruded billet for a fixed area reduction and the given friction factor. In Fig. 8-3 is shown the effect of friction factor on the configuration of the free surface and velocity field for the extruded billet having a constant aspect ratio and reduction of area. Diagrams of the velocity variables versus various area reductions, aspect ratios, and friction factors are shown in Fig. 9-1, Fig. 9-2 and Fig. 9-3, respectively.

In fig. 9-1, for a reduction of area smaller than 25%, the velocity could be seen to decrease with increasing reduction of area and for values of reduction of area 25-64%, it could be seen that the velocity increases with increasing reductions of area.

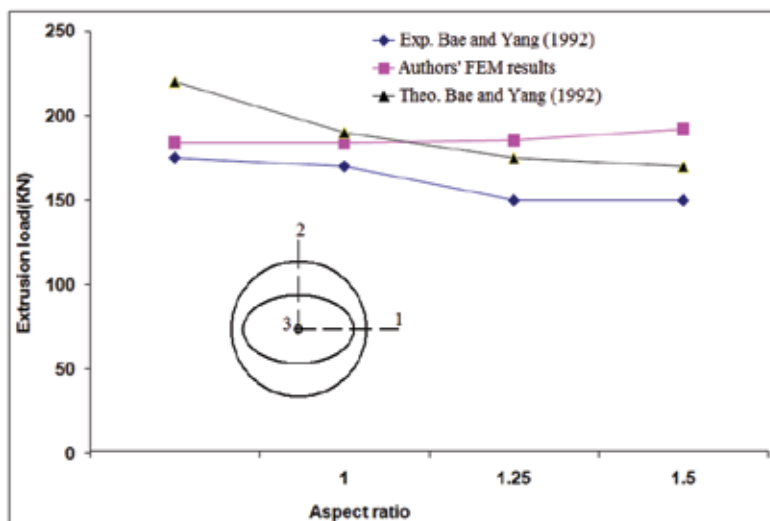


Fig. 6. Comparison of results with previous works for the backward extrusion of elliptical section from round billet

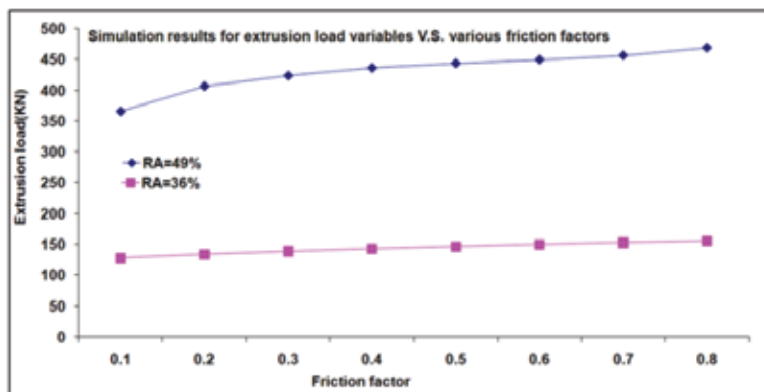


Fig. 7. Effect of frictional conditions on the extrusion load for the backward extrusion of elliptical section from round billet

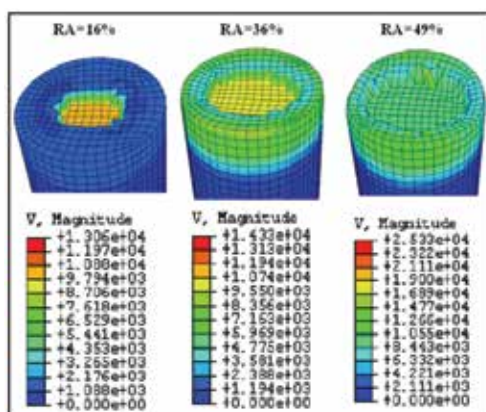


Fig. 8-1. Velocity distribution for the backward extrusion of elliptical sections from round billets-effect of reduction of area

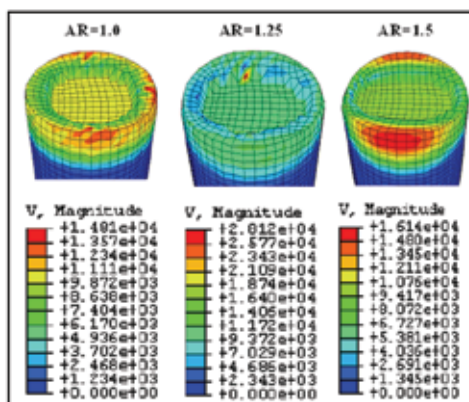


Fig. 8-2. Velocity distributions for the backward extrusion of elliptical sections from round billets-effect of aspect ratio

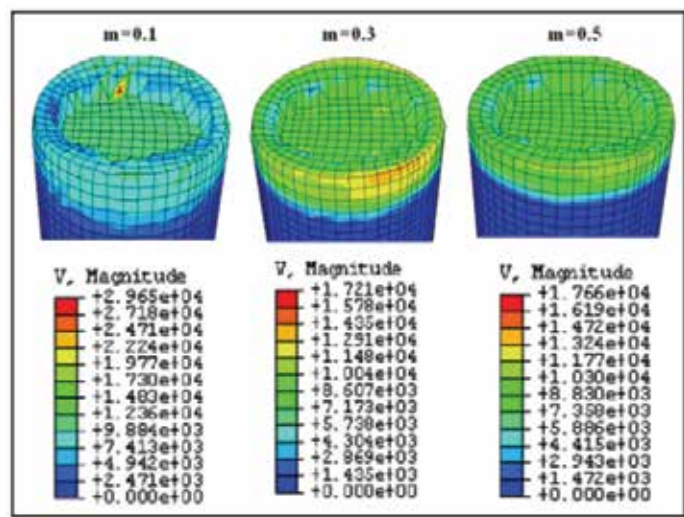


Fig. 8-3. Velocity distributions for the backward extrusion of elliptical sections from round billets-effect of frictional conditions

In Fig. 9-2, the effect of aspect ratio on the velocity has been shown. As could be seen from figure components of the velocity in directions 1 and 2 do not change appreciably with the aspect ratio while the component of the velocity in the axial direction (direction 3) changes very much with a variation in the aspect ratio and it is this component that has the major share of the total magnitude of the velocity. In Fig. 9-3, the effect of the friction factor on the velocity field has been shown. Except for very small values of the friction factor, there seems to be no appreciable difference between the different components of the velocity field. Simulation results for grid deformation could be seen in Fig. 10. Because of thinning in the work piece wall for a RA = 64%, the material flows toward punch and is distorted excessively; therefore, the material flow encounters difficulties.

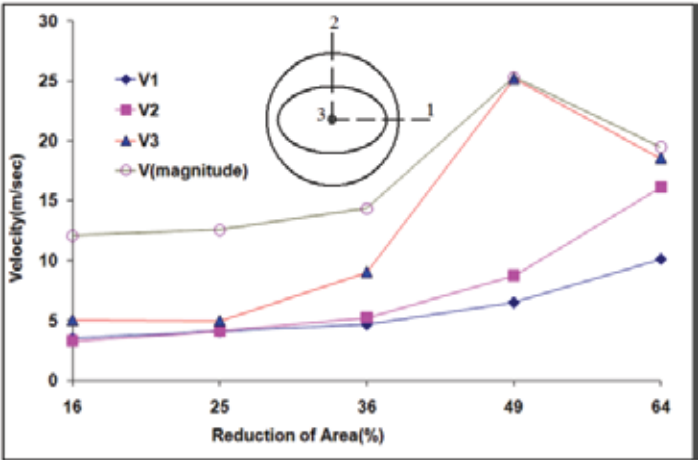


Fig. 9-1. Components of the velocity distribution- effect of reduction of area

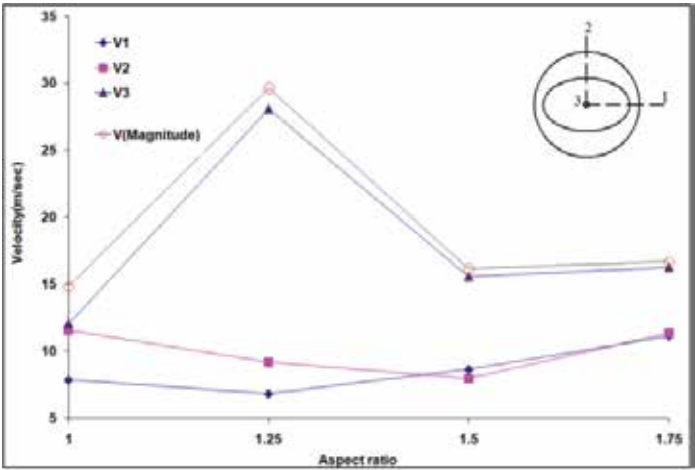


Fig. 9-2. Components of the velocity distribution- effect of aspect ratio

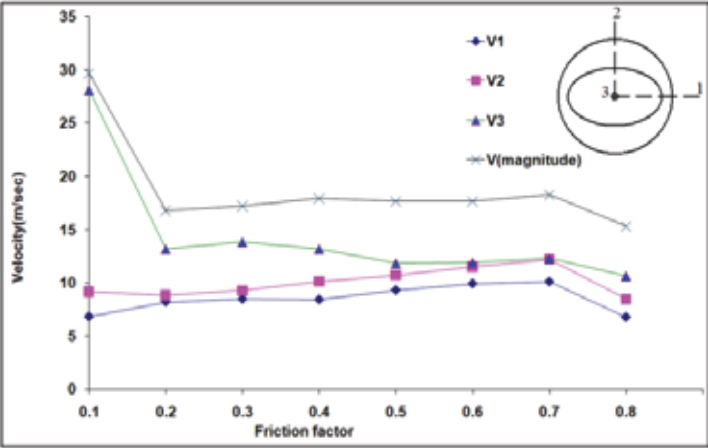


Fig. 9-3. Components of the velocity distribution- effect of frictional conditions

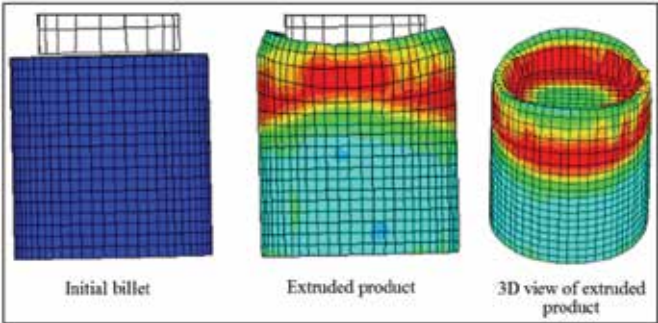


Fig. 10. Grid deformation for the backward extrusion of elliptical section from round billets

4.2 Circular Billet-Rectangular Punch Results

4.2.1 Extrusion Load

Fig. 11, Fig. 12 and Fig. 13 show the results of FEM simulation for internally rectangular sections from round billets. In Fig. 11, a graph of extrusion load versus reduction of area is given for the backward extrusion of internally rectangular sections from circular billets. Comparison of other theoretical and experimental results with the FEM simulation of the authors is also seen in this figure. Authors' results show better agreement with the experimental values at lower values of the reduction of area. In Fig. 12, a graph of extrusion load versus aspect ratio is given for the backward extrusion of internally rectangular sections from circular billets. Authors' results show better agreement with the experimental values at lower values of aspect ratio. The variation of the extrusion load with frictional conditions has been simulated and presented in figure 13 which shows the expected trend.

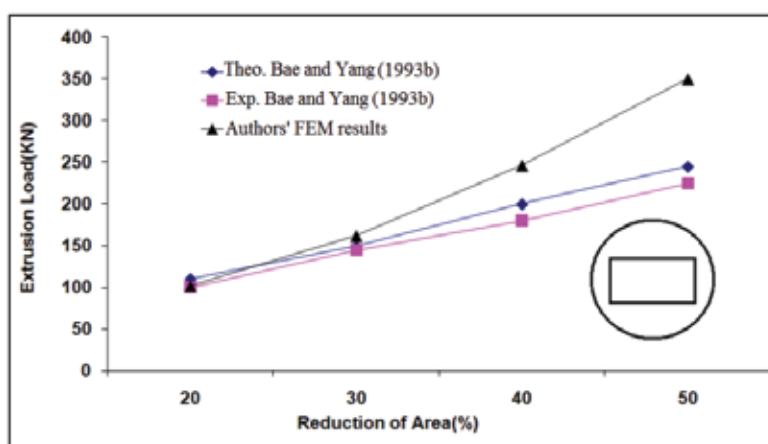


Fig. 11. Extrusion load versus reduction of area for the backward extrusion of rectangular sections from round billets

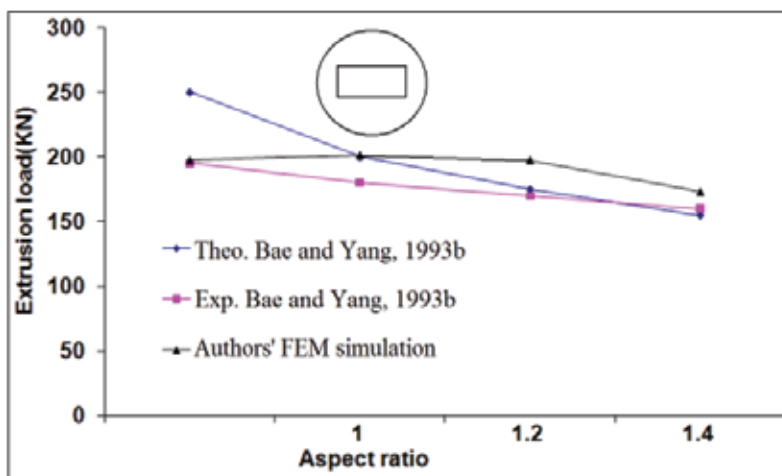


Fig. 12. Extrusion load versus aspect ratio for the backward extrusion of rectangular sections from round billets

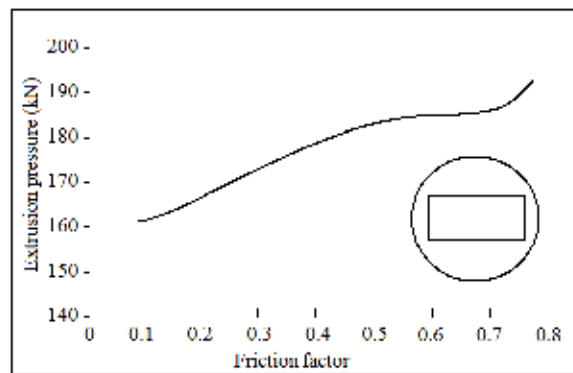


Fig. 13. Extrusion load versus friction factor for the backward extrusion of rectangular sections from round billets

3.2.2 Velocity Field Distribution

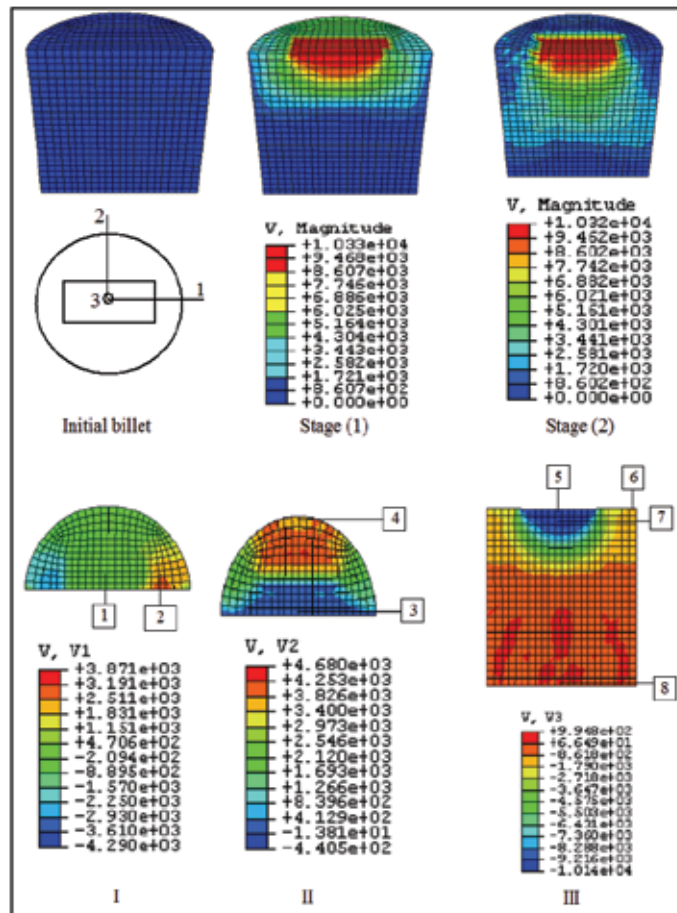


Fig. 14. Velocity distributions for the backward extrusion of rectangular section from circular billet

For the same extrusion, the grid deformation and spatial velocity distribution at nodes are also given in Fig. 14. Starting from the initial stage of the extrusion process and going through to the end, five different stages are observed. The values of velocity distributions in the part are also shown for each stage. The velocity field is shown by three directional velocities: V1, V2, and V3. The magnified figure of one stage for contour distribution of V1, V2, and V3 is shown in Fig. 14, respectively. The values of V1 were greater at locations nearer to the punch corner than elsewhere and it is lowest at the center of free surface of the extruded part at position (1). Also, V1 decreases from position (1) to (2), for the nodes near the wall of container. Values of V2 were greater at locations (3) and (4) and between these sections, V2 has the lowest value. The value of V3 was greatest at position (5) and decreases as one move toward (6), for the nodes of the work piece near the container wall. From position (7), near the free surface of workpiece toward position (8), for the nodes near the container lower section, V3 decreases gradually. These results were not obtained by other methods in other references.

4.3 Circular Billet-Hexagonal Punch Results

4.3.1 Stress and Strain Distribution

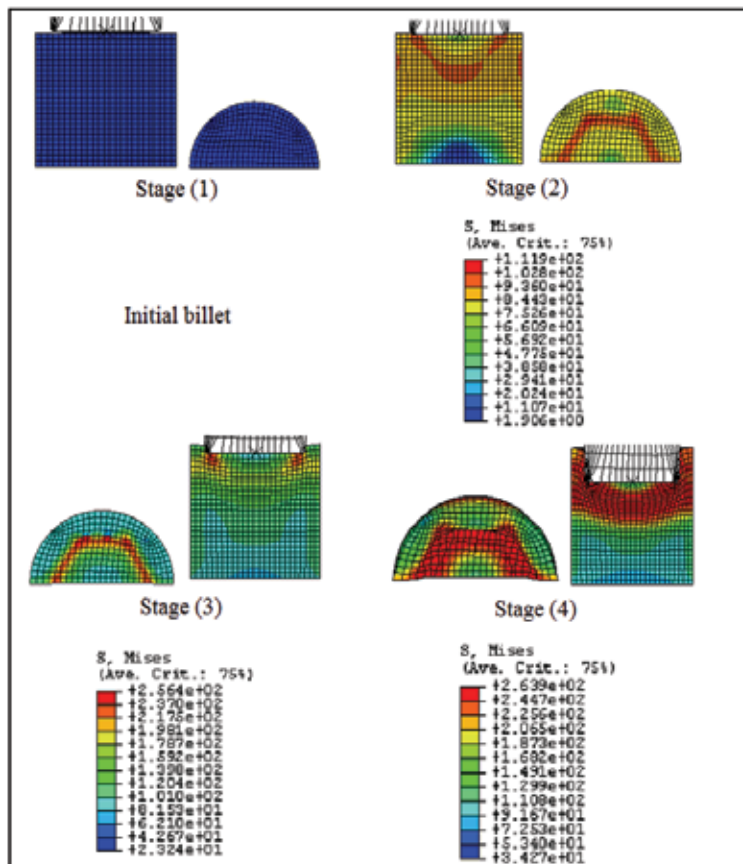


Fig. 15. Stress distribution for the backward extrusion of rectangular sections from round billets

For the same extrusion, the grid deformation and stress distribution are also given in Fig. 15. Starting from the initial stage of the extrusion process and going through to the end, four different stages are observed. The values of stress distribution in the part are also shown for each stage. A reduction of area of 49% was used for these simulations. It is illustrated that in which part of the billet the stress is greatest. The results for the equivalent strain distribution for the backward extrusion of internally hexagonal and externally circular shapes are shown in Fig. 16 for the four stages of the simulation. It could be seen that, as expected, the highest values of strain occur at where the deformation is at its highest level. As shown in these stages, because of maximum deformation in the surface of the work piece which is in contact with the corner of the punch, the plastic equivalent effective strain is greater than elsewhere. Also, in the initial stage of the process, contours of strains are distributed in broad areas and gradually decrease through the end of the process. The reason is that in the initial stage, with increasing stress, the plastic area develops and includes larger regions of the billet. However, after material flows and plastic deformation develops further, the deformation area is fixed and strains confine in the plastic areas. It should be mentioned that these results are very similar to those given in (Bae and Yang, 1993a).

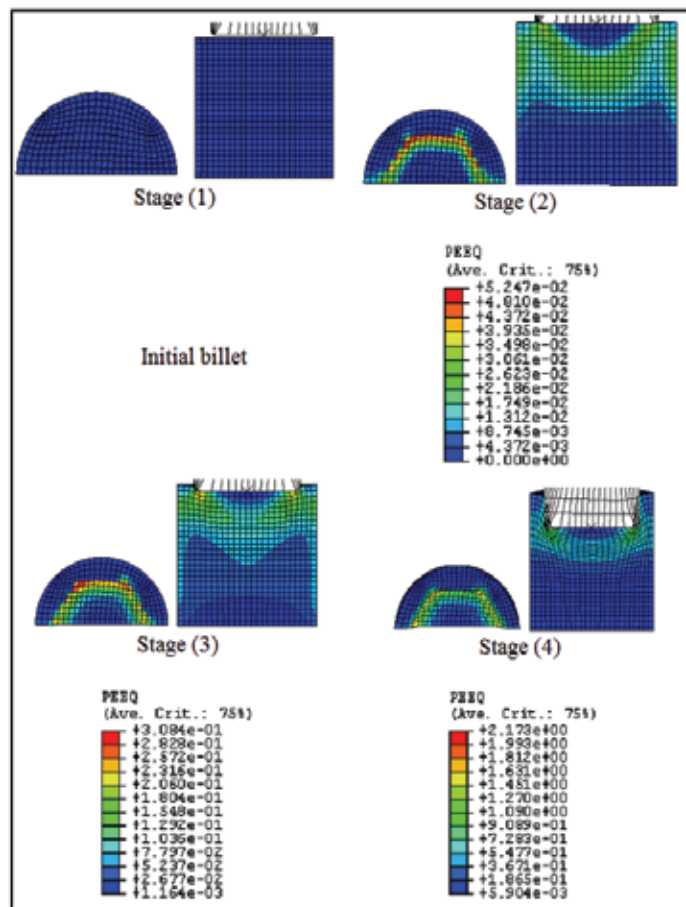


Fig. 16. Strain distribution for the backward extrusion of rectangular sections from round billets

4.4 Square Billet- Circular Punch Results

4.4.1 Extrusion load

Fig. 17 illustrates the area reduction influence on the extrusion load in backward extrusion process of internally circular-shaped tubes from square billets. The extrusion load increases with increasing area reduction for the given friction factor. Comparison of FE simulation with the experimental and theoretical values showed good agreement.

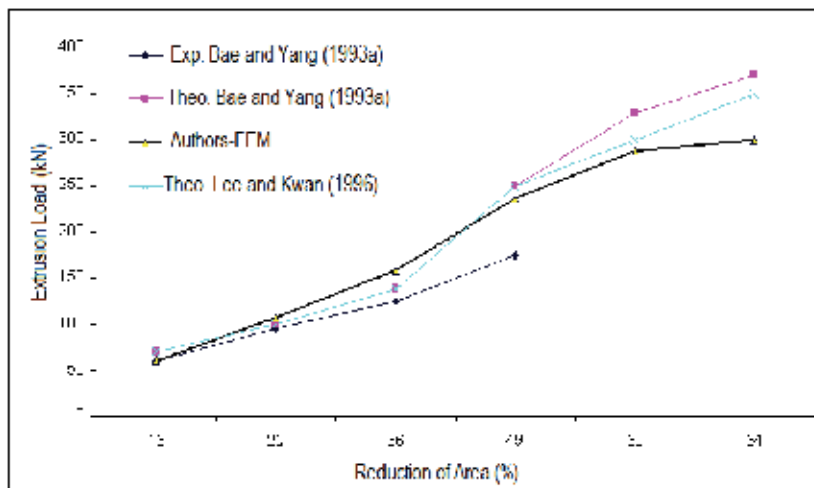


Fig. 17. Extrusion load versus reduction of area for the extrusion of internally circular and externally square shaped section

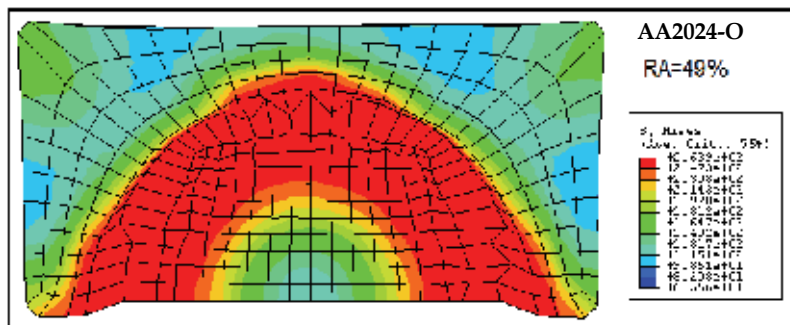


Fig. 18. Grid deformation for the extrusion of internally circular and externally square shaped sections

4.4.2 Stress distribution and grid deformation

A stress distribution contour of the backward extrusion of circular tubes with square sections is indicated in Fig. 18 as FE simulation result. The Von-Mises stress increases from the center of the section towards the internal wall of work piece then it diminishes toward external wall. It seems that because of material flows under the punch and the frictional velocity discontinuities, the stress increases. However as the material flow develops further towards the external surface it passes over the punch area and since there are no obstacle in

the path of the material flow it decreases at this position. In the vertical direction of the deformation area and under the punch, the conical shaped dead zone is created which acts rigidly and moves downward with the punch like part of it.

4.5 Hexagonal Billet- Circular Punch Results

4.5.1 Distribution of velocity and configuration of the free surface of the extruded billet

The Simulation results for the influence of the reduction in area on the velocity field have been shown in Fig. 19. It could be observed that as the reduction in area increases from 16

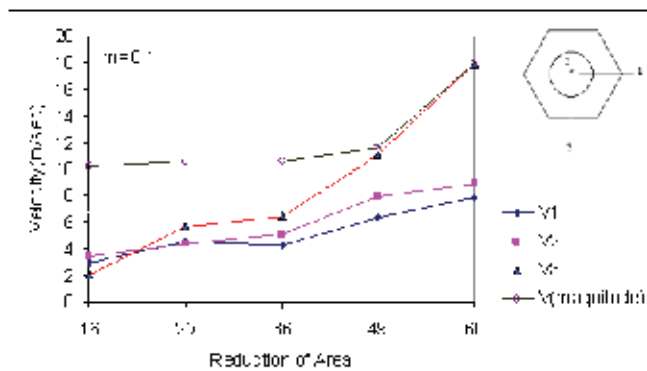


Fig. 19. Velocity components for the extrusion of internally circular and externally hexagonal shaped sections

to 49 percent, all three components of velocity (V_1 , V_2 , and V_3) have similar magnitudes but at higher reductions the component V_3 suddenly rises sharply. The reason for this seems to be the shape of the circular tube which is being extruded. At lower reductions, the thickness of tube walls is such that the flow of material happens evenly in all directions. However at 60% area reduction, it is clear that the components of velocity in the directions (1) and (2) follow the same trend as before while the component of velocity in direction (3) increases with sharp trend. This phenomenon can be explained by thinning of extruded wall which causes faster flow of material in this direction.

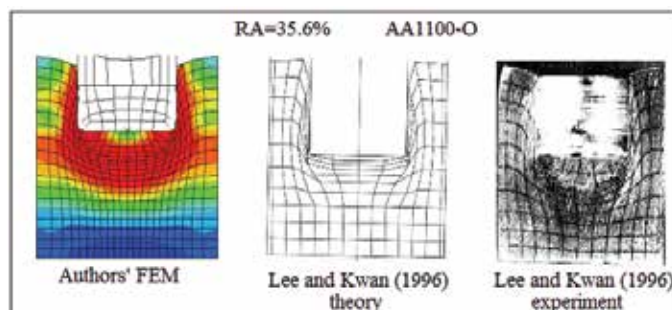


Fig. 20. Grid deformation for the extrusion of internally circular and externally hexagonal shaped sections

4.5.2 Grid deformation

Fig. 20 illustrates the grid deformation for the backward extrusion of internally circular shaped tubes from hexagonal billets, using FE simulation as compared with previous experimental and theoretical results (Lee and Kwan, 1996). FE outcomes are closer to the experimental results than the upper bound solutions. As illustrated in the figure, maximum deformation is observed in the corners and walls of the workpiece.

4.6 Rectangular Billet- Circular Punch Results

4.6.1 Strain distribution

Four stages of backward extrusion process have been simulated in the Fig. 21. Plastic equivalent effective strain contours of internally circular and externally rectangular shaped tubes are shown in this figure. It could be seen that, as expected, the highest values of strains occur at where the maximum deformation exists. The plastic equivalent effective strain in the areas where workpiece contacts with punch corner is greater than elsewhere. Consequently, maximum deformation happens in the zone of the workpiece in contact with the punch corner. In addition, in the initial stage of the process, contours of strains are distributed in broad areas and gradually decrease through the end of process.

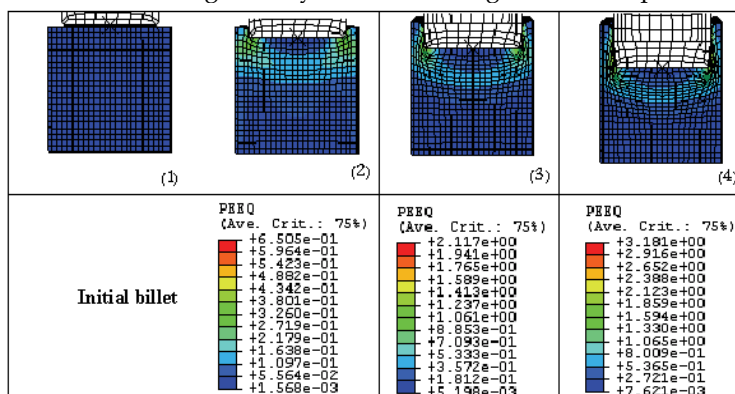


Fig. 21. Four stages of backward extrusion process for a rectangular billet and a circular punch giving the effective strain contours

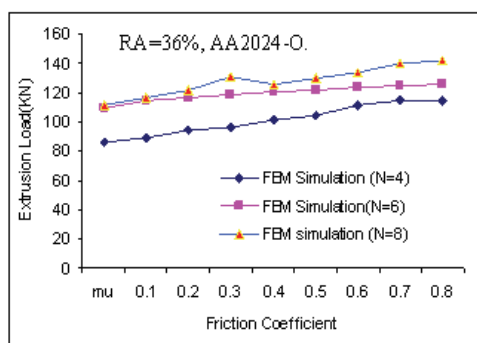


Fig. 22 Comparison of results for the backward extrusion of different polygonal sections.

This is because, in the initial stage, with increasing stress, the plastic area develops and includes larger region of billet, however, after material flows, there is no need for the load to increase and the deformation area is fixed and strains are confined to smaller plastic areas.

4.7 Extrusion load variables versus friction factor for polygonal billets and circular punches

The results of FE simulation for backward extrusion of internally circular shaped tubes from polygonal billets are depicted in figure 22 as the diagram of extrusion load versus friction factor. It could be concluded that with increasing the number of polygonal edges, the extrusion load goes up. This is due to the increase of the area of contact of the billet and the container, having in mind that the reduction of area has been kept constant; hence the friction force raises and causes the extrusion load to increase.

5. Conclusions

Finite element simulation of the backward extrusion process was carried out for shaped sections with internally elliptical, rectangular, hexagonal and externally circular shapes and also for the internally circular and externally polygonal sections.

- The 3D simulation predictions for the extrusion force for lower reduction of areas were closer to experimental observations whereas for higher reductions they were further apart.
- Some results such as the velocity distribution during different stages of the extrusion process for rectangular tubes and other results such as extrusion load, stress, and strain distribution for hexagonal tubes were also obtained which were not given in the previous work.
- The 3D velocity field distribution and the free surface configurations were presented by this simulation.
- It was concluded from the results given here that hollow sections with large reductions and thin walls are more prone to distortions due to the large differences in the components of velocity in different directions.
- A detailed analysis of strain, stress and velocity field for the backward extrusion of shaped section was presented which assisted a better understanding of the process.

6. Future Work

Backward extrusion of thin sections requires high loads and precision tooling; special designs in tooling could reduce the extrusion load considerably and eliminate the requirements for the high precision tooling. Therefore investigation into the tooling design and process design could be done by FEM simulation to evaluate the feasibility of such ideas before embarking on experimenting and manufacturing. This is currently under investigation by the authors.

7. References

- Abrinia, K. and Orangi, S. (2007). A Finite Element Simulation for the Backward Extrusion of Internally Hollow Circular Sections from Polygonal Billets, *Journal of Faculty of Engineering of the University of Tehran*, 40(6), 771-780,1026-0803.
- Abrinia, K. and Gharibi, K. (2008). An investigation into the extrusion of thin walled cans, *Int J Mater Form, Suppl 1*:411-414, 1960-6206.
- Abrinia K. and Orangi S., (2009) Investigation of Process Parameters for the Backward Extrusion of Arbitrary-Shaped Tubes from Round Billets Using Finite Element Analysis, *J Mat Eng Perform*, 18, 1201-1208, 1059-9495.
- Bae, W.B. and Yang, D.Y. (1992). An Upper-Bound Analysis of the Backward Extrusion of Internally Elliptic-Shaped Tubes from Round Billets, *J Mat Proc Tech*, 30(1-2), 13-20, 0924-0136.
- Bae, W.B. and Yang, D.Y. (1993-a). An Analysis of Backward Extrusion of Internally Circular-Shaped Tubes from Arbitrarily-Shaped Billets by the Upper-Bound Method, *J Mat Proc Tech.*, 36(2), 175-185, 0924-0136.
- Bae, W.B. and Yang, D.Y. (1993-b). An Upper-Bound Analysis of the Backward Extrusion of Tubes of Complicated Internal Shapes from Round Billets, *J Mat Proc Tech*, 1993, 36(2), 157-173, 0924-0136.
- Bakhshi-Jooybari, M., Saboori, M., Hosseinipour, S.J., Shakeri, M. and Gorji, A., (2006). Experimental and numerical study of optimum die profile in backward rod extrusion, *J Mat Proc Tech*, 177, 596-599, 0924-0136.
- Guo, Y.M., Yokouchi, Y. and K. Nakanishi, (2000). Hot Backward Extrusion Comparative Analyses by a Combined Finite Element Method, *Int J Mech Sci*, 42, 1867-1885.
- Im, Y.T., Kang, S.-H., Cheon, J.-S. and Lee, J.H. (2004). Finite element simulation of tip test with an aluminum alloy, *J Mat Proc Tech*, 157-158, 171-176, 0924-0136.
- Lee, R.-S., and Kwan, C.T. (1996). A Modified Analysis of the Backward Extrusion of Internally Circular-Shaped Tubes from Arbitrarily Shaped Billets by the Upper-Bound Elemental Technique, *J Mat Proc Tech*, 59, 351-358, 0924-0136.
- Lin, Y.T. and Wang, J.P. (1997). A New Upper-Bound Elemental Technique Approach, *J Comput Struct*, 65(4), 601-611, 0045-7949.
- Long, H. (2006), Quantitative evaluation of dimensional errors of formed components in cold backward cup extrusion, *J Mat Proc Tech*, 177, 591-595, 0924-0136.
- Moshksar, M.M. and Ebrahimi, R. (1998) An Analytical Approach for Backward Extrusion Forging of Regular Polygonal Hollow Components, *Int J Mech Sci*, 40(12), 1247-1263, 1867-1885.
- Orangi, S. and Abrinia, K. (2005) 3D simulation of backward extrusion process for production of internally shaped circular tubes sections from arbitrarily shaped tubes, *ISME International Conference*, Isfahan University of technology, Isfahan, Iran, May 2005.
- Saboori, M., Bakhshi-Jooybari, M., Noorani-Azad, M. and Gorji, A., (2006). Experimental and numerical study of energy consumption in forward and backward rod extrusion, *J Mat Proc Tech*, 177, 612-616, 0924-0136.
- Shin, T.J., Lee, Y.H., Yeom, J.T., Chung, S.H., Hong, S.S., Shim, I.O., Park, N.K., Lee, C.S. and Hwang, S.M. (2005). Process optimal design in non-isothermal backward extrusion of a titanium alloy by the finite element method, *Comput Methods Appl Mech Engrg*, 194, 3838-3869, 0045-7825.

- Uyyuru, R.K.U. and Valberg, H., (2006). Physical and numerical analysis of the metal flow over the punch head in backward cup extrusion of aluminum, *J Mat Proc Tech*, 172, 312-318, 0924-0136.
- Wriggers, Peter (2008) Nonlinear Finite Element Method, Springer, ISBN-10: 3540710000.

Finite element analysis on v-die bending process

Sutasn Thipprakmas

*King Mongkut's University of Technology Thonburi
Thailand*

1. Introduction

Bending is a common metalworking process used in sheet-metal forming to fabricate curve-shaped products of various sizes, such as parts of automobiles, aircraft, and ships shown in Fig. 1(a), in addition to various consumer products, such as kitchenware and sanitary products, depicted in Fig. 1(b).



(a) Automobile product

(b) Consumer product

Fig. 1. Examples of bent products

The bending process is based on engineering science and has a large variety of applications. In addition, bending also features in many sheet metal-forming processes, such as the deep drawing and stamping processes sketched in Fig. 2.

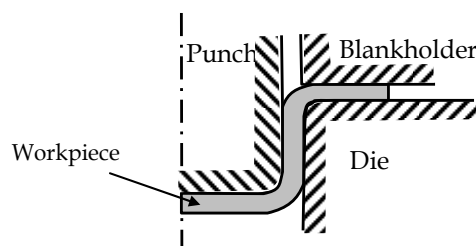


Fig. 2. Bending in deep-drawing process

According to DIN 8586, the bending process is subdivided into two groups as follows:

- (1) Bending by using a linear die motion, in which the tool moves linearly to bend the workpiece; for example, the wiping-die bending and U-die bending processes, shown in Fig. 3(a).
- (2) Bending by using a rotating die motion, in which a tool moves in rotations to bend the workpiece, as shown in Fig. 3(b). The basic advantages of rotary bending are less bending-force requirement, elimination of the use of a blankholder, and a final bending angle greater than 90 degree.

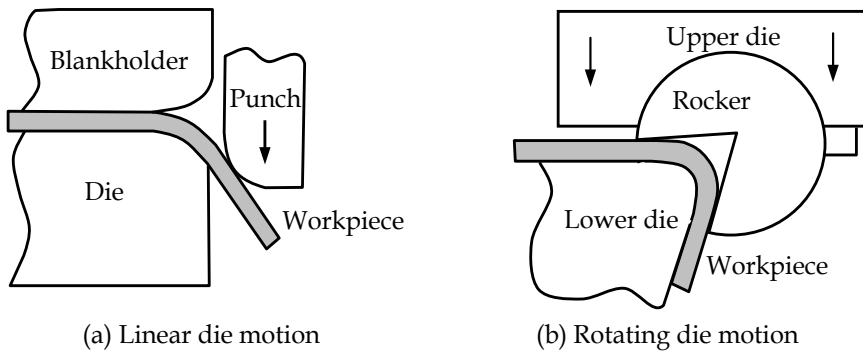


Fig. 3. Classification of bending process according to DIN 8586

Bending is a manufacturing process by which a force, corresponding to a given punch displacement, acts on the workpiece. The workpiece is initially bent in an elastic region. As the process continues, the workpiece is deformed by plastic deformation, thereby changing its shape. In the bending process, the bending load increases until the elastic limit of the material is exceeded. The material state enters the plastic deformation region and sheet metal can be formed. Specifically, the stress generated in the workpiece is greater than the yield strength but lower than the ultimate tensile strength of the material. The workpiece initially deforms where the bending moment is the greatest. For example, in the case of the V-die bending process, the process of permanent deformation starts directly underneath the punch. During the bending operation, the outer surface of the workpiece generates the greatest stretch, which then propagates inward toward the neutral plane. Similarly, the inner surface of the workpiece generates the greatest compression, which also then propagates inward toward the neutral plane. These distributions of the stresses are represented in Fig. 4.

This bending process results in formation of both a crack that usually occurs on the outer surface and a wrinkle that usually occurs on the inner surface. In addition to these crack and wrinkle defects, the major problem in the bending process is the accuracy of the part, including the part's geometry and dimension. These problems are directly influenced by the parameters of the bending process, such as properties and thickness of the material, angle and radius of bending, and the bending stroke (Schuler, 1998), (Lange, 1985). These problems were investigated in many earlier researches, where the finite element method (FEM) and related experimental protocols were conducted. The sequence of sheet-metal bending was optimized using a robot (Shigeru & Atsushi, 2002).

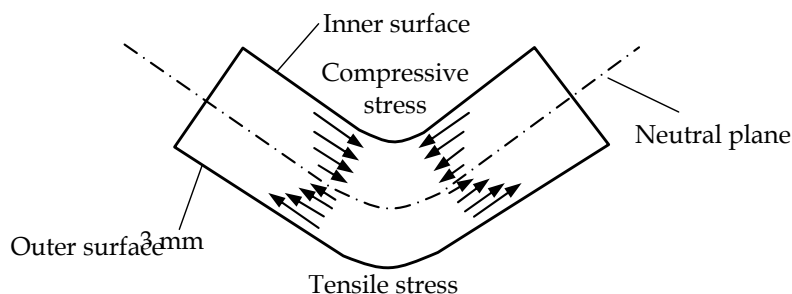


Fig. 4. Stress distribution in the V-die bending process

The plastic collapse of aluminum extrusions in biaxial bending was investigated (Belingardi & Peroni, 2007). The spring-back deformation during bending of a tube was studied (Da-xin et al., 2009). The deformation behavior of a thin-walled tube in rotary-draw bending under push assistant-loading conditions was investigated (Li et al., 2010). The modeling of the inelastic bending of a metal sheet by applying thermal coupling was studied (Woznica & Klepaczko, 2003). The bending moment and spring-back was studied in the pure bending of anisotropic sheets (Alexandrov & Hwang, 2009). The process of sheet-metal bending was investigated using computer-aided techniques (Duflou et al., 2005). A thermomechanical model was identified for air bending (Canteli et al., 2008). The influence of the coining force on spring-back reduction in a V-die bending process was investigated (Leu & Hsieh, 2008). The spring-back in wiping-die bending processes was examined by applying both an experimental approach and the response surface methodology (Mkaddem & Saidane, 2007). The air bending and spring-back of stainless steel-clad aluminum sheets was studied (Yilamu, et al., 2010). The bending of thin sheets of stainless steel was studied by a raster-scanned low-power CO₂ laser (Vsquez-Ojeda & Ramos-Grez, 2009). The modeling of spring-back, strain rate, and Bauschinger effects for the two-dimensional steady-state cyclic flow of sheet metal subjected to bending under tension was investigated (Sanchez, 2009).

Most of the previous research works studied the spring-back phenomenon, including the effects of process parameters on it. Therefore, in this chapter, the spring-go phenomenon has been investigated thoroughly and will be discussed in depth. It will also serve as the basis for theoretically explaining the spring-go phenomenon that occurs during the bending process in comparison to the spring-back phenomenon. In addition, the effects of process parameters, including the radius and height of the punch, have been investigated and will be clearly discussed. Next, the spring-back and spring-go phenomena, including the effects of the punch radius and punch height, were investigated using the FEM; the FEM simulation results were also validated by experiments. On the basis of material-flow and stress-distribution analyses, the FEM simulation results showed that both spring-back and spring-go phenomena, in addition to the effects of punch radius and punch height, could be theoretically elucidated. In addition, the FEM simulation results also showed the possibility of using the FEM to predict spring-back and spring-go characteristics.

2. Principle of V-die Bending Process

The V-die bending process is the bending of a V-shaped part in a single die. The principle of the V-die bending process is shown in Fig. 5. The workpiece is bent between a V-shaped punch and die. The force acting on the punch causes punch displacement and then the workpiece is bent. The workpiece is initially bent as an elastic deformation. With continued downward motion by the punch, plastic deformation sets in when the stresses exceed the elastic limit. This plastic deformation starts on the outer and inner surfaces directly underneath the punch. The greatest tensile stress is generated on the outer surface, whereas the greatest compressive stress is generated on the inner surface; these stresses decreasingly propagate inward toward the workpiece. Therefore, crack formation usually occurs on the outer surface and a wrinkle usually occurs on the inner surface. The initial bending stage—the so-called “Air bending”—starts the moment the punch establishes contact with the workpiece and is completed either when the legs of the workpiece become tangential to the faces of the die or when the smallest internal radius of the workpiece becomes smaller than the radius of the punch. As the process continues, after completion of air bending, the bending is focused on the three points of the punch and the two faces of the die. The contact points between the workpiece and die are shifted toward the centerline of the die, and the legs of the workpiece try to close around the punch. As the punch proceeds further, the legs of the workpiece establish contact with the punch, and it is pressed to open up again until the bend angle approaches the die angle (Schuler, 1998), (Lange, 1985). The clearance between the punch and the die in V-die bending is commonly dependent of the workpiece thickness. The usual thickness of the workpiece in the V-die bending process ranges from approximately 0.5 to 25 mm.

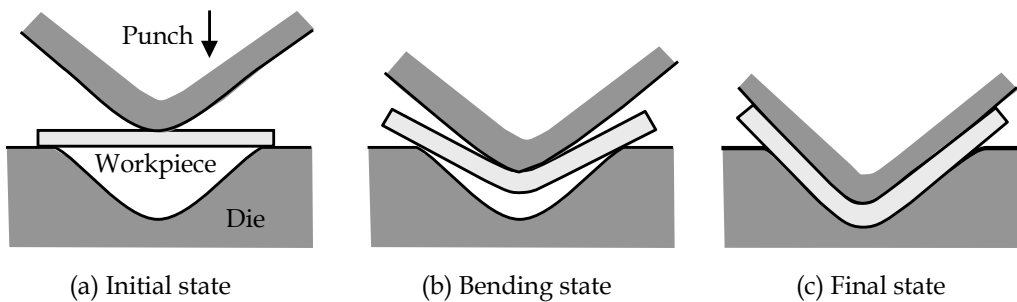


Fig. 5. Principle of the V-die Bending Process

3. Classification of V-die Bending Process

The basic advantages of the V-die bending process are as follows:

- (1) a simple tool design,
- (2) an economical setup time, and
- (3) an enormous range of sizes and complex shapes that can be fabricated for the part.

Therefore, the V-die bending process is generally used in both the Press Brake machine and the Press machine associated with press-form tooling. However, the basic disadvantage of

the use of the Press Brake machine is low productivity. The Press Brake machine is suitable for production of only a small quantity or the prototype of a product.

In addition to the classification of the V-die bending process based on the Press machine, the V-die bending process can be classified into two categories, namely, partial and full V-bending, as depicted in Fig. 6. Partial V-die bending (Fig. 6a) uses a V-bending punch containing a punch smaller than the workpiece; thus, it cannot establish contact with the whole surface of the workpiece on the die side. In Fig. 6(b), a full V-die bending is shown, where the workpiece completely establishes contact with the surfaces on both sides of the punch and die.

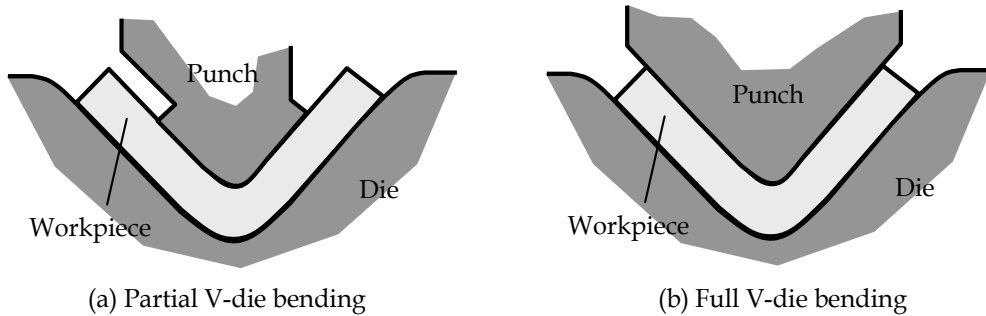


Fig. 6. Classification of the V-die Bending Process

4. Problems in V-die bending process

Considering the accuracy of the produced part, the product in the sheet metal-forming process is set higher requirements consisting of precise dimensions. These consist of fundamentally the same requirements as for the V-die bending process. Unwanted deformation generally occurs in the bending zone (underneath the punch), especially, the thick part of the workpiece and sharp bends, which results in a decrease in the accuracy of the part, as sketched in Fig. 7.

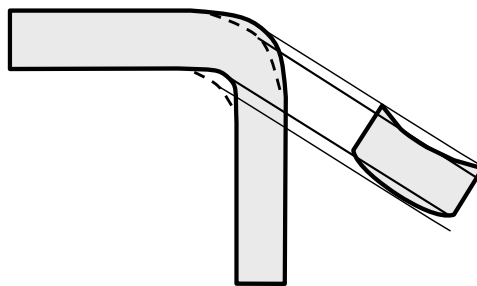


Fig. 7. Unwanted Deformation of the Cross Section during Bending

During the bending process, sufficient material flow to replace the material displaced is ensured to prevent the weakening and fracturing taking place at the workpiece corners (Schuler, 1998), (Lange, 1985). In addition, achieving the desired bend angle is very difficult and is the major problem in the bending process, especially in mass production. The reason is the occurrence of spring-back or spring-go feature. This problem is the key factor that

decreases the quality of the bent part. In addition, in practice, applying control over the spring-back or spring-go feature is very difficult because this feature is dependent on many process parameters, such as the material property, material thickness, and the angle, radius, and stroke of bending (Schuler, 1998), (Lange, 1985). Many earlier researches have investigated the spring-back and spring-go features to resolve these two in the bending process that is applied in many industrial problems by using the FEM and related experiments. The spring-back of CK67 steel sheet in V-die and U-die bending processes was investigated (Bakhshi-Jooybari et al., 2009). The spring-back control of sheet-metal air-bending process was studied (Wang et al., 2008). The variation of elastic modulus during plastic deformation and its influence on spring-back was investigated (Yu, 2009). The spring-back in the wipe-bending process of sheet metal was investigated using a neural network (Kazan et al., 2009). The spring-back of stainless steel sheet metal in V-bending dies was determined (Tekaslan et al., 2008). The spring-back prediction, compensation, and optimization were investigated using numerical analysis (Meinders et al., 2008). The spring-back characteristics of aluminum sheet-metal alloys under warm forming conditions were investigated using the numerical method (Kim & Koc, 2008). The spring-back and spring-go phenomena were investigated using the FEM (Thipprakmas & Rojananan, 2008). The spring-back in the bending of aluminum sheets was investigated using the FEM (East et al., 2002). The impact of physical parameters on spring-back appearance in U-draw bending was investigated by using FEM (Papeleux et al., 2002). The process parameters, such as die clearance, die radius, and step height, and their effects on spring-back were studied using the FEM (Ling et al., 2005). The amount of spring-back in steel sheet-metal of 0.5-mm thickness in a bending die was investigated (Tekaslan et al., 2006). The spring-back in sheet-metal bending was investigated using FEM (Panthi et al., 2007). The amount of spring-back on the spring material was investigated using FEM (Jin et al., 2000). The effects of punch height in partial V-bending process were investigated (Thipprakmas, 2010).

As mentioned above, the principle of the V-die bending is not very difficult to understand; however, a great degree of expertise and skills are required as this is an important technique for increasing the productivity and for improving product quality. Therefore, the characteristic features of bending must be theoretically understood. In addition, the effects of the process parameters related to the geometry and properties of the material must be strictly considered to achieve high-quality bent parts. To fully understand this process, the characteristic features of bending and the effects of the process parameters will be completely dealt with through theoretical investigation and explanation in this chapter.

5. Mechanism of Spring-back Phenomenon

According to the plastic deformation theory, the material is generally divided into two zones: the elastic and the plastic zones. The elastic property tries to maintain the material in the initial shape, whereas the plastic property tries to retain the material in the deformed shape. In the sheet metal-bending process, the bending load increases until the elastic limit of the material is exceeded and then the material state enters the plastic deformation zone. The outer surface of the material generates the tensile stress, which propagates inward toward the neutral plane. Vice versa, the inner surface of the material generates the compressive stress and it propagates inward toward the neutral plane. Because of the stress distributions, this phenomenon causes the formation of a small elastic band around the

neutral plane, as shown in Fig. 8. As the bending force is removed at the end of the bending stroke, the inner surface-generated compressive stress tries to enlarge the workpiece and the outer surface-generated tensile stress tries to shrink. In contrast, the elastic band remains in the bent parts trying to maintain its original shape, resulting in a partial recovery toward its initial shape. This elastic recovery is called “spring-back”. Thus, the workpiece tries to spring back and the bent part slightly opens out, as shown in Fig. 9.

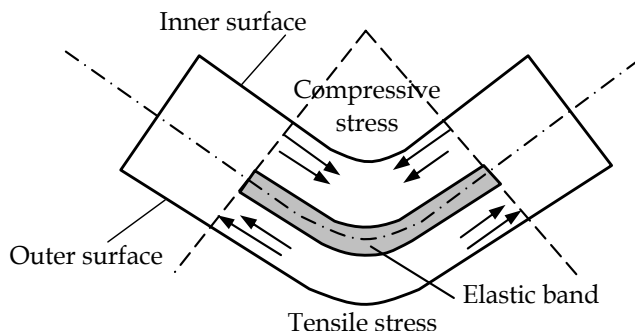


Fig. 8. Illustration of the Elastic Band in the Bent Part

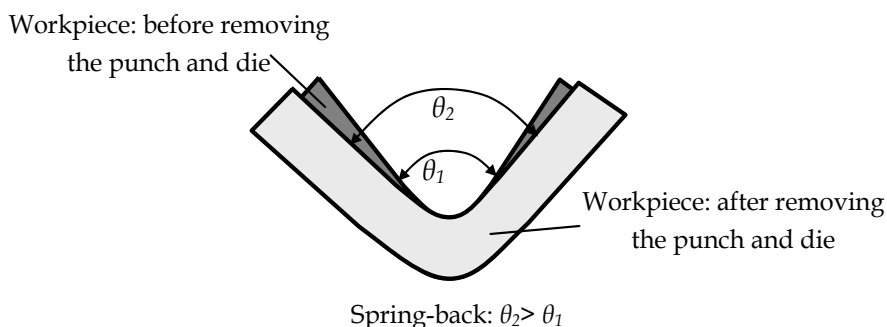


Fig. 9. Illustration of Spring-back Feature

6. Mechanism of Spring-go Phenomenon

In addition to the spring-back feature, the spring-go feature usually occurs in a bending process. The spring-back characteristic occurs as the bent part slightly opens. In contrast, the spring-go characteristic occurs as the bent part slightly closes, as shown in Fig. 10.

The concept of the spring-go phenomenon has rarely been theoretically clarified in previous researches. However, the spring-go phenomenon was investigated in the author's previous researches. FEM simulation was used as a tool for investigating and understanding the theoretical basis of the spring-go phenomenon. The FEM-simulation model used in the previous study by the current author is shown in Fig. 11.

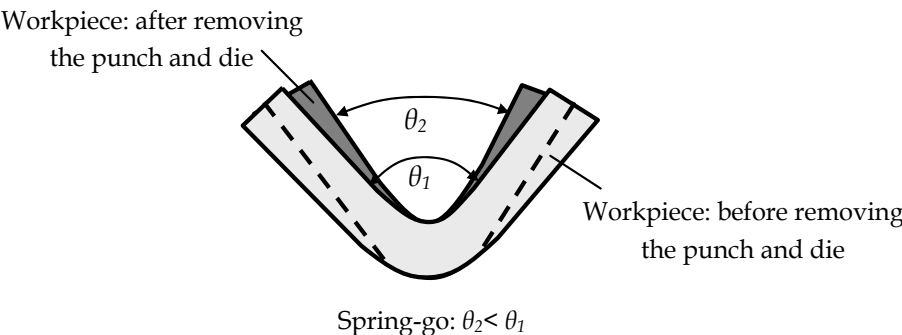


Fig. 10. Illustration of Spring-go Feature

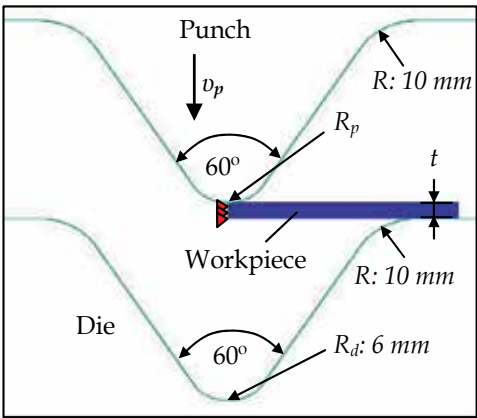


Fig. 11. FEM-Simulation Model

Simulation model	Plane strain model
Object type	Workpiece : Elasto-plastic Punch/Die : Rigid
Bended material	A1100-O ($\sigma_B=92.5\text{MPa}$, $\lambda=46\%$) ($t= 3\text{mm}$, length 70 mm)
Friction coefficient (μ)	0.1 $\bar{\sigma} = 160.4\bar{\epsilon}^{0.22}$
Flow curve equation	
Die radius (R_d)	6 mm
Angular punch radii (R_p)	1, 2, 3, 4, 5, 6 mm
Punch height (H_p)	6, 16, 26, 36 mm
Bending angle (θ)	60°
Punch velocity (v_p)	0.05 mm/sec

Table. 1. FEM-Simulation Conditions and Material Property

The two-dimensional plane strain with a simulation model of 35-mm length and 3-mm thickness was investigated. The commercial analytical code (DEFORM-2D) for a two-dimensional implicit quasi static FEM was used as the FEM-simulation tool. The workpiece material was aluminum A1100-O (JIS), which was set as an elasto-plastic type, with the rectangular meshes consisting of approximately 3500 elements. The punch and die were set as rigid type. Table 1 shows the details of the FEM-simulation conditions and the material properties. Similar to the results of previous research, on the basis of the material-flow analysis and stress distribution, the spring-go phenomenon has been investigated thoroughly and will be discussed in depth compared with the spring-back phenomenon.

6.1 Material-Flow Analysis

Referring to the author's previous research (Thipprakmas & Rojananan, 2008), Fig. 12 shows the FEM simulation analysis of the material flow feature in the workpiece with reference to the punch radius. The FEM simulation results showed the same feature of material flow during a small bending stroke, as shown in Fig. 12(a-1), (b-1), and (c-1). Specifically, the material was allowed to flow into the die along the punch direction. As the punch proceeded further, before the workpiece contacted the sides of the punch, as the punch radius decreased, both the velocity of the leg of the workpiece (zone A) and the gaps at the zones B and C increased, as shown in Fig. 12(a-2), (b-2), and (c-2).

After the workpiece contacted the side of the punch, as shown in Fig. 12(a-3), (b-3), and (c-3), the workpiece was pushed backward toward the die, resulting in an S-curve-shaped material-flow characteristic; therefore, a reverse-bending feature was generated. It was observed that the punch could proceed further as the punch radius decreased. Before removing the punch, in the case of small punch radii as shown in Fig. 12(a-4), the punch pressed the workpiece so that the latter was in contact with the die corner, whereas the leg of the workpiece had not yet completely contacted the side of the die. Therefore, an S-curve-shaped material-flow feature was generated. In contrast, the punch radius increased as shown in Fig. 12(b-4) and (c-4), the leg of the workpiece was already in contact with the die's side and therefore, the punch cannot push the workpiece into the die corner, resulting in a gap between the workpiece and the die (zone D). Therefore, the S-curved feature of the material flow was reduced; the S-curved material flow decreased as the punch radius was increased. Importantly, this S-curved material flow affected the stress distribution, resulting in the reversed bending characteristic; therefore, the required bending angle could not be achieved.

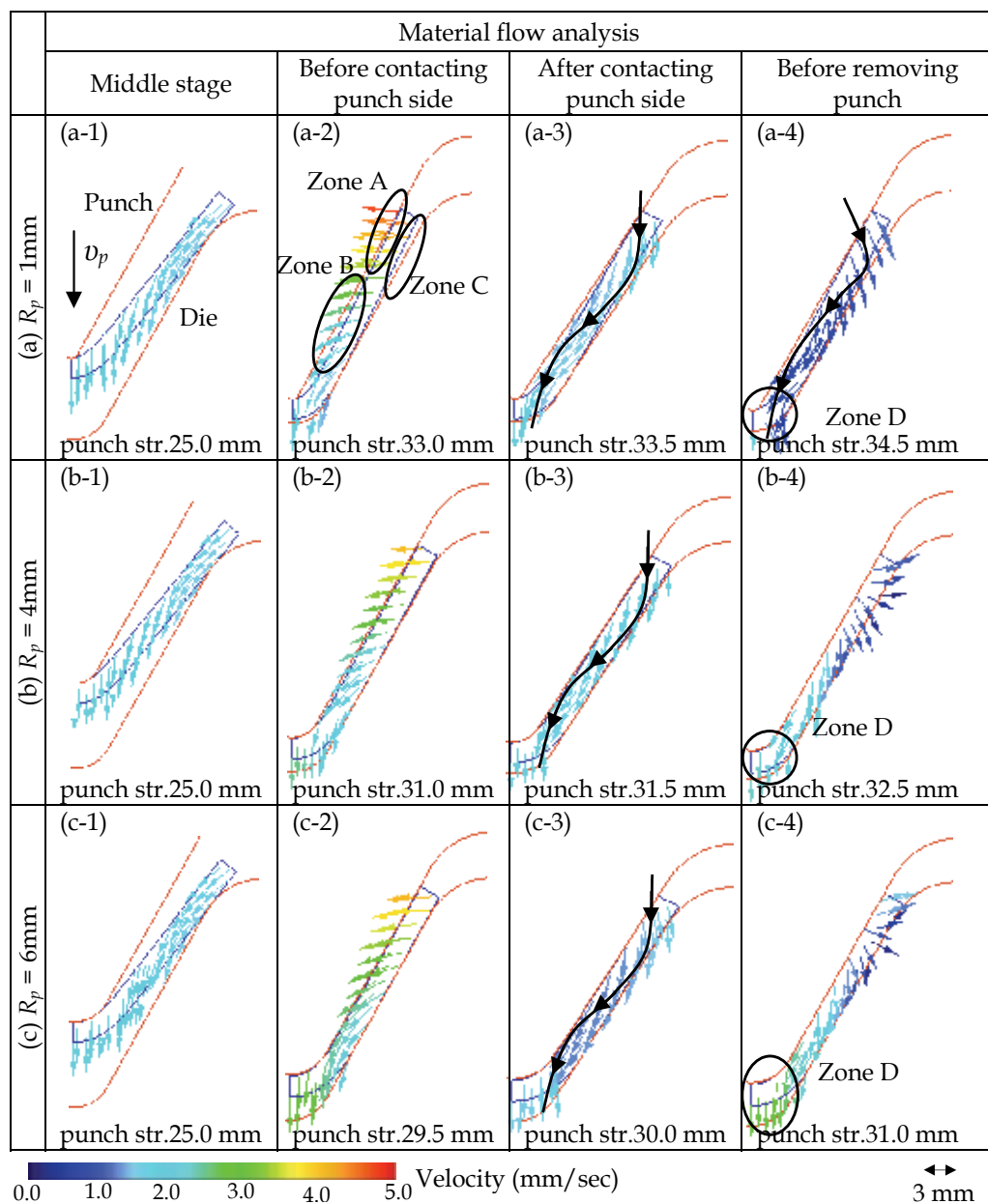


Fig. 12. Material-Flow Analysis ($A1100\text{-O}$, $t = 3\text{ mm}$, $\theta = 60^\circ$, $R_d = 6\text{ mm}$) (Thipprakmas & Rojananan, 2008)

6.2 Stress-Distribution Analysis

As mentioned in the material-flow analysis section, the material flow feature affected the stress-distribution analysis. This was also investigated and confirmed in the author's previous research (Thipprakmas & Rojananan, 2008). The FEM simulation results of the

stress distribution with reference to the punch radius are shown in Fig. 13. The stress distribution was similarly analyzed in the workpiece before the workpiece contacted the punch side, as shown in Fig. 13(a-1), (a-2), (b-1), (b-2), (c-1), and (c-2). The FEM simulation results agreed well with the bending theory (Schuler, 1998), (Lange, 1985). Specifically, tensile stress was generated on the outer surface of the workpiece, and it decreasingly propagated inward toward the neutral plane.

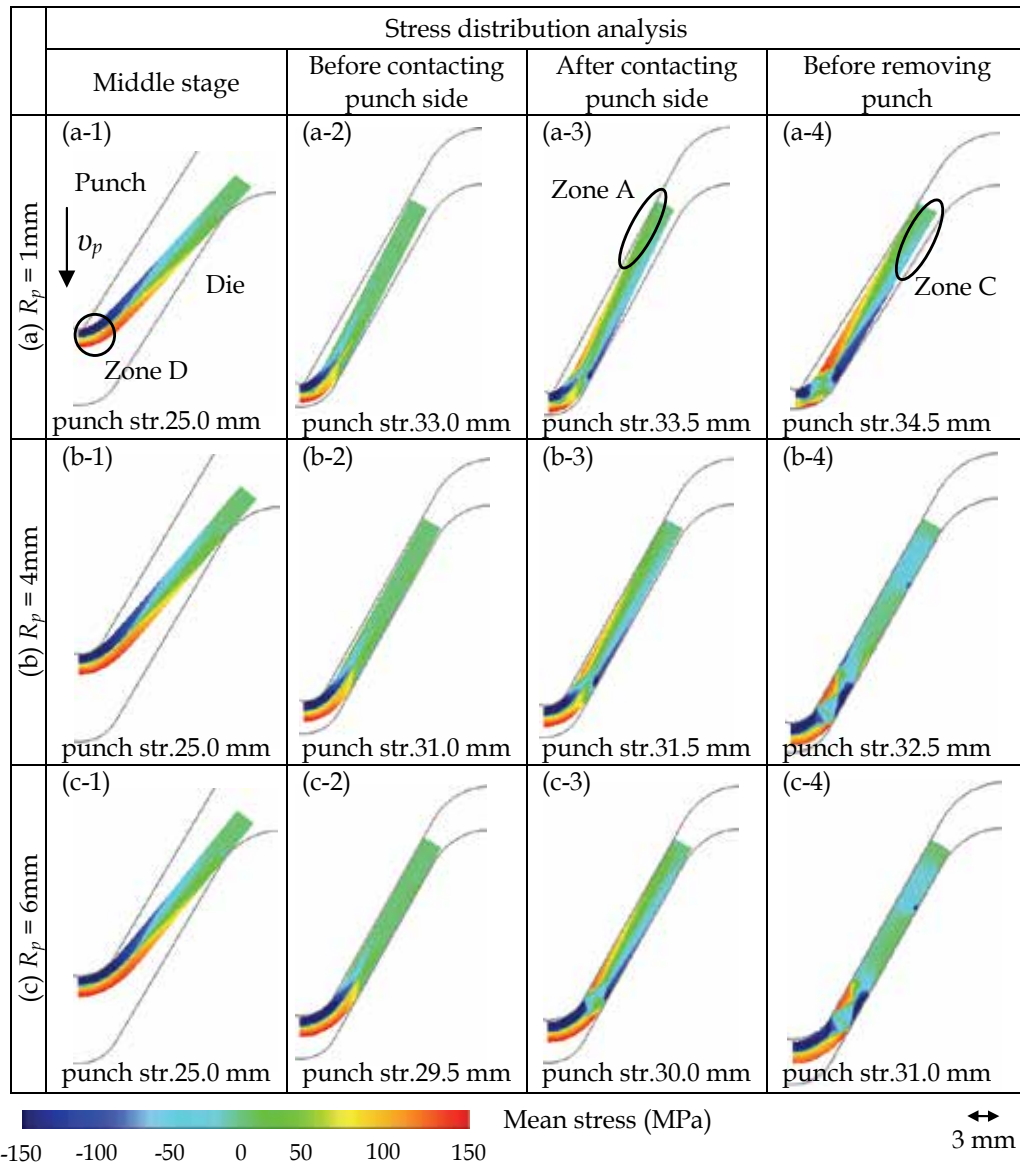


Fig. 13. Stress-Distribution Analysis (A1100-O, $t = 3\text{ mm}$, $\theta = 60^\circ$, $R_d = 6\text{ mm}$) (Thipprakmas & Rojananan, 2008)

Similarly, the compressive stress generated on the inner surface of the material decreasingly propagated inward toward the neutral plane. In addition, the maximum tensile and compressive stresses were generated at the zone of the bending radius, which decreasingly propagated along the leg of the workpiece. However, the bend-allowance zone increased as the punch radius increased; therefore, the analyzed stress band along the leg of the workpiece was increased as the punch radius increased. At the stage of the leg of workpiece contacting the punch side (Fig. 13(a-3), (b-3), and (c-3)), the material was pushed backward to the side of the die, resulting in an S-curve-shaped material-flow feature, in which the reverse-bending feature was generated. Specifically, tensile stress was generated on the punch side, whereas the compressive stress was generated on the die side, both of which were of reverse orientation compared with the stress distribution in the bending-radius zone. This reverse-bending feature increased as the bending stroke increased, until the end of the bending stroke. Before removing the punch (Fig. 13(a-4), (b-4), and (c-4)), the reverse-bending feature was strongly retained in the leg of the workpiece because the punch could not completely push this leg to contact the die side. This reverse bending was decreased when the punch compressed on the workpiece in the cases where the leg of the workpiece was pushed backward to contact the die side. The FEM simulation results of the stress distribution corresponded with the bending theory, and they also illustrated the correspondence with the material-flow features. The S-curve-shaped material flow resulted in the reverse bending of the leg of the workpiece. This reverse-bending feature affected the required bending angle.

6.3 Spring-back and Spring-go Characteristics

As mentioned above, due to the material-flow and stress-distribution features, the required bending angle was difficult to achieve. The stress distribution at the bend-allowance zone tried to slightly open the leg of the workpiece to the initial shape. In contrast, in the reverse-bending zone, where the tensile and compressive stresses were generated on the punch and die side, respectively, the leg of the workpiece also tried to open by the effect of the spring-back action. However, in this zone, the leg of the workpiece was reversely bent with reference to the bend-allowance zone; therefore, the leg of the workpiece tried to spring-back by slightly opening on the punch side. To predict the occurrence of spring-back or spring-go, the compensation of stress on the bend-allowance and reverse-bending zones was strictly considered. After compensating the whole stress on the workpiece, if the leg of the workpiece tried to slightly open to the punch side, it would be called "Spring-go". This spring-go feature caused the bending angle to be smaller than the required bending angle. In contrast, if the leg of the workpiece tried to slightly open to the die side, it would be called "Spring-back". This spring-back feature caused the bending angle to be larger than the required bending angle. This explanation was theoretically clearly described in the author's previous research (Thipprakmas & Rojananan, 2008), and the results are shown in Fig. 14.

7. Effects of Process Parameters

In all processes, the process parameters are the most important factors that control the quality of the products and the production cost. Selection of optimized process parameters could fabricate high-quality products, in addition to reducing the production cost and time.

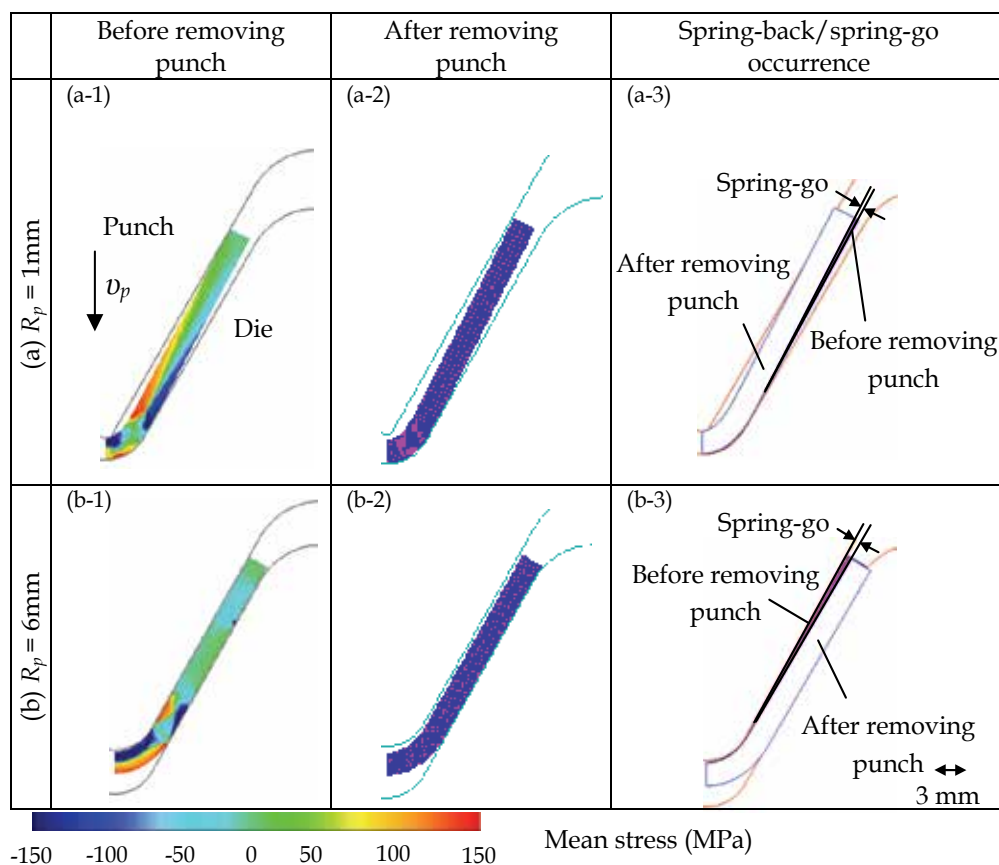


Fig. 14. Characteristics of Spring-back and Spring-go Mechanisms (A1100-O, $t = 3\text{ mm}$, $\theta = 60^\circ$, $R_d = 6\text{ mm}$) (Thipprakmas & Rojananan, 2008)

To prevent the occurrence of spring-back or spring-go and achieve the required bending angle in the V-die bending process, the effects of process parameters must be clearly understood. Many previous researchers studied the effects of process parameters, but most of them were aimed at the spring-back problem. Therefore, in this section, the effects of process parameters on the spring-back and spring-go effects that were investigated in the author's previous researches (Thipprakmas & Rojananan, 2008), (Thipprakmas, 2010), are clearly described. Among the investigated process parameters, the ones that mainly affect the required bending angle are the radius and height of the punch.

7.1 Punch Radius

The effects of punch radius were investigated in many previous researches and increasing the punch radius resulted in the increase of spring-back. However, in this section, the theoretical explanation of this effect will be discussed. In addition to the spring-back feature, the effects of the punch radius on the spring-go feature were also investigated and will be discussed in depth. Based on the author's previous research (Thipprakmas & Rojananan,

2008), the punch radius ranging from 1 to 6 mm was investigated. The FEM-simulation results for the spring-back and spring-go values are shown in Fig. 15. The FEM simulation results showed that the amount of spring-go decreased as the punch radius increased, whereas the amount of spring-back increased as the punch radius increased. A punch radius of 3 mm gave spring-back and spring-go values of zero; that is, the required bending angle of 60° could be achieved. To further elucidate clearly, a comparative stress-distribution analysis is shown in Fig. 16. As the punch radius increased, the bending stroke decreased because the punch pushed the leg of the workpiece backward to contact the side of the die. This process resulted in the decreasing of the reverse-bending zone. In addition, the bending zone at the punch radius increased as the punch radius increased, resulting in an increasing of the stress-distribution band and thus the amount of spring-back was increased.

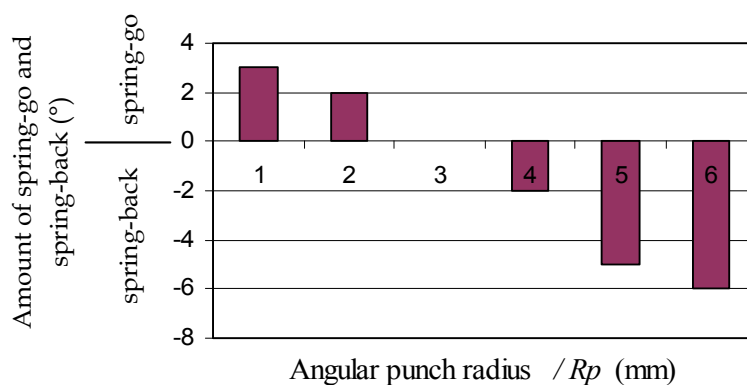


Fig. 15. The Values of Spring-back and Spring-go (A1100-O, $t = 3\text{mm}$, $\theta = 60^\circ$, $R_d = 6\text{mm}$) (Thipprakmas & Rojananan, 2008)

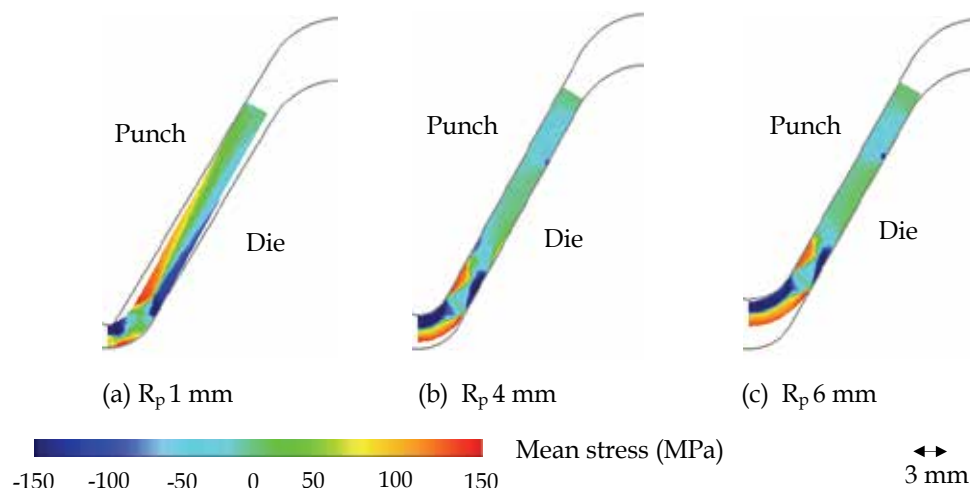


Fig. 16. Stress-Distribution Analysis with Reference to the Punch Radius (A1100-O, $t = 3\text{mm}$, $\theta = 60^\circ$, $R_d = 6\text{mm}$) (Thipprakmas & Rojananan, 2008)

7.2 Punch Height

According to the classification of the V-die bending process mentioned above, there are two categories, as shown in Fig 6; (1) full V-die bending, and (2) partial V-die bending. The partial V-die bending uses a V-shaped punch that is smaller than the workpiece; thus, the punch cannot push the whole leg of the workpiece backward. Therefore, the punch height greatly affects the required bending angle.

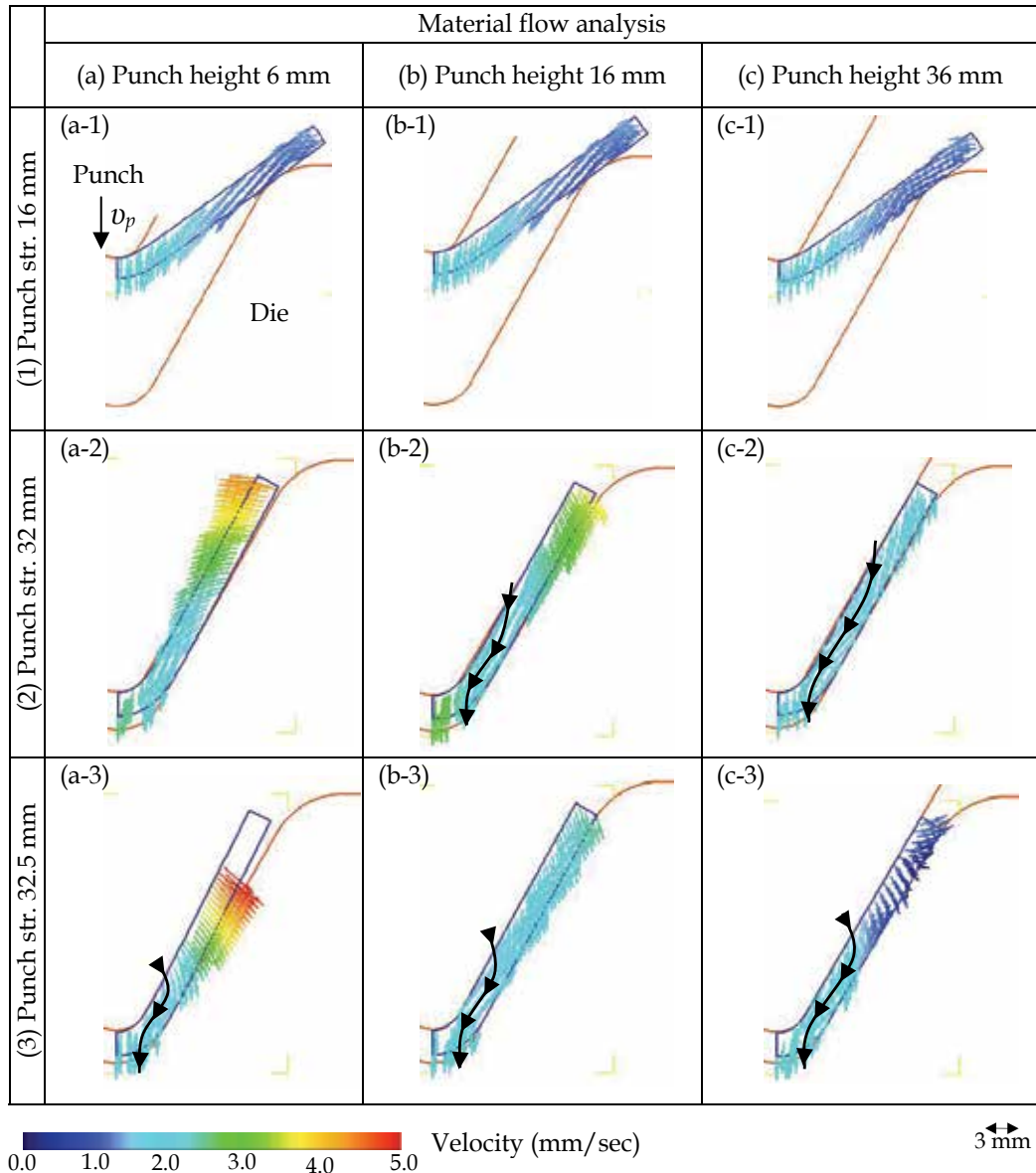


Fig. 17. Material Flow with Reference to the Punch Height (A1100-O, $t = 3$ mm, $\theta = 60^\circ$, $R_p = 4$ mm, $R_d = 6$ mm) (Thipprakmas, 2010)

This effect was investigated in the author's previous research using FEM simulation and additional experiments (Thipprakmas, 2010). Fig. 17 shows the material flow in the workpiece with reference to different punch heights. Before the punch contacts the leg of the workpiece, the manner of material flow was illustrated as shown in Fig. 17(a-1), (b-1), and (c-1). With a bending stroke of approximately 32 mm, the leg of the workpiece had not yet contacted the die in the case of a punch height of 6 mm, as shown in Fig. 17(a-2). In contrast, it was pushed backward to the die side in the cases of punch heights of 16 and 36 mm, resulting in an S-curve-shaped material flow, as shown in Fig. 17(b-2) and (c-2). It was also observed that the S-curve-shaped material flow increased as the punch height increased. At the end of the bending stroke, the small height of the punch could not push the whole leg of the workpiece to contact the side of the die, resulting in a smaller bending angle. This gap between the leg of the workpiece and the die's side decreased as the punch height increased. The S-curve-shaped material flow resulted in the reverse-bending zone, with a reversed stress distribution on the leg of the workpiece.

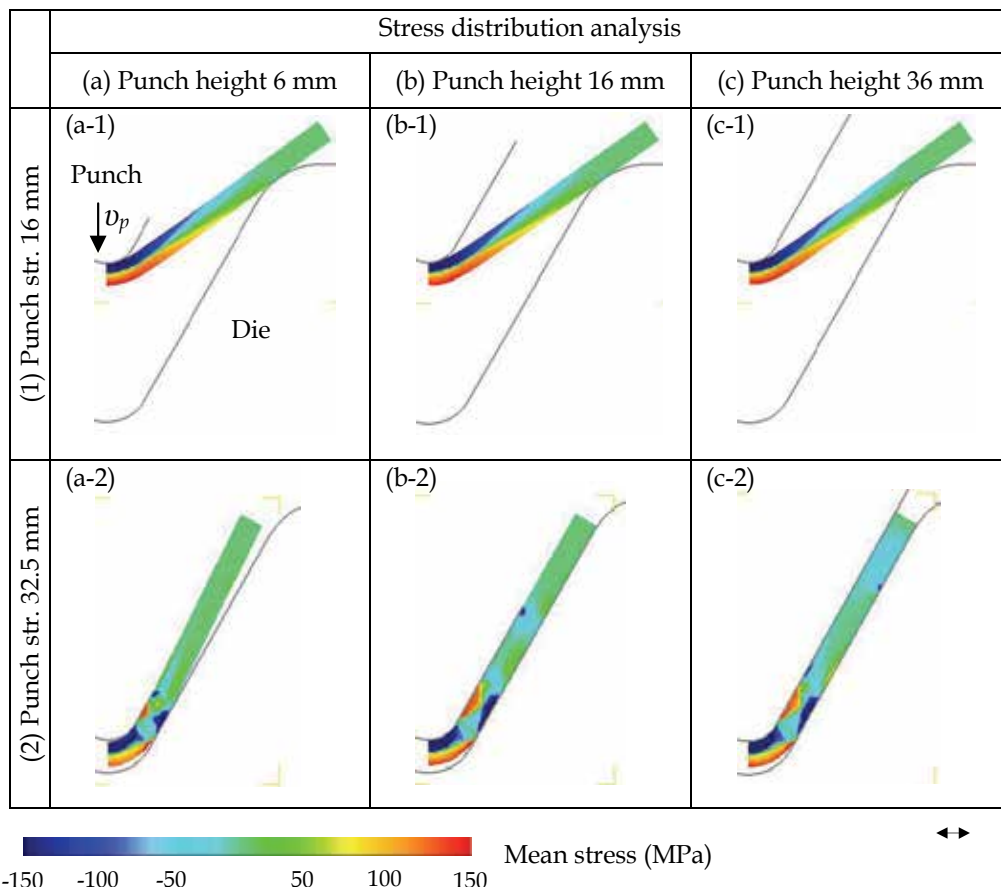


Fig. 18. Stress-Distribution Analysis with Reference to the Punch Height (A1100-O, $t = 3$ mm, $\theta = 60^\circ$, $R_p = 4$ mm, $R_d = 6$ mm) (Thipprakmas, 2010)

The stress-distribution analysis with reference to the punch heights is shown in Fig. 18. Before the leg of the workpiece contacted the punch side, as shown in Fig. 18(a-1), (b-1), and (c-1), an identical manner of stress distribution was determined, which corresponded well with the bending theory. At the end of the bending stroke, Fig. 18(a-2), (b-2), and (c-2), reverse bending was generated due to the S-curve-shaped of the material flow. This reverse-bending zone was increased as the punch height increased, and it affected the required bending angle. Therefore, to achieve the required bending angle, a balance between (i) compensating the gap between the leg of the workpiece and the die side and (ii) the stress distribution on the reverse-bending zone and bend-allowance zone was strictly considered. Referring to the author's previous research (Thipprakmas, 2010), Fig. 19 shows the comparison of the bending angle with reference to the punch heights. The FEM-simulation results illustrated that the suitable punch heights to achieve the required bending angle were approximately 21 and 30 mm. The spring-go effect occurred in the cases where the punch height was smaller than 21 mm and larger than 30 mm. In contrast, the spring-back effect was observed in the case of punch heights ranging between 21 and 30 mm. As clarified above, a very small punch height caused a large gap between the leg of the workpiece and the die side, resulting in the over bending of the part. Although a small reverse bending was generated, after compensating, the bending angle obtained was smaller than the required bending angle. Vice versa, a very large punch height caused a large reverse-bending zone, which overcame the spring-back in the bend-allowance zone; therefore, the spring-go effect was generated. However, for the punch height range 21–30 mm, the reverse bending could not suppress the spring-back in the bend-allowance zone; therefore, the spring-back effect was generated.

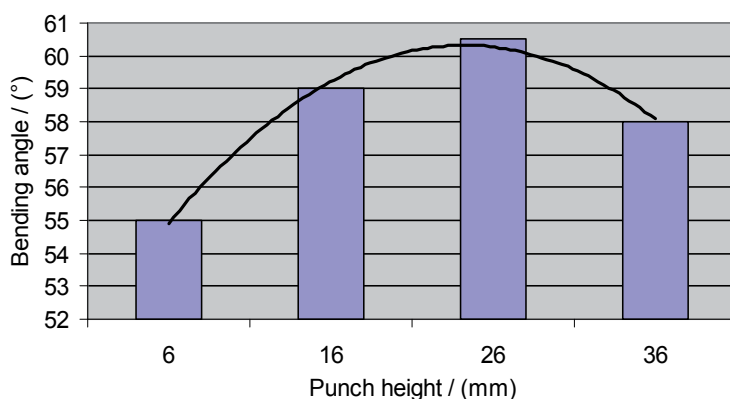
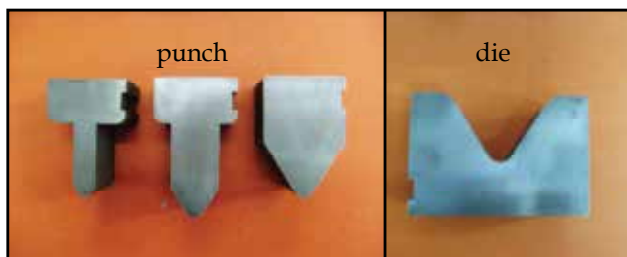


Fig. 19. Bending Angle Analyzed by FEM with Reference to the Punch Height (A1100-O, $t=3$ mm, $\theta=60^\circ$, $R_p=4$ mm, $R_d=6$ mm) (Thipprakmas, 2010)

8. Validation of FEM simulation results

Laboratory experiments were carried out to validate the FEM-simulation results. Referring to the author's previous research (Thipprakmas, 2010), the experiments were carried out on an aluminum A1100-O (JIS) workpiece with the dimensions of 70-mm length, 30-mm width, and 3-mm thickness. The 5-ton universal tensile-testing machine (Lloyd Co. Ltd.) was used as the press machine. In Fig. 20, the V-bending die is shown, in addition to the tensile-

testing machine with the V-bending die set. The bending force was recorded to compare with the bending force obtained by FEM. The bending angle was measured using the profile projector. Comparing the results, the bending force and the bending angle obtained by FEM showed a good agreement with those obtained by the experiments. These results indicated that FEM simulation could be used as a tool for theoretically clarifying the mechanisms of spring-back and spring-go, in addition to studying the effects of process parameters on these effects. Furthermore, FEM can be used to predict the bending angle.



(a) V-bending punch and die



(b) V-bending die set assembled on tensile test machine

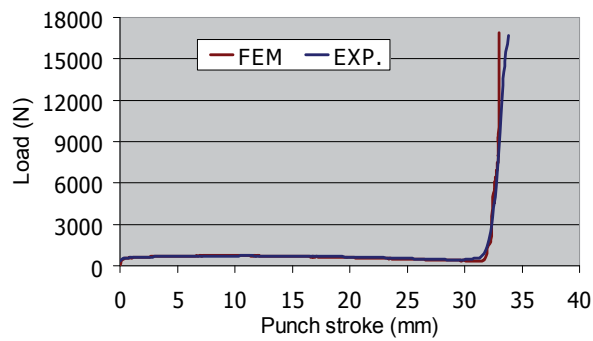
Fig. 20. The V-bending die and the Tensile-testing Machine (Thipprakmas, 2010)

8.1 Bending Force

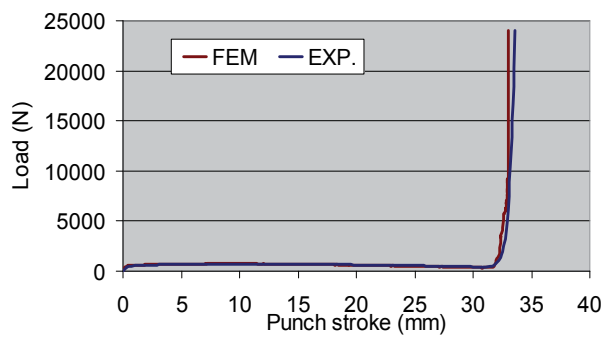
As shown in Fig. 21, the bending forces obtained by FEM showed good agreement with those obtained by the experiments. They were somewhat constant and rapidly increased near the end of the bending stroke. This bending force corresponded well with the bending theory. However, as the punch height increased, the increase in the reverse-bending feature, resulted in an increase of the bending force.

8.2 Bending angle

The comparison of the bending angles obtained by experiments and from FEM analysis is shown in Fig. 22. The FEM-simulation results showed good agreement with the experimental results, with an error of approximately 1%.



(a) Punch Height of 16 mm



(b) Punch Height of 36 mm

Fig. 21. Comparison of the Bending Force between FEM Simulation and Experimental Results (A1100-O, $t=3$ mm, $\theta=60^\circ$, $R_p=4$ mm, $R_d=6$ mm) (Thipprakmas, 2010)

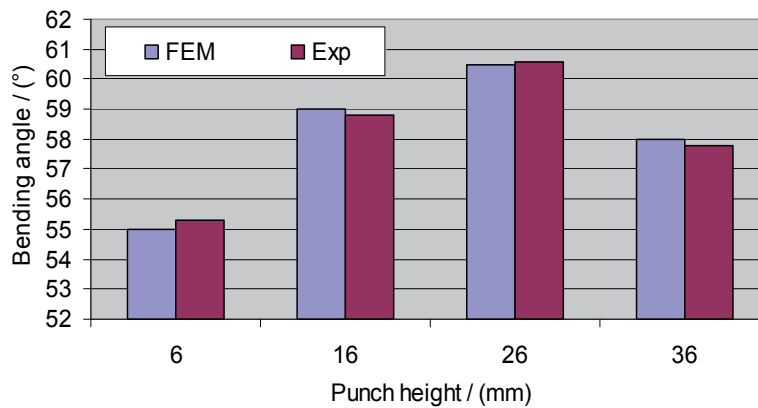


Fig. 22. Comparison of Bending Angles Obtained by FEM Simulation and Experiments (A1100-O, $t=3$ mm, $\theta=60^\circ$, $R_p=4$ mm, $R_d=6$ mm) (Thipprakmas, 2010)

9. Conclusions

FEM simulation was applied to investigate the spring-back and spring-go phenomena in a V-die bending process and to investigate the effects of process parameters, including radius and height of the punch. The FEM-simulation results were validated by laboratory experiments. The FEM-simulation results showed that the spring-back and spring-go phenomena could be theoretically elucidated based on material-flow - and stress-distribution analyses. The generation of an S-curve-shaped material-flow feature caused the reversal of stress distribution on the leg of the workpiece. This reverse-stress distribution that the tensile and compressive stresses generated on the punch and die sides, respectively, resulted in a reverse-bending zone on the leg of the workpiece, and the leg of the workpiece tried to move slightly closer to the punch. After compensating the whole stress in the workpiece, the spring-go phenomenon was clearly explained when the stress generated in the reverse-bending zone overcame the stress generated in the bend-allowance zone. In contrast, the spring-back was established when the stress generated in the bend-allowance zone suppressed the stress generated in the reverse-bending zone. In addition to clearly understanding the process, the parameters affecting the spring-back and spring-go features were investigated. The FEM-simulation results illustrated that the radius and height of the punch significantly affected the spring-back and spring-go processes, in addition to the bending angle obtained. The punch radius affected the material flow. Specifically, the S-curve-shaped material flow was stronger as the punch radius decreased. This S-curve-shaped material flow resulted in reverse-stress distribution and the reverse-bending zone in the leg of the workpiece, which affected the spring-back and spring-go phenomena. The reverse-stress distribution and reverse-bending zone increased as the punch radius decreased. Therefore, the amount of spring-go decreased as the punch radius increased, whereas the amount of spring-back increased as the punch radius increased. The FEM-simulation results also illustrated that the effects of the punch height could be theoretically explained on the basis of material-flow and stress-distribution analyses. The application of a very small punch height caused a large gap formation between the leg of the workpiece and the die's side; in addition, a small reverse-bending zone was generated. In contrast, the application of a very large punch height caused a large reverse-bending zone but no gap formation between the leg of the workpiece and the die side. After compensating the whole stress distribution on the workpiece, the effect of the spring-back generated in the bend-allowance zone was not sufficient to overcome the large gap and the reverse-bending zone; therefore, the spring-go effect was generated, in which the bending angle obtained was smaller than the required bending angle in both the cases of application of very small and very large punch heights. However, if the spring-back generation on the bend-allowance zone could suppress the gap between the leg of the workpiece and the die's side, and the reverse-bending zone, spring-back was generated, in which the obtained bending angle was larger than the required bending angle. Therefore, suitable process parameters must be strictly considered to achieve the required bending angle by balancing (i) the compensation of the gap between the leg of the workpiece and the die side, and the stress distribution on the reverse-bending zone and (ii) the stress distribution on the bend-allowance zone. The FEM-simulation results were validated by laboratory experiments. The FEM-simulation results showed a good agreement with those obtained by the experiments in terms of the bending force and bending angle. These results indicated that FEM simulation could be used as a tool for the theoretical elucidation of the mechanisms of the spring-back and

spring-go phenomena, in addition to studying the effects of process parameters. Furthermore, to save production cost and time, it can be applied to industrial problems to predict the bending angle and further determine suitable process parameters.

10. References

- Alexandrov, S. & Hwang, Y.M. (2009). The bending moment and spring-back in pure bending of anisotropic sheets, *International Journal of Solids and Structures* 46, Vol. 46, pp. 4361–4368.
- Atsushi, K. & Shigeru, A. (2002). Optimized bending sequences of sheet metal bending by robot, *Robotics and Computer Integrated Manufacturing*, vol. 18, pp. 29–39.
- Bakhshi-Jooybari M., Rahmani B., Daezadeh V. & Gorji A. (2009). The study of spring-back of CK67 steel sheet in V-die and U-die bending processes, *Materials and Design*, Vol. 30, pp. 2410–2419.
- Belingardi, G. & Peroni, L. (2007). Experimental investigation of plastic collapse of aluminium extrusions in biaxial bending, *International Journal of Mechanical Sciences*, Vol. 49, pp. 554–566.
- Canteli, J.A., Cantero, J.L. & Miguelez, M.H. (2008). Experimental identification of a thermo-mechanical model for air bending, *Journal of Materials Processing Technology*, Vol. 203, pp. 267–276.
- Da-xin, E., Hua-hui, He., Xiao-yi, Liu. & Ru-xin, N. (2009). Spring-back deformation in tube bending, *International Journal of Minerals, Metallurgy and Materials*, Vol. 16, pp. 177–183.
- Dufloy, J.R., Vancza, J. & Aerens, R. (2005). Computer aided process planning for sheet metal bending: A state of the art, *Computers in Industry* 56, Vol. 56, pp. 747–771.
- East, V., Darendeliler, H. & Gokler, M.I. (2002). Finite element analysis of spring-back in bending of aluminium sheets, *Materials and Design*, Vol. 23, pp. 223–229.
- Jin, M., Thipprakmas, S., Matsumoto, N., Souma, S. & Murakawa, M. (2000). Design optimization of die and punch for bending of spring materials, *Proceeding of the 51st Japaness Joint Conference for the Technology of Plasticity*, pp. 481–482, Japan.
- Kazan R., Firat M. & Tiryaki A.E. (2009). Prediction of spring-back in wipe-bending process of sheet metal using neural network. *Materials and Design*, Vol. 30, pp. 418–423.
- Kim H.S. & Koc M. (2008). Numerical investigations on spring-back characteristics of aluminum sheet metal alloys in warm forming conditions, *Journal of Manufacturing Processes*, Vol. 204, pp. 370–383.
- Lange K. *Handbook of metal forming*. New York: McGraw-Hill Inc.; 1985. p. 19.1–35.
- Leu, D.K. & Hsieh, C.M. (2008). The influence of coining force on spring-back reduction in V-die bending process, *Journal of Materials Processing Technology*, Vol. 196, pp. 230–235.
- Li, H., Yang, H., Zhan, M. & Kou, Y.L. (2010). Deformation behaviors of thin-walled tube in rotary draw bending under push assistant loading conditions, *Journal of Materials Processing Technology*, Vol. 210, pp. 143–158.
- Ling, Y.E., Lee, H.P. & Cheok, B.T. (2005). Finite element analysis of spring-back in L-bending of sheet metal, *Journal of Materials Processing Technology*, Vol. 168, pp. 296–302.

- Meinders T., Burchitz I.A., Bonte M.H.A. & Lingbeek R.A. (2008). Numerical product design: Spring-back prediction, compensation and optimization, *International Journal of Machine Tool and Manufacturing*, Vol. 48, pp. 499-514.
- Mkaddem, A. & Saidane, D. (2007). Experimental approach and RSM procedure on the examination of spring-back in wiping-die bending processes, *Journal of Materials Processing Technology*, Vol. 189, pp. 325-333.
- Panthi, S.K., Ramakrishnan, N., Pathak, K.K. & Chouhan, J.S. (2007). An analysis of springback in sheet metal bending using finite element method (FEM), *Journal of Materials Processing Technology*, Vol. 186, pp. 120-124.
- Papeleux, L. & Ponthot, J.P. (2002). Finite simulation of springback in sheet metal forming, *Journal of Materials Processing Technology*, Vol. 126, pp. 785-791.
- Sanchez, L.R. (2009). Modeling of spring-back, strain rate and Bauschinger effects for two-dimensional steady state cyclic flow of sheet metal subjected to bending under tension, *International Journal of Mechanical Sciences*, Vol. 52, pp. 429-439.
- Schuler. *Metal forming handbook*. Berlin, Heidelberg, New York: Springer-Verlag; 1998. p. 366-373.
- Tekaslan, O., Seker, U. & Ozdemir, A. (2006). Determining springback amount of steel metal has 0.5 mm thickness in bending dies, *Materials and Design*, Vol. 27, pp. 251-258.
- Tekaslan O., Gerger N. & Seker U. (2008). Determination of spring-back of stainless steel sheet metal in "V" bending dies, *Materials and Design*, Vol. 29, pp. 1043-1050.
- Thipprakmas S. & Rojananan S. (2008). Investigation of spring-go phenomenon using finite element method, *Materials and Design*, Vol. 29, pp. 1526-1532.
- Thipprakmas S. (2010). Finite element analysis of punch height effect on V-bending angle, *Materials and Design*, Vol. 31, pp. 1593-1598.
- Vsquez-Ojeda, C. & Ramos-Grez, J. (2009). Bending of stainless steel thin sheets by a raster scanned low power CO2 laser, *Journal of Materials Processing Technology*, Vol. 209, pp. 2641-2647.
- Wang J., Verma S., Alexander R. & Gauc J. (2008). Spring-back control of sheet metal air bending process, *Journal of Manufacturing Processes*, Vol. 10, pp. 21-27.
- Woznicaa, K. & Klepaczkob, J.R. (2003). Modeling of inelastic bending of a metal sheet with thermal coupling, *International Journal of Mechanical Sciences* 45, Vol. 45, pp. 359-372.
- Yilamu, K., Hino, R., Hamasaki, H. & Yoshida, F. (2010). Air bending and spring-back of stainless steel clad aluminum sheet, *Journal of Materials Processing Technology*, Vol. 210, pp. 272-278.
- Yu H.Y. (2009). Variation of elastic modulus during plastic deformation and its influence on spring-back, *Materials and Design*, Vol. 30, pp. 846-850.

Analysis of Welding Residual Stresses and Its Applications

Byeong-Choon Goo, Jung-Won Seo and Seung-Yong Yang
Korea Railroad Research Institute
Korea

1. Introduction

Welding has been one of the most effective joining methods for structural steel from its advent in the middle of 1930s. Welding has increased productivity of shipbuilding, bridge building, construction, etc. Though welding has many advantages, it has several drawbacks such as welding deformation, residual stresses and welding defects. Regarding structural integrity, welding residual stresses are a key factor to be considered. A lot of study on welding residual stresses has been conducted in experimental and numerical areas, but the effect of the residual stresses on mechanical behavior, fatigue life, etc. has not been clearly explained until yet. Some results show the effect of welding residual stresses on fatigue lives is significant. On the other hand, other results show specimens with tensile residual stress have even longer fatigue lives than parent materials. These unclear conclusions seem to be mainly concerned with residual stress relaxation during fatigue tests and the metallurgical difference of the welded materials and parent materials. Therefore, understanding of residual stress relaxation by an applied loading is very important for fatigue life prediction and structural integrity analysis of welded structures. The improvement of numerical analysis methods has contributed very much to the understanding of role of residual stresses in weldments. In particular, finite element method (FEM) has been a powerful tool for analysis of welding process and related problems. In the following, we will deal with several numerical analysis problems concerned with welding. They are welding residual stresses, welding residual stress relaxation, residual stress effect on the fatigue and fracture of welding details. All finite element analyses were carried out by using ABAQUS code.

2. Analysis of residual stress relaxation due to mechanical loading

A lot of study on welding residual stress has been conducted in experimental (Iida and Takanashi, 1998) and numerical areas (Yang and Goo, 2005), but the effect of the residual stresses on fatigue life has not been clearly explained until yet. In this study, we studied the residual stress relaxation by the finite element analysis. Initial residual stresses of a welded specimen were generated by the simulation of welding process. Residual stress relaxation under applied loading was examined by elastoplastic finite element analysis.

2.1 Generation of residual stress by welding

Two dimensional welding analyses are carried out to generate initial residual stresses in the plate shown in figure 1. Longitudinal welding is carried out along the middle line of the plate. The width of the weld bead is 4 mm. After welding, metallurgical structure is changed. In the experimental study (Iida and Takanashi, 1998), a large hole enough to rule out the effect of the metallurgical structure change is pierced. The process of hole drilling is simulated by using the element removal command. Figure 2 shows the finite element mesh. A half of the specimen is considered. The bottom of the specimen ($y=-95$) is fixed and longitudinal loading is applied to the opposite side. All finite element analyses were carried out by ABAQUS Code. Figures 3 and 4 show physical and material properties used for the simulation. A kinematic hardening rule was used. First of all, temperature field by welding is obtained. By using the temperature field, elastoplastic finite element analysis is performed. Finally the hole drilling analysis is carried out. The distribution of residual stress σ_{yy} along the x-axis from the hole center is shown in figure 5. Before the hole drilling, the maximum residual stress σ_{yy} is nearly the same as the yield stress of the base material. An interesting point is that after the hole drilling the residual stress increases.

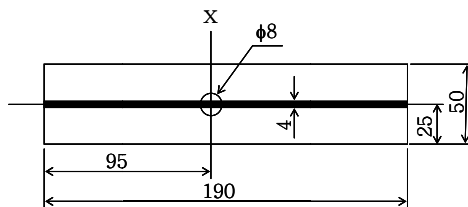


Fig. 1. Welded plate, thickness=4



Fig. 2. Finite element model

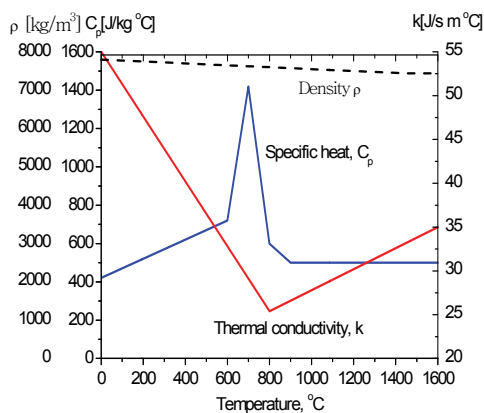


Fig. 3. Material properties: conductivity k , density ρ , and specific heat, c_p .

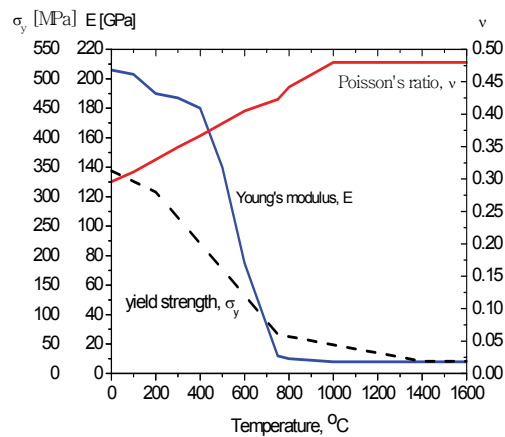


Fig. 4. Young's modulus E , Poisson's ratio ν , and yield stress σ_y

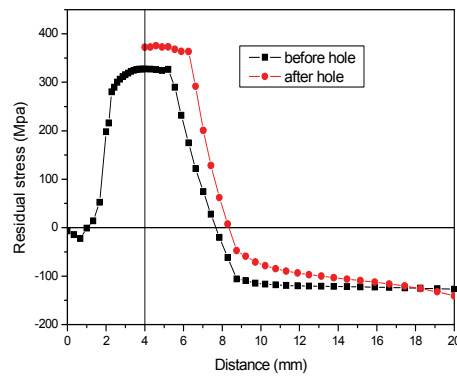


Fig. 5. Residual stress σ_{yy} before and after the hole drilling

2.2 Residual Stress Relaxation by applied loading

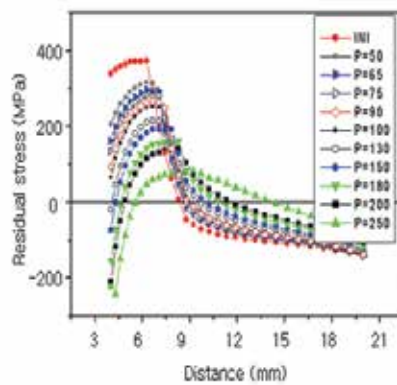


Fig. 6. Residual stress variation by applied loads with stress ratio, $R=0$

Residual stress relaxation under mechanical loading was studied by the finite element analysis. Mechanical loads were applied to the finite element model shown in figure 1. The plate has initial residual stresses from the welding and hole drilling. Figure 6 shows the residual stress relaxation under a series of loads. The loads increase to the maximum value from zero and decrease linearly to zero.

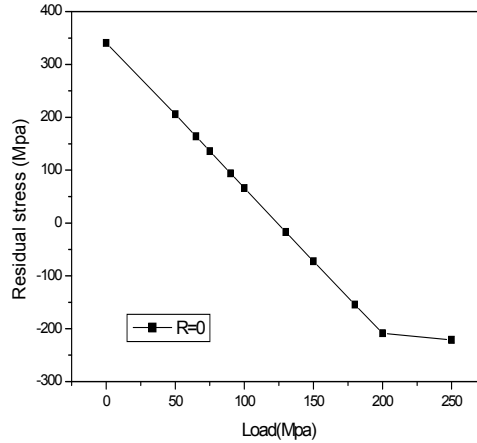


Fig. 7. Variation of residual stress, σ_{yy} at the hole edge by the applied loads

Figure 7 shows the variation of residual stress σ_{yy} at the hole edge ($x=4$). The residual stress decreases linearly until the applied load increases to 200 MPa. But when the applied load is 250 MPa, the magnitude of the residual stress relaxation is similar to that in case of 200 MPa. When the applied load is 250 MPa, the residual stress after unloading exceeds the compressive yield stress. Figure 8 shows this process. To quantify the effect of the applied loads on the residual stress relaxation, the normalized residual stresses after unloading are represented as a function of the normalized applied loads as shown in figure 9. The relation is expressed as:

$$\sigma_{res} / \sigma_{res_ini} = a(\sigma_{res_ini} + \sigma_{app}) / \sigma_y + b \quad (1)$$

where σ_{res} : residual stress after unloading; σ_{res_ini} : initial residual stress; σ_{app} : applied load; σ_y : initial yield stress; a, b: constants. $a=-0.87$ and $b=1.88$ are obtained. It is found that for the material considered in this study, the welding residual stress, σ_{yy} decreases to zero when $(\sigma_{res_ini} + \sigma_{app}) / \sigma_y = 2.3$.

To examine the behavior of residual stress relaxation in case that fully reversal loads (Load ratio, $R=-1$) are applied, a tension-compression load of 200 MPa is applied to the same welded specimen used above. In figure 10, the simulation results are shown and compared with the results in case of $R=0$. The relaxation of the residual stress, σ_{yy} at the hole edge is negligible.

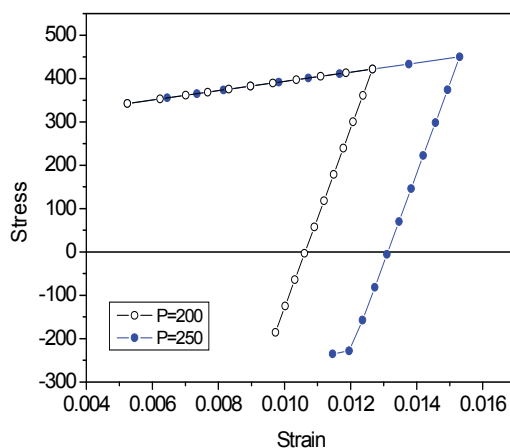


Fig. 8. Finite element analysis of the process of residual stress relaxation

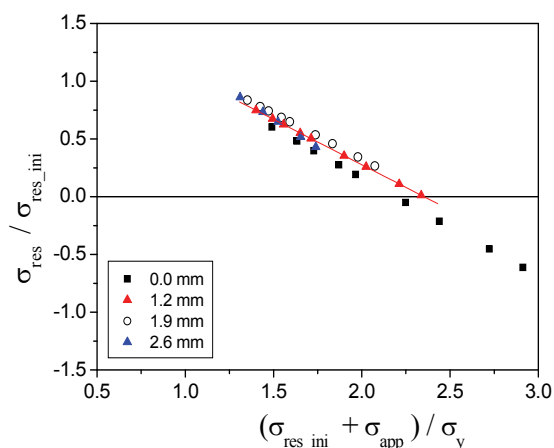


Fig. 9. Normalized residual stresses vs. applied loads. Distance is from the hole edge

From the hole edge to 3mm, the residual stress relaxation occurs. The stress-strain history is shown in figure 11 when the applied load is 200 MPa. Figure 12 shows the simulation results of residual stress relaxation under various applied loads. The residual stress at the hole edge decreases to a certain value and then increases as the applied loads increase. Figure 13 compares the simulation results when $R=0$ and -1 .

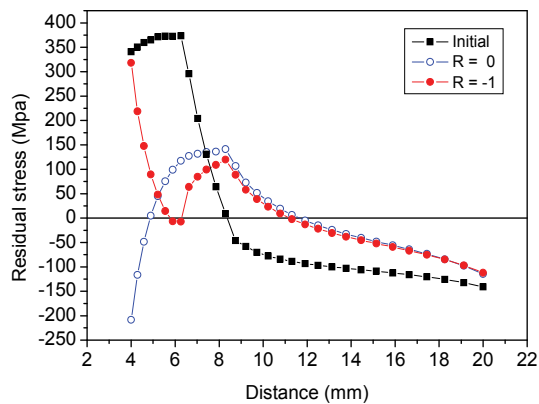


Fig. 10. Effect of stress ratio on the residual stress relaxation. Distance is from the hole center.

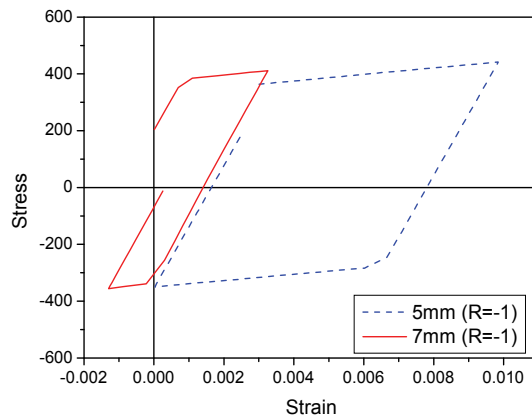
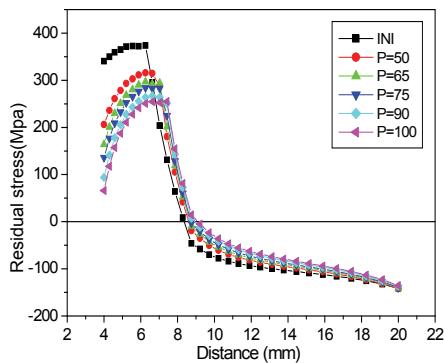
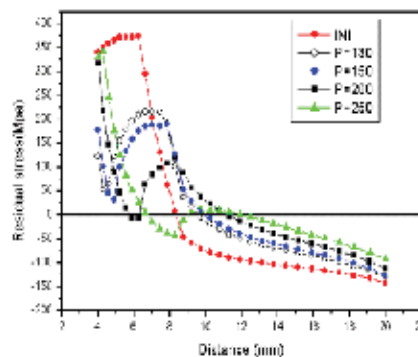


Fig. 11. History of $\sigma - \varepsilon$ under $R=-1$. Distance is from the hole center.



(a)



(b)

Fig. 12. Residual stress relaxation in case of $R=-1$

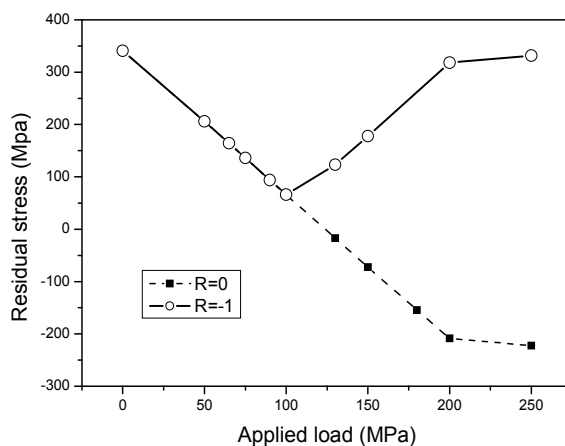


Fig. 13. Residual stress relaxation at the hole edge when $R=0$ and -1

3. Fatigue crack propagation under residual stresses

In general, cyclic loading is applied to mechanical structures such as rolling stock during the service time, and the structures reach to the final failure due to accumulated fatigue damage. To extend the lifetime of metallic structures under fatigue loading, several methods have been tried, for example extending the time period of crack nucleation or repairing whole section with cracks fully penetrating through the thickness of the section, etc. Song (Song, 2001) has proposed to generate compressive residual stress around the crack tip by heating to retard or stop the fatigue crack growth.

In this chapter, we study the effect of welding residual stresses on fatigue behavior by finite element method. The compact tension specimen is chosen as the analysis object, and the welding line is perpendicular to the direction of the crack propagation. Temperature, residual stress distributions and residual stress redistribution with fatigue crack growth were calculated on the commercial package ABAQUS. There are several procedures for crack growth, for example element extinction, node release and interface-cohesive element, etc. In this paper, to simulate the fatigue crack growth, a cohesive zone model was used along the symmetric line in the crack direction. To account for the damage accumulation due to fatigue, a law of damage evolution was included in the constitutive equation of the cohesive zone model. The analysis was also applied to the specimen not welded. Comparing results obtained from the welded and not-welded specimens, how the residual stresses exert influence on the fatigue behavior was examined.

3.1 Analysis of welding residual stress

To obtain residual stress distribution by welding, it is necessary first to conduct analysis for transient heat transfer, and then obtain mechanical stress distribution from the temperature distribution in the structure. A compact tension (CT) specimen is chosen as the analysis object. Figure 14 shows the two-dimensional finite element modeling of the CT specimen. Only the top half of the full specimen is analyzed by using the symmetry condition to the

center line. The height of the model is 75 mm and the distance between the notch and the right hand side of the specimen (i.e., length of the center line) is 100 mm. Heat input of $3.68 \times 10^8 \text{ J/m}^2\text{s}$ is applied on the surface of the element along the C-D line from top to bottom with speed of 2.42 mm/s. The distance between the welding line C-D and the notch is 60 mm. Convective boundary condition with ambient temperature (20°C) is applied to the boundary lines of the specimen, and heat transfer does not take place across the symmetry line. The material parameters including the latent heat of melting, density variation with temperature, etc. are illustrated in Figure 15. Temperature distribution captured at a time during the welding procedure is illustrated in Figure 16. The maximum temperature is about 1500°C .

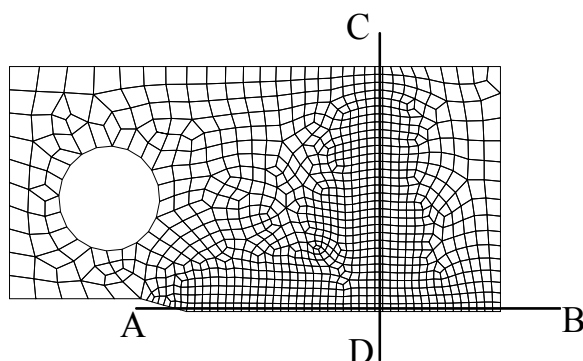


Fig. 14. Finite element mesh for CT specimen. Height of the model is 75mm, and the distance between the notch and the right hand side boundary is 100mm. Along the A-B line, the interface-cohesive elements will be located in the fatigue analyses.

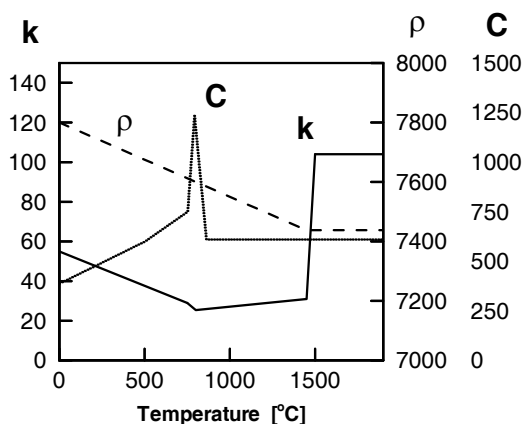


Fig. 15. Variation of conductivity k , density ρ and specific heat C with temperature.

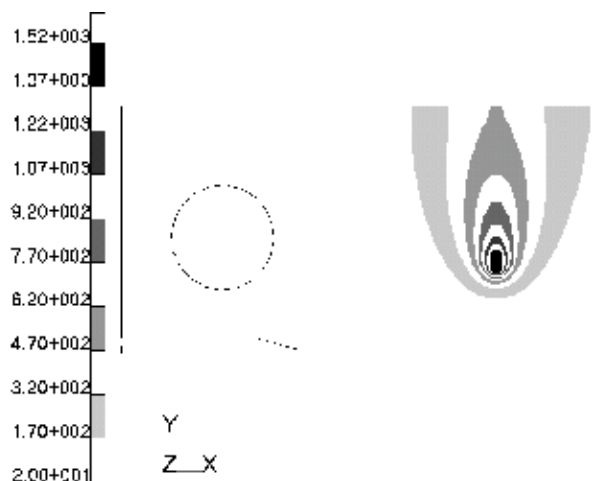


Fig. 16. Temperature distribution at a time during the welding procedure.

The stress distribution induced by the welding was calculated assuming temperature dependent elastic-plastic constitutive behavior for the specimen. The temperature-dependent variation of Young's modulus, Poisson's ratio and approximate strain-hardening behavior of mild steel are given in Figure 17. The resultant residual stress distribution in the CT specimen is shown in Figure 18. It is observed that compressive vertical residual stress, σ_{22} , is generated near the notch by welding. The profiles of the vertical and horizontal residual stresses along the A-B line are drawn in Figure 18 (b). The magnitude of the horizontal residual stress turns out to be small compared to that of the vertical residual stress.

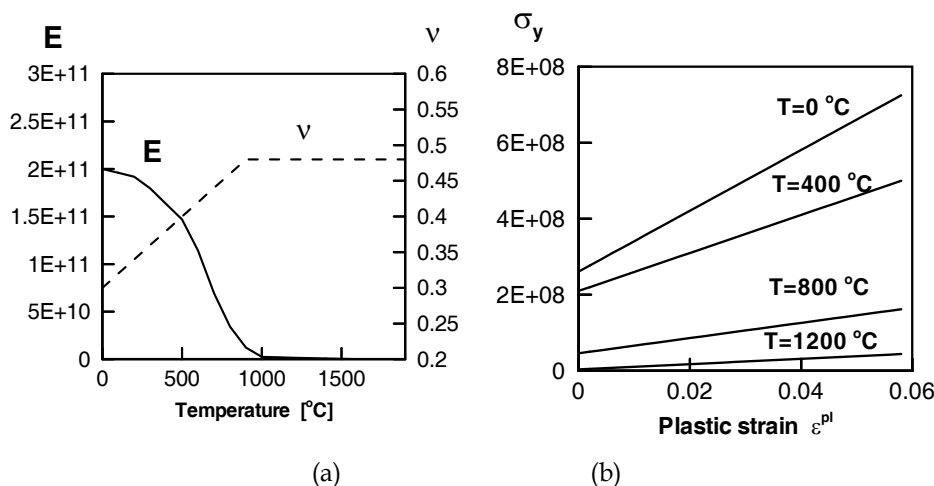


Fig. 17. (a) Variation of Young's modulus and Poisson's ratio with temperature. (b) Yield strength and plastic strain curves at several temperatures.

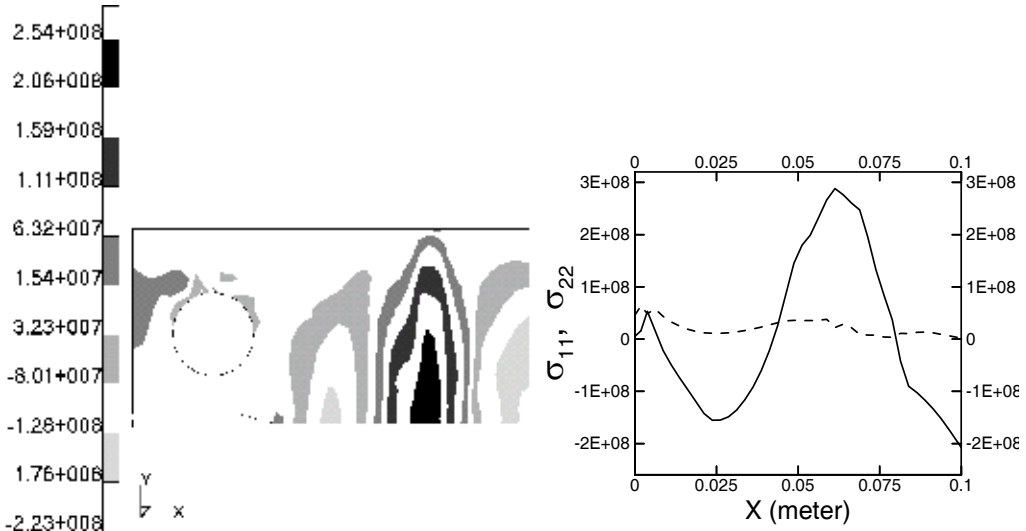


Fig. 18. (a) Vertical residual stress σ_{22} distribution by welding along the C-D line. (b) Welding residual stress σ_{11} and σ_{22} distributions along the A-B line in Figure1; dashed line is σ_{11} and solid line is σ_{22} .

3.2 Cohesive zone model for fatigue crack growth

Cohesive zone models have been developed by many researchers (Needleman, 1987; Rose, Ferrante and Smith, 1981; de-Andres, Perez and Ortiz, 1999). In the viewpoint of cohesive zone model, fracture is regarded as a gradual phenomenon in which separation takes place across an extended crack tip, or cohesive zone, and is resisted by cohesive forces. De-Andres et al. (de-Andres, Perez and Ortiz, 1999) used three-dimensional irreversible cohesive zone model to account for the accumulation of fatigue damage by allowing unloading irreversibility. Roe and Siegmund (2003) developed a constitutive law for the cohesive zone model for fatigue crack growth simulation. In their model, the cohesive forces were reduced by the accumulation of fatigue damage, and a time evolution equation for a damage parameter was used to calculate the current state of damage. Here, we used their time evolution equation of damage to study the fatigue crack growth behavior.

To describe the cohesive laws used in this chapter, we confine our interest in two-dimensional case. An effective traction t is given for a given effective displacement, δ , as follows (de-Andres, Perez and Ortiz, 1999).

$$t = \begin{cases} \lambda \sigma_c \frac{\delta}{\delta_c} e^{-\frac{\delta}{\delta_c}} & \text{if } \delta = \delta_{\max} \text{ and } \dot{\delta} \geq 0 \\ \frac{t_{\max}}{\delta_{\max}} \delta & \text{if } \delta < \delta_{\max} \text{ or } \dot{\delta} < 0 \end{cases}$$

where $\lambda = \exp(1)$ and σ_c is the cohesive strength of material and δ_c is a characteristic opening displacement. The first equation is for the loading case, and the second one is for unloading case. δ_{\max} is the maximum attained displacement by loading, and $t_{\max} = t(\delta_{\max})$ in the first equation. In this model, the $t-\delta$ response follows different path for unloading, and the path with unloading is always directed to the origin. For the material separation vector $\delta = \delta_s \mathbf{e}_s + \delta_n \mathbf{e}_n$ where δ_s and δ_n are the sliding and normal opening displacements at the cohesive interfaces, respectively, the effective displacement δ is defined in different ways for normal opening and compression. That is, for normal opening ($\delta_n \geq 0$), the effective displacement and the traction vector are given by

$$\delta = \sqrt{\beta^2 \delta_s^2 + \delta_n^2}$$

$$\mathbf{t} = \frac{t}{\delta} (\beta^2 \delta_s \mathbf{e}_s + \delta_n \mathbf{e}_n)$$

where β is a weighting factor. Meanwhile, for normal compression ($\delta_n < 0$),

$$\delta = \sqrt{\beta^2 \delta_s^2}$$

$$\mathbf{t} = \frac{t}{\delta} \beta^2 \delta_s \mathbf{e}_s + t_n \mathbf{e}_n$$

$$t_n = k_s \left. \frac{\partial t}{\partial \delta} \right|_{\delta=0} \delta_n = k_s \lambda \sigma_{c0} \frac{\delta_n}{\delta_c}$$

The last expression means that the cohesive material behaves as a linear spring for normal compression. The stiffness of the compressive spring is taken as a factor k_s times the slope of the effective traction-separation curve at the origin. σ_{c0} is used instead of σ_c to emphasize that the slope is fixed as a constant by the initial value of the cohesive strength. Refer to later sections to see the meaning of varying cohesive strength. The cohesive constitutive equations were implemented using the user interface UEL of ABAQUS. Figure 19 illustrates loading responses predicted for pure normal displacement and pure sliding by the cohesive traction-separation laws. We can observe that the traction force varies linearly to the displacement in the case of normal compression. The shear traction is symmetric to the origin with the sliding displacement. Meanwhile, if the normal opening and sliding displacements are applied simultaneously, the maximum tensile normal traction resisted by the material will be smaller than the cohesive strength of material, σ_c .

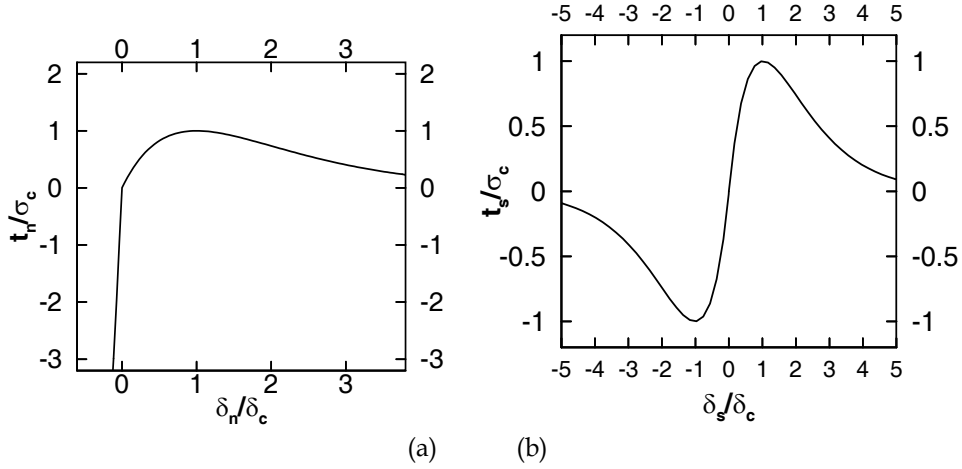


Fig. 19. (a) Normal traction vs. normal displacement relation with $\delta_s = 0$. (b) Shear traction vs. shear displacement relation with $\delta_n = 0$.

With the cohesive constitutive equations described above, cyclic loading with the maximum traction below the cohesive strength will result in infinite life. To establish a cohesive zone model that can capture finite life effects, the cohesive zone constitutive equations should include evolving cohesive properties with time. Roe and Siegmund (Roe and Siegmund, 2003) applied damage mechanics to the development of an evolution equation for a damage parameter. The damage parameter was incorporated into the traction-separation law via the cohesive traction. That is, the cohesive constitutive equation for damage accumulation under cyclic loading is given by using the current cohesive strength defined by

$$\sigma_c = (1 - D)\sigma_{c0}, \quad 0 \leq D \leq 1$$

where D is the damage parameter, and σ_{c0} is the initial cohesive strength. To compute the current state of damage, a damage evolution equation was proposed as follows.

$$\dot{D} = \frac{|\dot{\delta}|}{\delta_D} \left(\frac{t}{\sigma_c} - \frac{\sigma_f}{\sigma_{c0}} \right) H(\bar{\delta} - \delta_f), \quad \dot{D} \geq 0$$

where H denotes the Heavyside function, $\bar{\delta} = \int |\dot{\delta}| dt$, σ_f is an endurance limit below which cyclic loading can proceed infinitely without failure, and δ_D , δ_f are positive material constants. $|\dot{\delta}|$ is defined by

$$|\dot{\delta}| = \begin{cases} \sqrt{\beta^2 \dot{\delta}_s^2 + \dot{\delta}_n^2} & \text{for } \delta_n \geq 0 \\ \sqrt{\beta^2 \dot{\delta}_s^2} & \text{for } \delta_n < 0 \end{cases}$$

3.3 Results of fatigue crack growth simulation

To investigate the effect of welding residual stress on fatigue behavior, compact tension (CT) specimens with and without residual stresses are analyzed by two-dimensional plane strain finite elements with thickness 1m. By symmetry of the problem only top half of the specimen is analyzed, and the cohesive interface elements are located along the symmetry line in the crack propagation direction. The material data used for the cohesive laws in this investigation are listed in Table 1. The values are presumed approximately, and they need to be calibrated by comparing with experimental works in the future. The surrounding matrix is modeled by elastic-plastic solid. The material properties of welded specimen can be different from toes of not-welded specimen because the microstructures are changed by welding. But, in this computational work, the same material properties are assumed before and after welding.

First, Figure 20 and Figure8 illustrate results of fatigue crack growth simulation of the specimen without welding residual stress. Repeated triangle type displacement loading oscillating between 0 and the maximum amplitude of 0.1 mm with period of 2 seconds was applied vertically at the top point of the hole in the CT specimen. The crack extension with respect to time is shown in Figure7. In this case, the level of the applied traction to the cohesive zone is less than the cohesive strength of the material, but as the cyclic loading proceeds fatigue crack propagates. The slope of the crack extension-time curves were

determined for a series of the stress intensity ranges $\frac{da}{dN} = C\Delta K^n$ at a point (15 mm away

from the notch along the A-B line) and plotted in a log-log scale. Circles in Figure 21 summarize these data. To obtain the value of the stress intensity factor, the mathematical formula (Anderson, 1995) or finite element analysis can be used. In this work, ABAQUS option **contour integral* was used to compute the stress intensity factor, and the result was almost the same as that by the formula. The numerically predicted data could be fitted to a Paris type power law $da/dN = C\Delta K^n$ with $n \approx 3.32$. Figure 22 illustrates the vertical stress σ_{22} distribution generated during the fatigue crack growth in load control simulation in which maximum point force of 5 MN is applied periodically at the top point of the hole. Small amount of compressive residual stress is induced behind the crack tip which causes closure of crack surfaces.

Next, to investigate fatigue behavior under the welding residual stress, specimen with the welding residual stress distribution obtained in the previous section was tested for cyclic point force loading with amplitude 5 MN. Figure 23 shows redistributions of welding residual stress with the fatigue crack propagation. Compressive vertical stress σ_{22} is redistributed as the fatigue crack grows into the CT specimen. To satisfy the equilibrium condition the tensile residual stress is also relaxed by little amount. In the figure, the position of crack tip corresponds to the location where $D = 0.5$. Triangles in Figure8 plot

the crack growth rate- $\frac{da}{dN} = C\Delta K^n$ behavior on a log-log scale for the case with welding residual stress. ($\frac{da}{dN} = C\Delta K^n$ is the values calculated for the specimen without welding residual stress.) As seen in the plot, the computation predicts reduction of the crack growth rate for the case with welding residual stress. One can expect that crack closure takes place by the compressive vertical welding residual stress, and the effective traction in the cohesive zone should be reduced. In this case, the effective stress intensity factor ΔK_{eff} , which is defined by the difference between ΔK and the stress intensity at which the crack opens, is less than ΔK , and the crack growth rate is expected to decrease.

σ_{c0}	1000 [M Pa]	δ_D	0.5 [mm]
δ_c	0.05 [mm]	σ_f	150 [M Pa]
δ_f	0.25 [mm]	k_s	2

Table 1. Numerical parameters for the cohesive zone model.

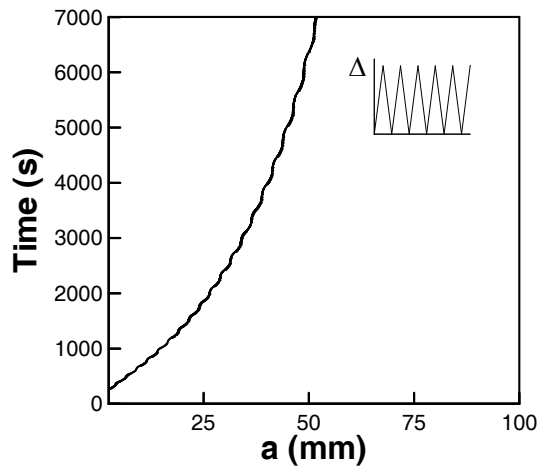


Fig. 20. Numerical fatigue crack growth result for constant amplitude fatigue without residual stress. Amplitude of the cyclic displacement loading was 0.1 mm. 7000 seconds corresponds to 3500 cycles.

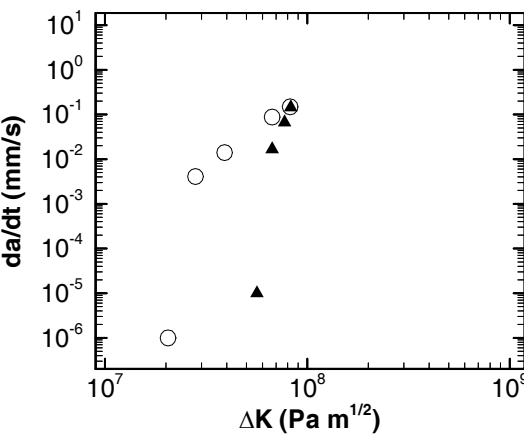


Fig. 21. Numerical fatigue crack growth rates with and without welding residual stress. Circles are results without welding residual stress, and triangles are results with welding residual stress

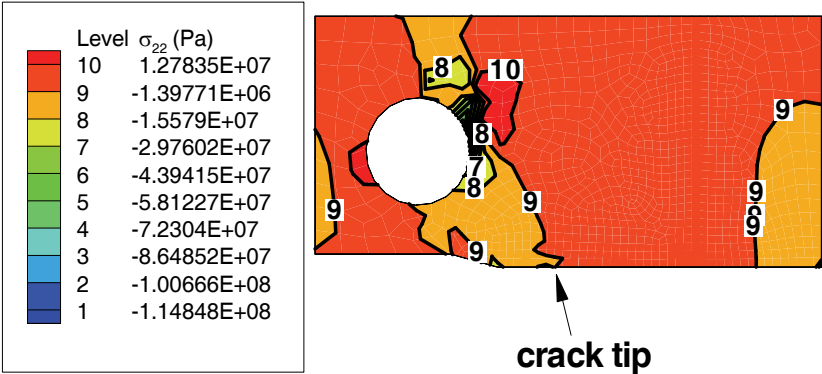
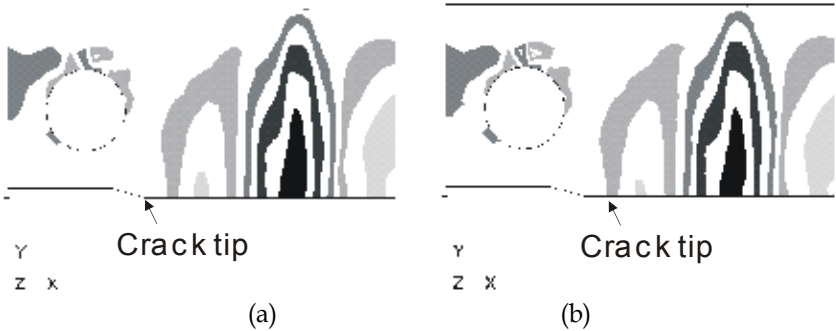


Fig. 22. Vertical residual stress σ_{22} distribution after 125 cycles. Vertical point force is applied periodically at the top of the hole. This analysis is for the case without welding residual stress. Behind the crack tip, small amount of compressive residual stress is generated.



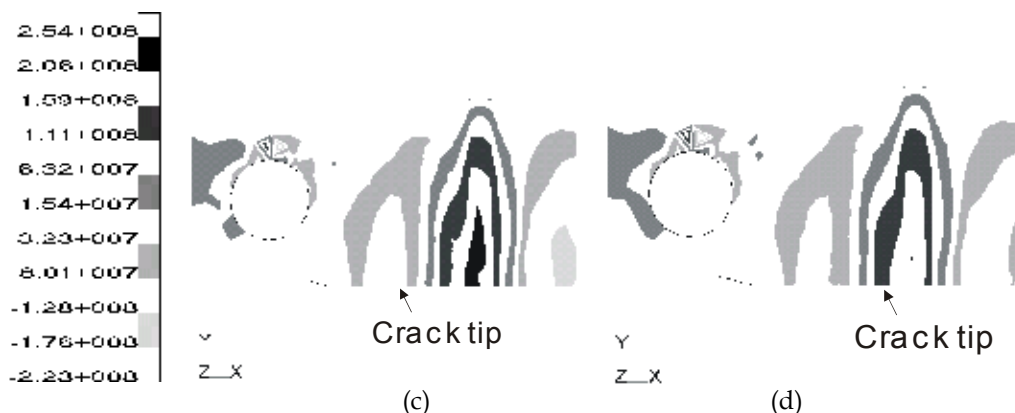


Fig. 23. Redistributions of welding residual stress σ_{22} with the crack growth. The contour plots are drawn at external-loading-free-states. Location of the crack tip is marked at each case. (a) After 50 cycles. (b) After 100 cycles. (c) After 200 cycles. (d) After 250 cycles.

4. Fatigue life estimation considering residual stresses

Residual stresses play a significant role in structural behavior of welded joints. According to our tensile fatigue tests at 20 Hz, $R=0.1$ on butt-welded joints, fatigue strength of as-welded specimens, that is, specimens having residual stress is higher than that of annealed specimens in the short life, but vice versa in the long life (Figures 24, 25).

To simulate this behavior, taking into account the residual stress relaxation and notch effect, we developed a model to evaluate fatigue lives of welded joints with residual stresses. To evaluate the fatigue lives of welded structures by a strain-life approach, it is necessary to identify the value of parameters through tests or literature. But it is difficult for field engineers to get the necessary data through tests from the viewpoint of time and cost. Therefore, some researchers (Lida, 2001; Reemsnyder, 1981; Lawrence, Burk and Yung, 1982) have been studying to obtain fatigue properties from simple tensile tests and hardness measurement. We surveyed literature and quoted a procedure to identify the fatigue parameters for fatigue life evaluation based on a local strain approach.

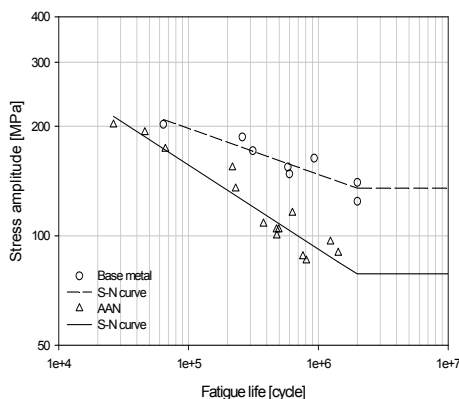


Fig. 24. S-N curves for base material and as-welded (AAN) specimens

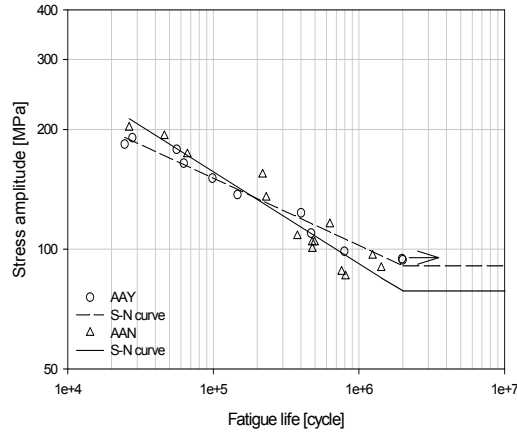


Fig. 25. S-N curves for annealed (AAY) and not annealed (AAN)

4.1 Relations between mechanical and fatigue properties

When mean stress effect is neglected, one of the equations frequently used for local strain-life approach is Coffin-Manson's (Vormwald and T. Seeger, 1987; Stephens, Fatemi, Stephens and Fuchs, 2001)

$$\frac{\Delta \varepsilon}{2} = \frac{\Delta \varepsilon_e}{2} + \frac{\Delta \varepsilon_p}{2} = \frac{\sigma'_f}{E} (2N_f)^b + \varepsilon'_f (2N_f)^c \quad (1)$$

Where $\Delta \varepsilon / 2$, $\Delta \varepsilon_e / 2$ and $\Delta \varepsilon_p / 2$ are total, elastic and plastic strain amplitude respectively. σ'_f , ε'_f , b and c are constants. N_f is the number of cycles to failure. Roessle and Fatemi (Lawrence, J. D. Burk and J. Y. Yung, 1982) obtained σ'_f and ε'_f as a function of Brinell hardness, HB and/or Young's modulus, E :

$$\sigma'_f = 4.25HB + 225 \text{ (MPa)}; \quad \varepsilon'_f = \frac{0.32(HB)^2 - 487HB + 191000}{E} \quad (2a, b)$$

And using $b = -0.09$ and $c = -0.56$, they expressed Eq. (1) with the Brinell hardness and Young's modulus:

$$\frac{\Delta \varepsilon}{2} = \frac{4.25HB + 225}{E} (2N_f)^{-0.09} + \frac{0.32(HB)^2 - 487HB + 191000}{E} (2N_f)^{-0.56} \quad (3)$$

To consider mean stress effect, we express Basquin's equation as a following formula:

$$\sigma_{max} = \Delta\sigma / 2 = \sigma'_f (2N_f)^b \quad (4)$$

where σ_{max} is maximum local stress. Multiplying Eq. (3) by Eq. (4), we obtain a parameter expressed as a function of the Brinell hardness and Young's modulus:

$$\frac{E\sigma_{max}\Delta\varepsilon}{2} = (4.25HB + 225)^2 (2N_f)^{-0.18} + (4.25HB + 225)[-0.32(HB)^2 - 487HB + 191000](2N_f)^{-0.65} \quad (5)$$

4.2 FE model and fatigue notch factors

We evaluated the fatigue lives of four kinds of specimens with a width of 25 mm (Figure 26) at a stress ratio of $R = 0.1$. The weld bead profiles necessary for modeling were measured on a contour measuring instrument, Mitutoyo CV-3000S4, with a resolution of 0.0002 mm. FE models are shown in Figure 27. A half of the specimen is considered. The obtained stress concentration factors, K_t , by FE analysis and fatigue notch factors, K_f , by Peterson's formula [Stephens, Fatemi, Stephens and Fuchs, 2001] are shown in Table 2.

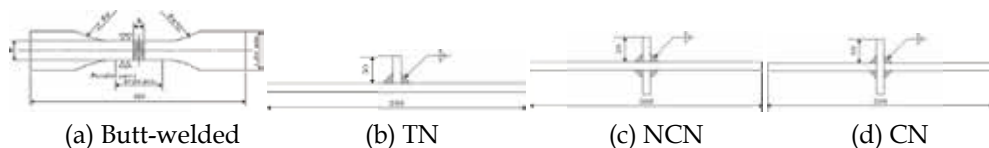


Fig. 26. Specimens for fatigue tests and FE analysis

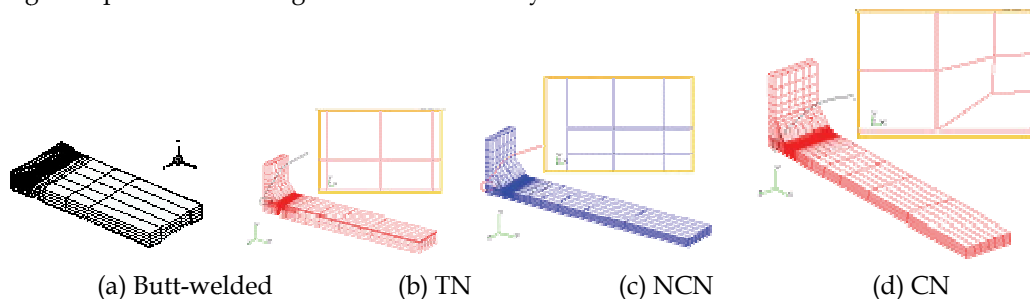


Fig. 27. FE analysis models

Specimens	Butt-weld	TN	NCN	CN
K_t / K_f	1.60/1.52	2.71/1.63	3.64/1.93	4.22/2.40

Table 2. Calculated stress concentration factors and fatigue notch factors

4.3 FE analyses of residual stress and residual stress relaxation

The analyses of welding residual stress and residual stress relaxation were carried out on ABAQUS. Weld beads are generated by the element birth technique supplied on ABAQUS. The input data for welding is shown in Table 3 and material constants such as density, ρ ,

conductivity, C , specific heat, k , Young's modulus, E , Poisson's ratio, ν (Davis, 1990), plastic strain-yield strength relationship, $\varepsilon_{pl} - \sigma_y$, used for the FE analyses are shown in Figure 28. They are dependent on temperature.

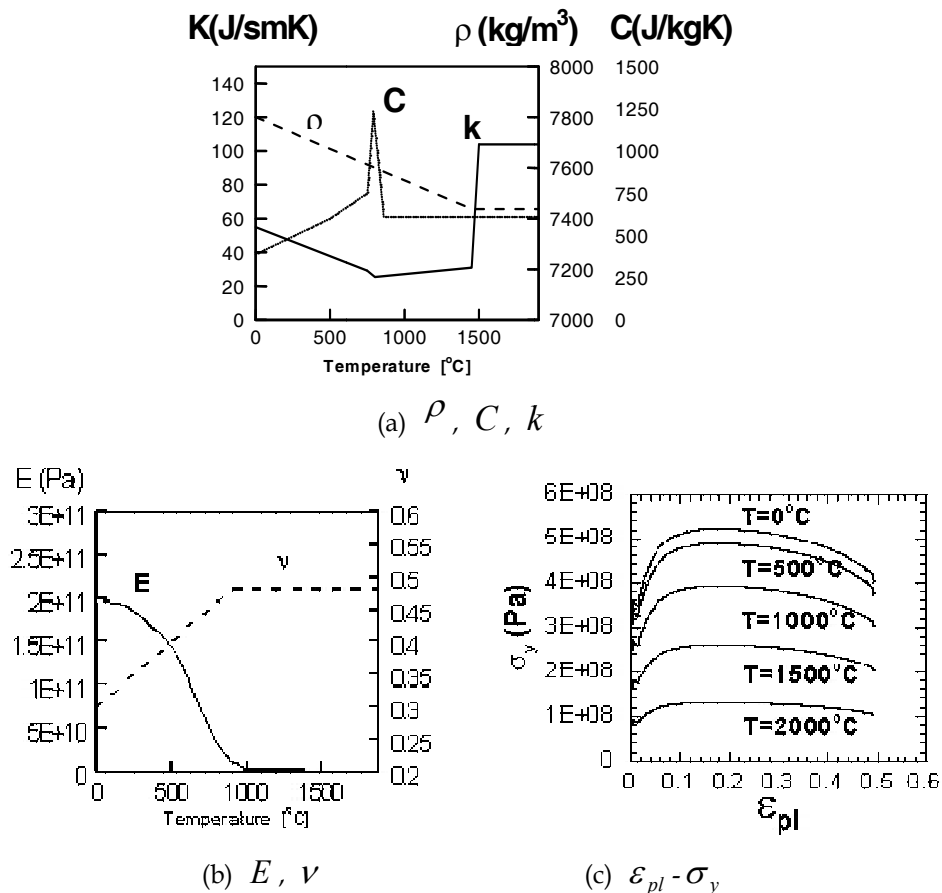


Fig. 28. Material properties used for FE analysis

Parameters	Filler metal temp. [°C]	Solidus temp. [°C]	Liquidation latent heat [°C]	Arc efficiency	Voltage [V]	Current [A]	Arc beam radius [m]	Heat convection coefficient [J/sm ² °C]
Butt-weld	2000	1465	1544	0.4	105	150	0.005	10
TN,NCN,CN	2000	1465	1544	0.4	28	270	0.005	10

Table 3. Parameters for weld analysis

Figure 29 shows the residual stress relaxation by an applied loading at a stress ratio $R = 0.1$. The loading increases linearly to S_{\max} and decreases linearly to S_{\min} . The positions A and D are shown in Figure 30.

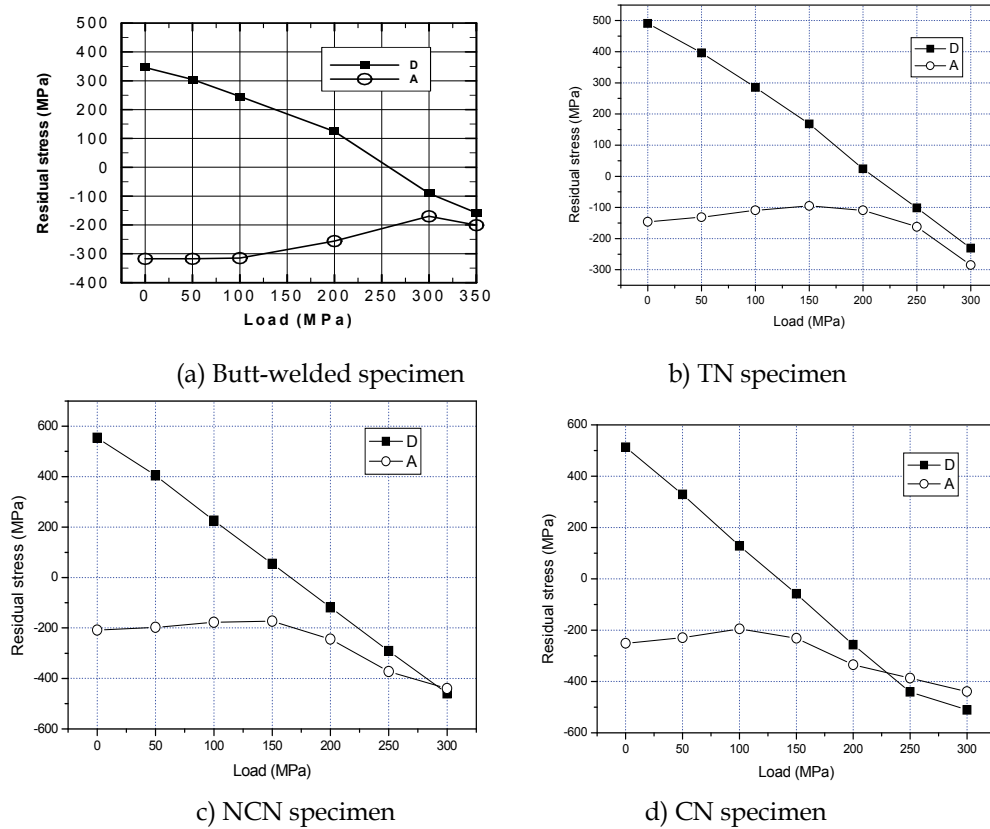


Fig. 29. Residual stress relaxation at A and D by applied loading



Fig. 30. Locations A and D in Figure 6 for the residual stress relaxation

The initial residual stress, σ_{11} (σ_{xx}), due to welding is tensile in the middle part of the specimen (near D, F, G), while compressive near the edge (near A, B, C). When the external loading is applied, the residual stresses are redistributed. At point D, the residual stress, σ_{11} , decreases with the applied loading. When the applied load, S_{\max} , is greater than a certain value, the residual stress at point D changes to compressive residual stress for all specimens. At point A, the residual stress, σ_{11} , increases to a certain value with the applied loading and then decreases. In case of the CN specimen, the residual stress at point A becomes larger than the residual stress at point D when the applied loading is larger than about 230 MPa. When the applied load is less than 230 MPa, cracks will begin at point D. But the applied load is larger than 230 MPa, cracks will begin at point A. This means the

locations of fatigue crack initiation depend on the magnitude of applied load.

4.4 Fatigue life evaluation of welded joints

We applied Eq. (5) to evaluate the fatigue lives of four kinds of welded joints aforementioned. We measured micro Vickers hardness of the butt-welded specimens with 5kg loading. Figure 31(a) shows the measuring points. Fatigue failure occurs at the notch roots located near the boundary of the weld metal and heat-affected zone. In this area, the Vickers hardness is about 150. The measured Brinell hardness ranged from 147 to 156 HB. For the evaluation of fatigue lives, we used the Brinell hardness 150 HB. Figures 32-35 show the predicted and experimental results. The experimental results were obtained from as-welded specimens. The predicted results with residual stress were obtained by considering the maximum initial residual stress obtained during the welding analysis above-mentioned. In case of the predicted results without residual stress, the initial residual stress was assumed to be zero. In the short life, the effect of residual stress on the fatigue strength is less than that in the long life. This means when the applied load is high, the initial residual stress relaxes much. But when the applied load is low, the mechanical behavior of the notches is elastic. In this case the residual stress plays a role as a mean stress.

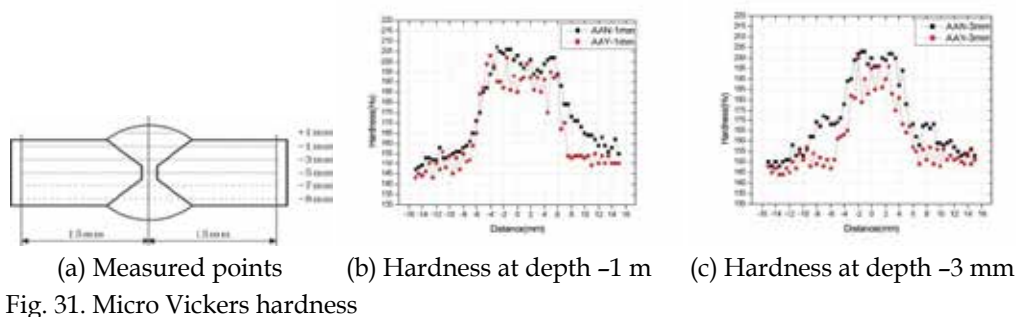


Fig. 31. Micro Vickers hardness

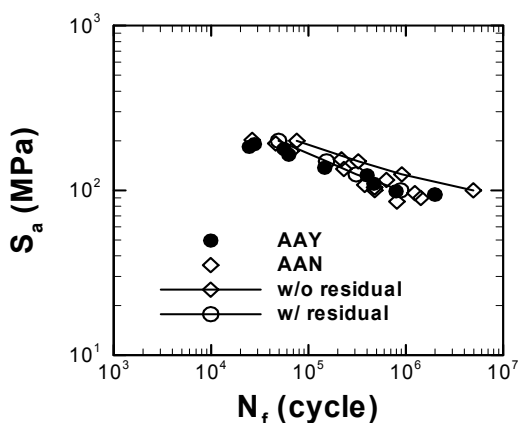


Fig. 32. Butt weld

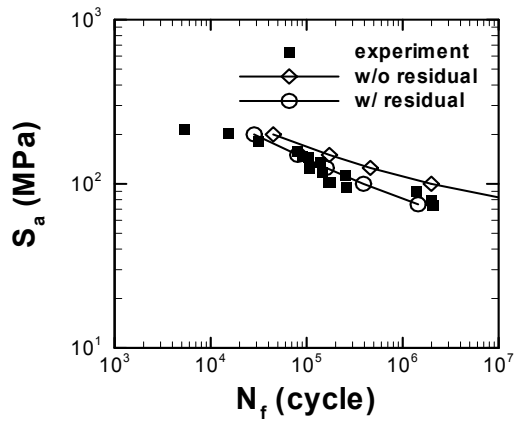


Fig. 33. T-shape fillet weld

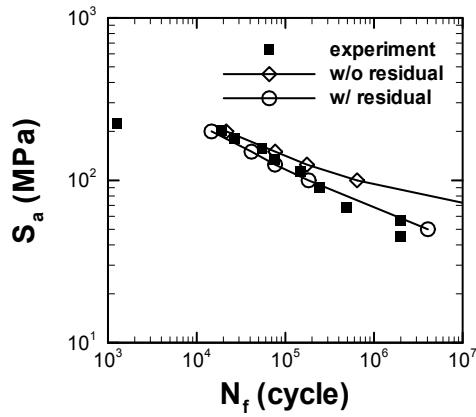


Fig. 34. Non load carrying cruciform

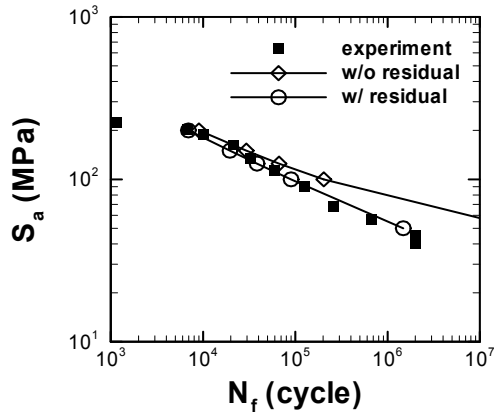


Fig. 35. Load carrying cruciform

5. Conclusions

Residual stresses play an important role in the behavior of welded components and structures. How to deal with them has been one of the challenging problems for engineers and researchers. One of the useful methods to relieve the residual stresses is to apply mechanical loading. When a cyclic zero-maximum mechanical load ($R=0$) is applied to a specimen with initial tensile residual stresses, the tensile residual stresses are relaxed if the sum of the applied load and initial residual stresses is big enough to produce yielding of the material. The tensile residual stresses may even change to compressive residual stresses. When a fully reversal load ($R=-1$) is applied to the same specimen, the effect of the residual stress relaxation is not significant. The applied compressive loads produced the tensile residual stresses. The process of residual stress relaxation can be simulated successfully by the finite element method.

The effect of welding residual stresses on fatigue behavior can be simulated too by the finite element method. We modeled the fatigue crack propagation under cyclic loading for two-dimensional compact tension specimens with and without welding residual stress. Initial welding residual stress was generated by conducting transient thermo elastic-plastic finite element analyses. Then a cohesive zone model with evolutionary damage law was applied to predict the fatigue crack growth under the residual stress. When compressive residual stress normal to the crack surface existed around the crack tip, the fatigue crack growth rate was slower. As the fatigue crack propagates, the welding residual stresses are redistributed. When estimating structural strength or fatigue life, residual stresses should be also considered. We proposed a parameter for fatigue life evaluation considering residual stress. According to the results, the locations of fatigue crack initiation depend on the magnitude of applied load, because residual stresses are redistributed, which change the weakest point under applied loading. The predicted results are in a better agreement with experimental results when we consider the welding residual stresses.

6. References

- Iida, K. and Takanashi, M., Welding in the World, **Vol.41**, 314 (1998).
- Yang, S. Y. and Goo, B. C. (2005). Key Engineering Materials, Vol. 297~300, p. 710.
- Goo, B. C., Yang, S. Y. and Seo, J. W. (2005). Key Engineering Materials, Vol. 297~300, p. 762.
- Song, H. C. (2001). Ph.D. thesis, Seoul National University, Korea.
- Needleman, A. (1987). J. Appl. Mech. Vol. 54, p. 525.
- Rose, J. H., Ferrante, J. and Smith, J. R. (1981). Phy. Rev. Let., Vol. 47, p. 675.
- De-Andres, A., Perez, J. L. and Ortiz, M. (1999). Int. J. Sol. Struct., Vol. 36, p. 2231.
- Roe, K. L. and Siegmund, T. (2003). Eng. Fract. Mech., Vol. 70, p. 209.
- Anderson, T. L. (1995). Fracture Mechanics, CRC Press, U.S.A.
- Reemsnyder, H. (1981). Experimentation and Design in Fatigue, pp. 273-295.
- Lawrence, F. V., Burk, J. D. and Yung, J. Y. (1982). ASTM STP 776, pp. 33-43.
- Vormwald, M. and Seeger, T. (1987). Residual stresses in Science and Technology, pp. 743-750.
- Stephens, R. I., Fatemi, A., Stephens, R. R. and Fuchs, H. O. (2001). Metal Fatigue in Engineering, John Wiley & Sons, Inc., pp. 196-205.
- Davis, J. R. (1990). Materials Handbook, 10th edition, Vol. 1, American Society for Metals, U.S.A.

Dynamic Finite Element Analysis on Underlay Microstructure of Cu/low-k Wafer during Wirebonding

Hsiang-Chen Hsu¹, Chin-Yuan Hu¹, Wei-Yao Chang¹,
Chang-Lin Yeh² and Yi-Shao Lai²

1 I-Shou University, Taiwan

2 Advanced Semiconductor Engineering, Taiwan

1. Abstract

The aim of present research is to investigate dynamic stress analysis for microstructure of Cu/Low-K wafer subjected to wirebonding predicted by finite element software ANSYS/LS-DYNA. Two major analyses are conducted in the present research. In the first, the characteristic of heat affected zone (HAZ) and free air ball (FAB) on ultra thin Au wire have been carefully experimental measured. Secondary, the dynamic response on Al pad/beneath the pad of Cu/low-K wafer during wirebonding process has been successfully predicted by finite element analysis (FEA).

Tensile mechanical properties of ultra thin wire before/after electric flame-off (EFO) process have been investigated by self-design pull test fixture. The experimental obtained hardening value has significantly influence on localize stressed area on Al pad. This would result in Al pad squeezing around the smashed FAB during impact stage and the consequent thermosonic vibration stage. Microstructure of FAB and HAZ are also carefully measured by micro/nano indentation instruments. All the measured data serves as material inputs for the FEA explicit software ANSYS/LS-DYNA. Because the crack of low-k layer and delamination of copper via are observed, dynamic transient analysis is performed to inspect the overall stress/strain distributions on the microstructure of Cu/low-k wafer. Special emphasizes are focused on the copper via layout and optimal design of Cu/low-k microstructure. It is also shown that the Al pad can be replaced by Al-Cu alloy pad or Cu pad to avoid large deformation on pad and cracking beneath the surface. A series of comprehensive experimental works and FEA predictions have been performed to increase bondability and reliability in this study.

2. Introduction

Although several advanced interconnection techniques have been developed recently, wirebonding process has been the most popular interconnection technology in integrated circuit (IC) devices packaging for its easy application and low cost for the past two decades. Ultrasonic thermosonic bonding (T/S bonding) technique has shown better reliability and

good interconnection among all wire bonding processes. The 4N (99.99%) ultra thin gold wire (0.8~1mil diameter) has been implemented from 0.18 mm process to 90 nm process in CMOS IC. In the 65 nm process, the copper interconnection was encapsulated by low permittivity intermetal dielectric (IMD) materials which easily soften microstructure of Cu/Low-K layers. Previous researches experimental illustrated failure modes of interfacial delamination and stress-induced void within the Cu/Low-K IMD layers. Because the complete mechanism of wirebonding process includes downward z-motion, impact and ultrasonic vibration stages, many material properties of gold wire were scarcely realized. Wirebonding process is therefore essential difficult to simulate by numerical analysis. Few papers published the reliability of wirebonding process by FEA prediction. However, some of the bonding material data are numerical assumptions. These difficulties can be overcome and will be fully explored as follow.

Due to thermal effect of EFO, bonding wire usually divided in three different zones, namely FAB, HAZ and as-drawn wire. The material behaviour of three zones in gold wire has been fully investigated by many earlier works [1-5]. However, elastic modulus (E) and Poisson's ratio (ν) at leveled working temperature for FAB and HAZ in Au wire are still scarcely known. The micro-Vickers indentation test was applied to obtain the Vickers hardness (HV) and then transfer to ultimate tensile stress (UTS). Nanoindentation test was also conducted to evaluate the surface elastic mechanical property around FAB. The bonding temperature (150-175°C) effect is taken into account in the tensile mechanical property. Thermo-tensile mechanical attributes for HAZ were measured by self-design pull test fixture. Micro interfacial friction at tiny surface is a major factor to reduce Al pad squeezing during impact and thermosonic vibration stages. The coefficient of frictional force (μ) at a tiny surface level was measured by Atomic Force Microscopy (AFM). It should be noted that no time-dependent materials are used because of fast wirebonding process. With sufficient experimental material properties obtained in this research, both 2-D and 3-D FEA models were developed to predict the dynamic response of wirebonding process.

The geometry of the overall structure included the entire underlay microstructure of Cu/Low-K IMD layer for FEA solid model is well defined. Wirebonding process predicted by FEA is simulated in three steps. In the first step, the capillary (also refer to the "tool") push FAB downward 10 μm within 0.7 ms to touch the pad. Secondary, the tool is continuous push FAB impact pad and the contact interface length became welded. The third step provides a slightly downward force and ultrasonic vibration, which refers to 120 KHz frequency, 1 μm amplitude within 4 ms vibration time. Loading procedure for the tool is: (1) total traveling time is 4.79 ms (2) the vertical displacement is 12.45 μm (3) the horizontal displacement is 2 μm . In addition, the final shape of FAB changes significantly and the small contact region between tool and FAB is always collapsed due to large plastic deformation. Thus, FEA model needs to be re-meshed very frequently during iteration approach. A fine mesh scheme is required to evaluate these plastic deformations with a sufficient accuracy.

3. Experimental Works

EFO for ultra thin Au wire was conducted by K&S 1488 plus wire bonder. The melting temperature for Au wire used in this research is found to be 1064°C, which brought into higher current and lower voltage. Only the tip of wire instantaneously jumps to melting temperature and becomes round-shape ball (FAB) due to gravity. Above FAB zone, a small

segment of wire (HAZ) is found to be highly affected by temperature. The working temperature is sharply decreasing from melting to ambient temperature which results in recrystallization of this region.

3.1 Mechanical Tensile Test

Based on ASTM standard F219-96, Nippon Micrometal 4N Au wire (diameter=25.4 μm) was performed for tensile pull test. Samples were prepared in 10mm gauge length, wire pull speed is controlled in 1mm/min and the yield strength is measured at 1% of total elongation. All data were repeatedly measured at least three times to receive consistent results. The Instron-3365 universal test system with $5\text{N}\pm 0.5\%$ load cell is used. Thermal effects (25, 125 and 200°C) on material properties are taken into account. It has been reported that the plastic behavior and fracture will occur in the HAZ area when external loading is applied. It is therefore clear that the breakage sites of EFO wire are in the neck somewhere between HAZ and FAB. Fig. 1 schematic illustrates sample preparation for as-drawn wire and EFO wire. Fig.2 shows the self-design pull test fixtures for EFO wire (HAZ).

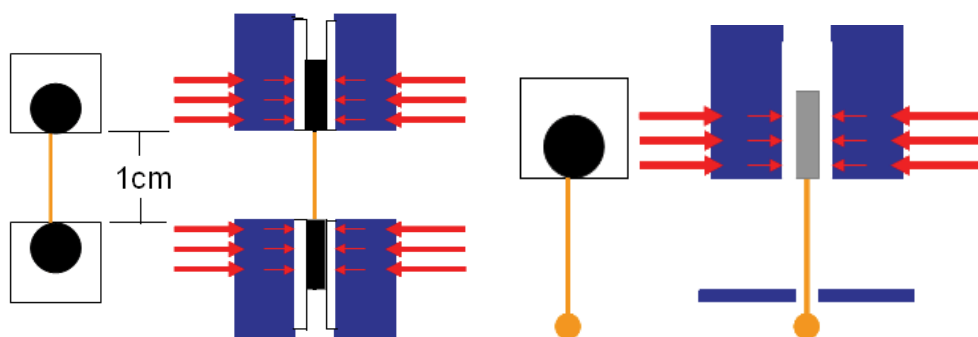


Fig. 1. Samples preparation (a) as-drawn wire (b) EFO wire (HAZ)

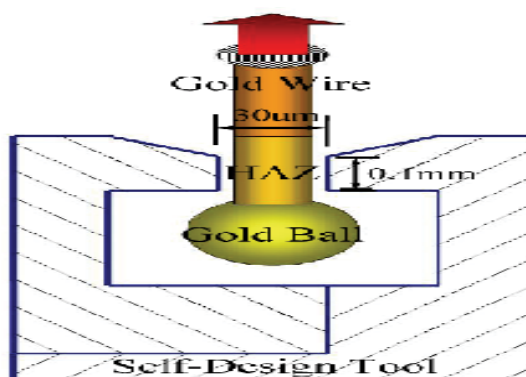


Fig. 2. Self-design tensile test fixtures for EFO wire (HAZ).

In the real-world wirebonding process, the impact displacement is 7.62 μm and the impact time is 1 ms. The contact speed (C/V) is defined as

$$C/V = (\text{impact displacement} / \text{impact time}) = 0.3\text{mil} / 1\text{ms} = 7.62\text{mm/s} \quad (1)$$

The pulling speed (V_s) in wire tensile test is therefore needed to be modified as

$$V_s = C/V = 7.62\text{mm/s} = 457.2\text{mm/min} \quad (2)$$

Because the pulling speed (7.62 mm/s) in wirebonding is much faster than the original tensile test speed (1 mm/s), strain-rate effects should be taken into account for the measured mechanical properties.

Engineering strain-rate ($\dot{\epsilon}$) is defined as

$$\dot{\epsilon} = \frac{de}{dt} = \frac{d(\frac{l_i - l_o}{l_o})}{dt} = \frac{1}{l_o} \frac{dl}{dt} = \frac{V_s}{l_o} \quad (3)$$

where e is engineering strain, l_o is gauge length and l_i is final length.

True strain-rate ($\dot{\epsilon}$) is defined as

$$\dot{\epsilon} = \frac{d\epsilon}{dt} = \frac{d(\ln(\frac{l_i}{l_o}))}{dt} = \frac{1}{l_i} \frac{dl}{dt} = \frac{V_s}{l_i} = \frac{l_o}{l_i} \dot{\epsilon} = \frac{1}{1+e} \dot{\epsilon} = \frac{\frac{V_s}{l_o}}{1+e} \quad (4)$$

where ϵ is the true strain. In this research, true strain-rate based on equation (4) is found to be $\dot{\epsilon} = 10^{-2}/s$ which implies the problem domain can be treated as quasi-static equation, i.e. inertial effect is neglected in the finite element prediction. For comparison, the strain-rate based on ASTM standard F219-96 can be calculated as $\dot{\epsilon} = 10^{-4}/s$.

The relationship between true stress (σ_T) and true strain (ϵ)/true strain-rate ($\dot{\epsilon}$) can be expressed by Power Law

$$\sigma_T = K(\epsilon)^n (\dot{\epsilon})^m \quad (5)$$

where K is the strength coefficient constant and m is the strain-rate sensitivity which is related to strain and temperature, i.e.

$$m = \left(\frac{\partial \ln \sigma_T}{\partial \ln \dot{\epsilon}} \right)_{\epsilon, t} = \frac{\dot{\epsilon}}{\sigma} \left(\frac{\partial \sigma_T}{\partial \dot{\epsilon}} \right)_{\epsilon, t} = \frac{\Delta \log \sigma}{\Delta \log \dot{\epsilon}} = \frac{\log(\sigma_T)_{t2} - \log(\sigma_T)_{t1}}{\log \dot{\epsilon}_2 - \log \dot{\epsilon}_1} = \frac{\log((\sigma_T)_{t2}/(\sigma_T)_{t1})}{\log(\dot{\epsilon}_2/\dot{\epsilon}_1)} \quad (6)$$

n is the strain harden exponent, i.e.

$$n = \frac{d(\log \sigma_T)}{d(\log \epsilon)} = \frac{\log(\sigma_T)_2 - \log(\sigma_T)_1}{\log(\epsilon_2) - \log(\epsilon_1)} = \frac{\epsilon}{\sigma_T} \frac{d\sigma_T}{d\epsilon} \quad (7)$$

Preliminary result shown that the strain harden exponent in equation (7) is a major factor for Au EFO wire to affect both large plastic deformation on FAB and squeezing on Al pad. In order to correlate engineering strain and true strain after necking, the Cowper-Symonds constituent is applied to evaluate the sensitivity of strain-rate.

$$\frac{\sigma_d}{\sigma_s} = 1 + \left(\frac{\dot{\epsilon}}{C} \right)^{\frac{1}{P}} \quad (8)$$

where σ_d is the dynamic flow stress, σ_s is the static flow stress which is defined as the average of ultimate tensile stress (UTS) and yield stress (σ_Y), C and P are the correlation constants. Mechanical tensile tests were performed according to low strain-rate $\dot{\epsilon} = 10^{-4} / s$ and high strain rate $\dot{\epsilon} = 10^{-2} / s$. Sensitivity constants C and P can then be determined.

True stress-true strain curves obtained from wire pull test at $\dot{\epsilon} = 10^{-2} / s$ for Au as-drawn wire and Au EFO wire are demonstrated in Fig. 3 and Fig. 4, respectively. Material properties in equation (5) to (8) are carefully evaluated and listed in Table 1.

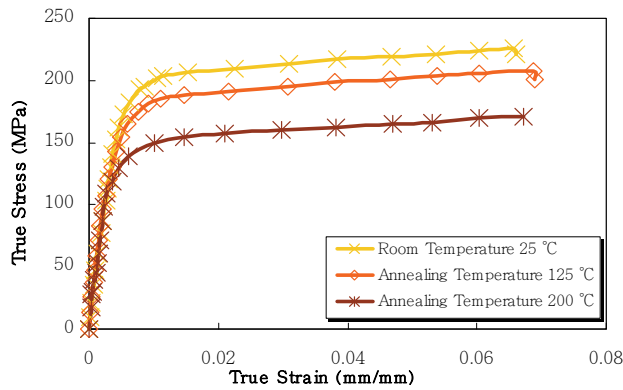


Fig. 3. Tensile mechanical properties for as-drawn wire (inclusion of thermal effects).

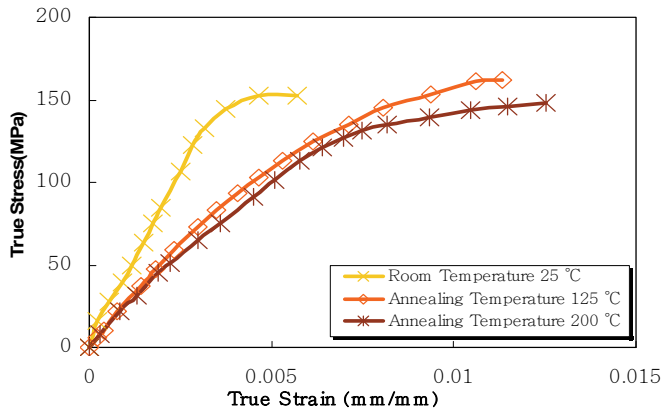


Fig. 4. Tensile mechanical properties for EFO wire (inclusion of thermal effects).

Material Property	E	σ_Y	UTS	n	K	m	C	P
Unit	Gpa	Mpa	Mpa		Mpa			
As-drawn wire 25°C	32	230	249	0.078	291	0.056	540	2.35
As-drawn wire 125°C	27.8	221	240	0.076	287	0.052	578	2.27
As-drawn wire 200°C	25.8	203	232	0.057	266	0.061	503	2.04
EFO wire 25°C -HAZ	14.4	91	97					
EFO wire 125°C -HAZ	13.8	82	84					
EFO wire 200°C -HAZ	12.24	70	75					

Table 1. Mechanical material property for Au as-drawn wire and Au EFO wire ($\dot{\epsilon} = 10^{-2} / s$).

3.2 Micro-Vickers Hardness Test

The Akashi MVK-H11 micro-Vickers hardness tester with a loading of 10 gf, a loading time of 5 sec. and a loading speed of 100 $\mu m/s$ was used to examine the hardness of FAB, HAZ and as-drawn wire on Au wire. The diameter of FAB is controlled at $62 \pm 1.5 \mu m$ to investigate the hardness. Fig. 5 shows the micro-Vickers hardness indentations from optical microscopy (OM) picture. Fig. 6 illustrates the Vickers hardness of gold wire (the origin is the center of FAB and measured every 50 μm until 1000 μm). The length of HAZ was discovered around 80-90 μm and the average of micro Vickers hardness for FAB, HAZ and as-drawn wire region are 41.7, 36.8 and 56 Hv, respectively.

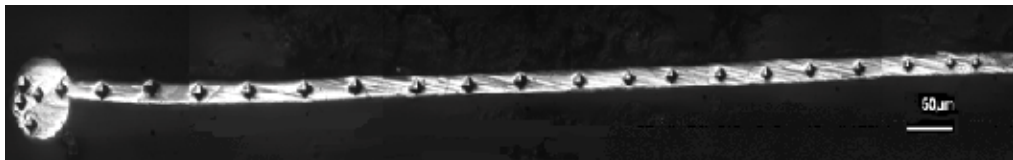


Fig. 5. Micro-Vickers hardness indentation (OM picture).

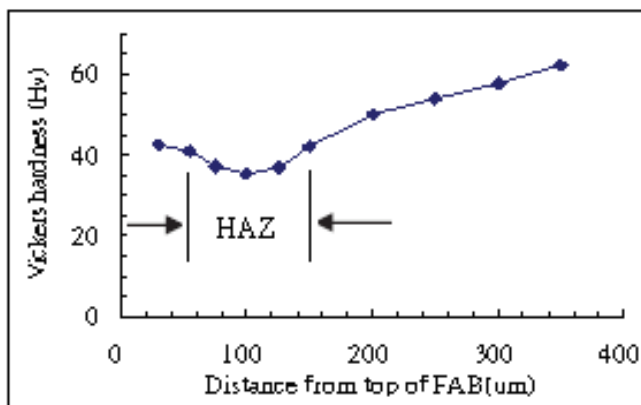


Fig. 6. Micro-Vickers hardness of FAB, HAZ and as-drawn wire region.

Because the hardness in horizontal x-direction plays an important role during ultrasonic vibration stage, the micro-Vickers hardness in FAB is also measured before/after vibration. Fig. 7 (a) schematically illustrates the micro hardness vertical measuring system and Fig. 7 (b) shows the corresponding hardness of FAB before and after ultrasonic vibration stage. The FAB hardness on both side dramatically increased after ultrasonic vibration. The influence of FAB hardness on large plastic deformation would be very useful for studying the Al pad squeezing during wirebonding process.

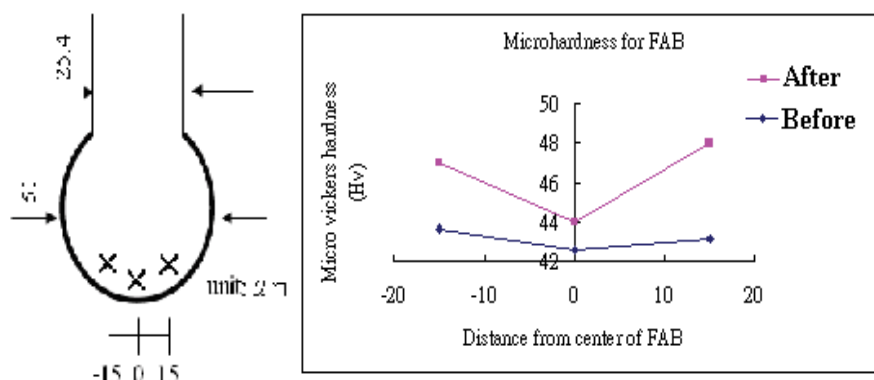


Fig. 7. (a)Micro hardness vertical measuring system (b)corresponding micro-Vickers hardness.

3.3 Nano-indentation Test

Because the impact and ultrasonic vibration stages of first bond occurred at the very surface level, tensile mechanical test and micro-Vickers hardness are not capable to illustrate the surface mechanical property on FAB. The MTS Nano Indenter XP System is available to measure continuous stiffness on the thin film of material. Fig. 8 shows loading and unloading curves for nanoindentation test on specific point 1-5 on the surface of FAB. Thermal effects after annealing were also studied. Table 2 lists the corresponding FAB surface elastic modulul for EFO wire.

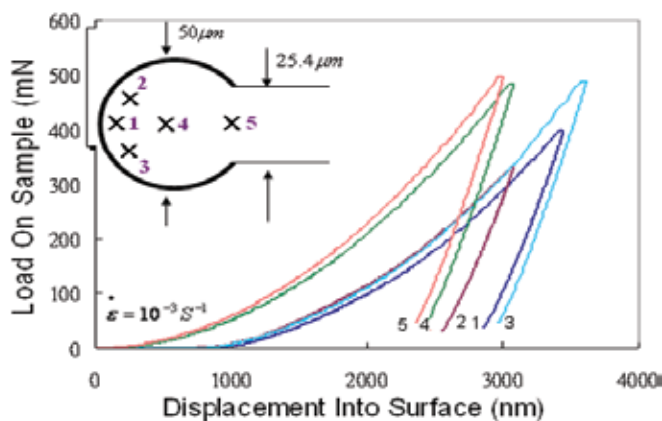


Fig. 8. Nanoindentation loading/unloading curves on FAB surface for EFO wire (25°C).

	Surface Elastic Modulus (GPa)		
	25 °C	Annealing 125 °C	Annealing 200 °C
Point 1	50.1	47.9	42.6
Point 2	52.4	48.1	42.8
Point 3	52.1	48.3	43.3
Point 4	55.1	52.4	50.4
Point 5	59.2	54.2	51.5

Table 2. Elastic modulus (GPa) measured by nano-indentation on FAB surface.

3.4 Micro-Interfacial Frictional Force Measurement

It has been reported [5-6] that surface frictional force between FAB and bond pad is a major factor to reduce pad shoveling around the smashed ball during thermosonic vibration stages. Traditional instruments are not available to measure the coefficient of frictional force (μ) at a tiny surface level. This can be overcome by using Atomic Force Microscopy (AFM) series - NT-MDT Solver P47. The measurement of the normal force (F_n) between probe and sample can be utilized the so-call "force-distance curve" measured by AFM. Then, F_n is defined as

$$F_n = k_n D \quad (9)$$

where k_n is the normal modulus of probe and D is the distance between setpoint and release contact point. By selecting a thin rectangle probe made by silicon, k_n can be determined as

$$k_n = \frac{Ewt^3}{4l^3} \quad (10)$$

where E is the elastic modulus of silicon, w is the width, t is the thickness and l is the length of probe. The measured frictional force (f) is defined as

$$f = \frac{l}{h} S_z V \frac{k_n}{1 + \nu} \quad (11)$$

where h is the height, S_z is the sensitivity of detector, V is the differential voltage and ν is the Poisson's ratio of silicon probe. The measured coefficient of micro-interfacial frictional force is given as

$$\mu = \frac{f}{F_n} \quad (12)$$

It should be noted that equation (12) is the coefficient of sample and silicon probe. By using the relative coefficient method developed by Oak Ridge National Laboratory, the coefficient of frictional force in equation (12) can be related to sample and probe

$$\mu = \sqrt{\mu_S^2 + \mu_P^2} \quad (13)$$

where μ_S is the coefficient of sample and μ_P is coefficient of probe. The silicon probe used in this research is NANOWORLD SSS-SEIHR-10, the coefficient of frictional force is given as $\mu_P = 0.12$. The dimension of the probe is given as: $l=225\mu m$, $w=33\mu m$, $h=10\mu m$ and $t=5\mu m$. Material property is given as: elastic modulus is 112.4 Gpa and Poisson's ratio is 0.28.

Fig. 9 illustrates a 3D scanned picture at speed 60.72 $\mu m/s$ and time 36 sec by AFM. Fig. 10 demonstrates the probing distance at 8.28 nm and 2 volt. The measured coefficient of frictional force in equation (11) for Au EFO ball is 0.24, and for Al pad is 0.17. From equation (13), the coefficient of frictional force for Au EFO ball is 0.2, and for Al pad is 0.14, respectively. Thus, the coefficient of micro-interfacial frictional force between Au EFO ball and Al pad is 0.244, which is easily applied to FEA.

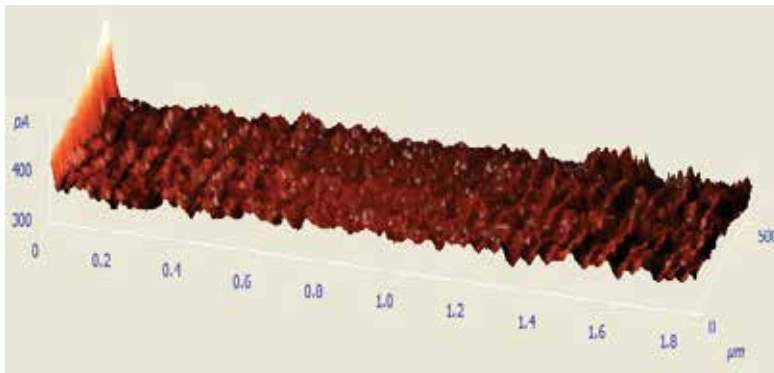


Fig. 9. A scanned 3D AFM picture for Al bond pad.

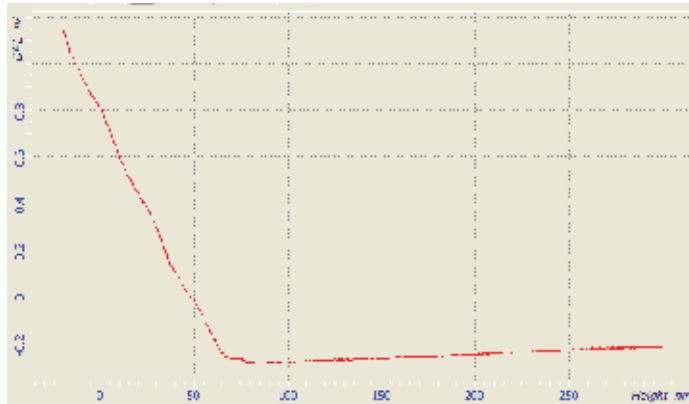


Fig. 10. AFM probing distance at $D=8.28\text{nm}$ and $V=2$ volt.

3.5 Microstructure of Cu/low-k Layer

The SEM picture for bonded FAB and the microstructure of Cu/low-K layer is demonstrated in Fig. 11. The underlay horizontal metal layers and vertical via arrays through the dielectric are designed to transmit the downward bonding force. The IMD material between horizontal layers is designed to absorb the bonding impact, which would

result in bending moment occurred at the junction of metal layer and vias. Some failure modes due to excessive bending tensile stress were observed during wirebonding process.

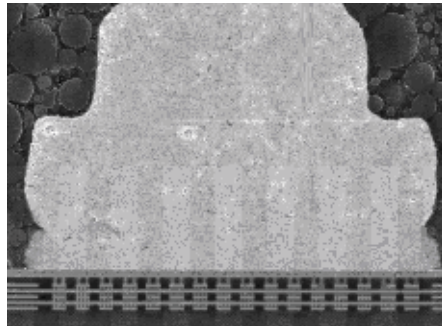


Fig. 11. Microstructure of bonded FAB and Cu/low-K layer.

4. Finite Element Prediction

Although the problem domain can be treated as a quasi-static problem, the explicit scheme is suggested to solve highly nonlinear large deformation on FAB and squeezing on pad. The time increment (Δt) is limited to micro second scale to improve the stability for numerical iterations in this research. Finite element commercial software ANSYS based on implicit method is first applied to define the material, solid model, mesh and boundary conditions. The core explicit scheme solver is LS-DYNA. For post processing, LS_Prepost 6 and Post 26 are applied to deal with time-history deformation, overall strain and stress distribution. Element type for analysis is plane 162 for 2-D model and solid 64 for 3-D model.

4.1 Modeling

Both 2-D and 3-D finite element models based on ANSYS/LS-DYNA codes are developed to simulate the wirebonding process. Since this research is focused on the stress beneath the bond pad, the capillary, bonded FAB and entire microstructure of Cu/Low-K layer should be well defined. Fig. 12 presents the finite element 2-D model and Fig. 13 illustrates the details of underlay microstructure of Cu/low-k layer.

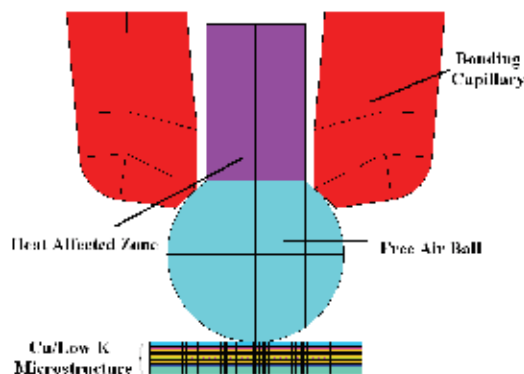


Fig. 12. Finite element 2-D wirebonding model.

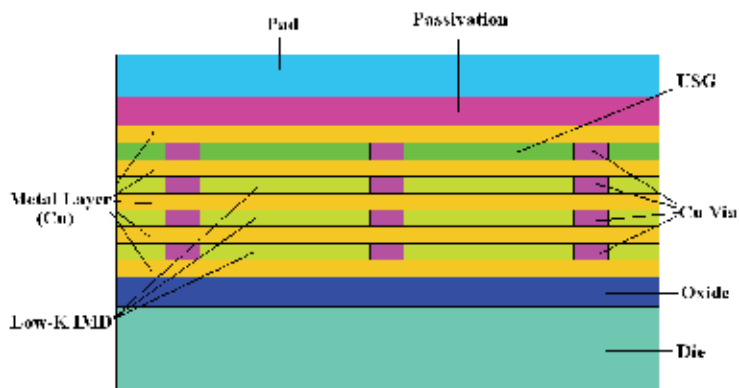


Fig. 13. Details of underlay microstructure of Cu/low-k layer.

Fig. 14 demonstrates the 3-D finite element model and Fig. 15 depicts the Cu via array in underlay microstructure. Detail dimension for underlay microstructure is listed in Table 3.

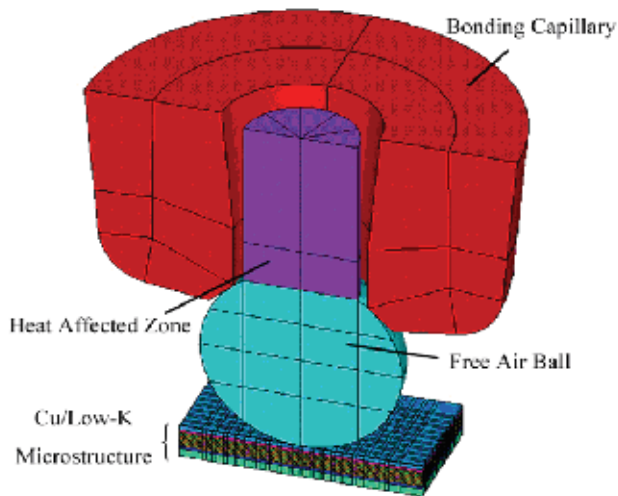


Fig. 14. Finite element 3-D wirebonding model.

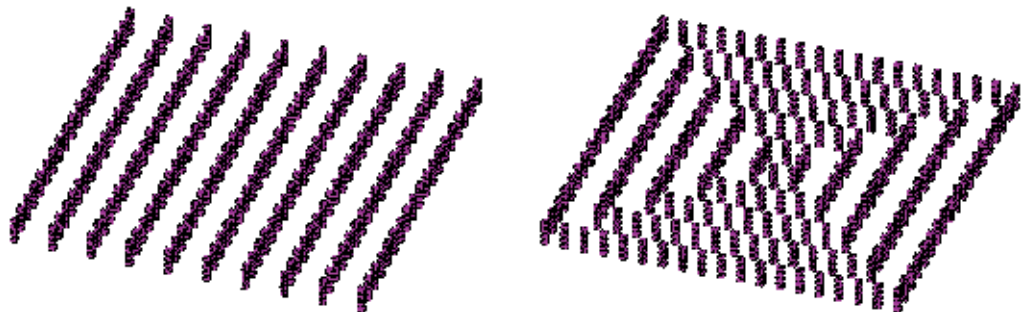


Fig. 15. Underlay Cu via (a) parallel array (b) loop array.

	Dimension
HAZ Diameter	25.4 μm
FAB Ball Diameter	50 μm
Cu/Low-k Layer	54×54×8 μm
via	0.4×0.4×0.4 μm
Al Pad Thickness	1 μm
Passivation Thickness	0.7 μm
Metal Layer	0.4 μm
USG Layer	0.4 μm
Low-k Layer	0.4 μm
Oxidation layer	0.7 μm

Table 3. Dimension for underlay microstructure.

4.2 Material Property

The final smashed ball is formed by the bonding capillary which is made of high elastic modulus ceramic. Therefore, the capillary is treated as a rigid body in the model. In order to study the overall stress distribution in the entire structure, the elastic-plastic material behavior is defined for all the material used in pad and microstructure beneath pad. Fig. 16 plots bi-linear tensile mechanical property for materials used in this research. As can be seen, low-K IMD material has a low stiffness and a very low yield stress which leads to easily weak the overall structure. Beside these, the silicon die is treated as a perfect plastic material. Table 4 lists the material properties for Cu/low-K IMD in the underlay microstructure.

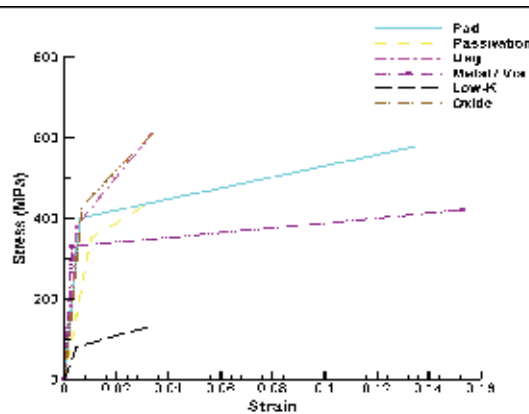


Fig. 16. Bi-linear material property for materials used in pad and underlay microstructure.

	ν	E (GPa)	$\rho(\text{g}/\text{cm}^3)$	σ_Y (MPa)
Pad	0.33	69	2.71	400
Passivation	0.24	32	1.31	350
USG	0.23	80	2	380
Copper Metal/ Via	0.38	121	8.91	330
Low-k IMD	0.30	18	2	80
Oxide	0.32	66	2.64	430
Die	0.23	161	2.33	—

Table 4. Material property for Cu/low-k IMD in the underlay microstructure

4.3 Meshing

Because the final shape of FAB changes significantly and a small contact region between FAB and capillary is always collapsed during iteration, a fine mesh scheme with sufficient accuracy is required to evaluate these plastic deformations. Fig. 17 demonstrates the mesh density for physical model at different number of element (a) 20440, (b) 41040, (c) 99998 and (d) 261488, respectively. Fig. 18 demonstrates the convergence of mesh density.

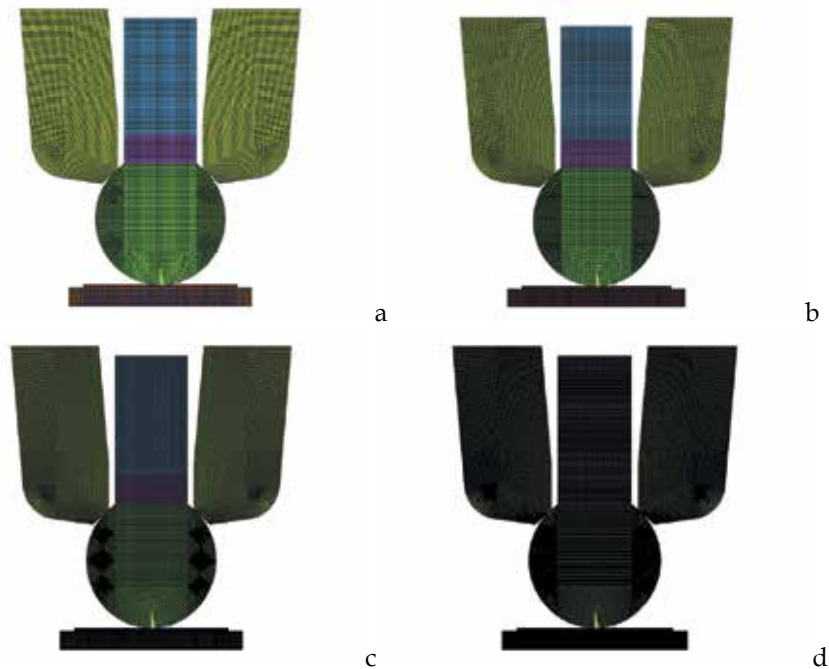


Fig. 17. Number of element for physical model (a) 20440 (b) 41040 (c) 99998 (d) 261488.

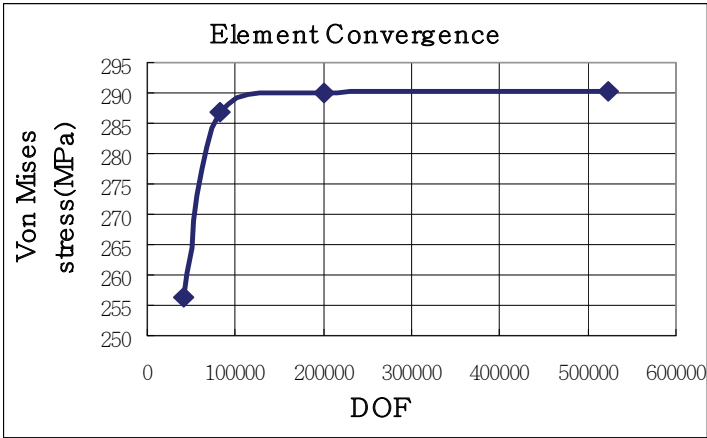


Fig. 18. Convergence of mesh density.

4.4 Boundary Conditions

Because the strain/stress in die is excluded and the tool is treated as a rigid body, boundary conditions for finite element model are illustrated in Fig. 19.

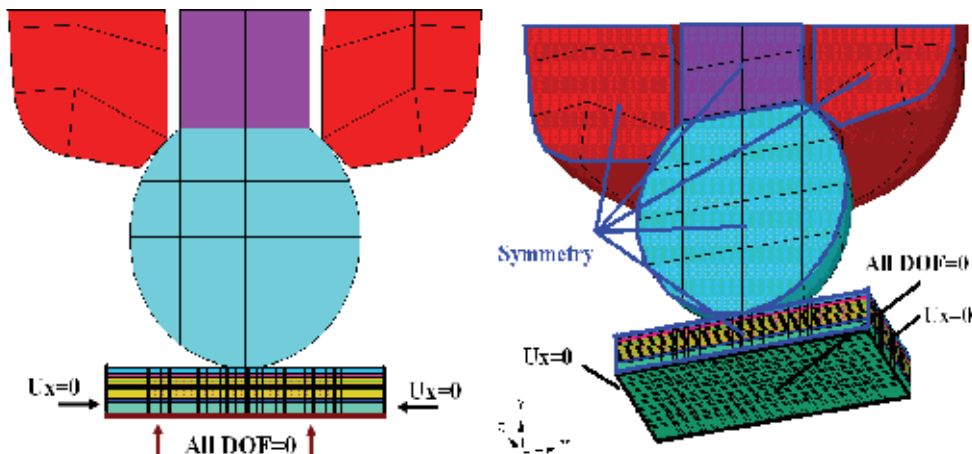


Fig. 19. Boundary conditions for (a) 2-D model (b) 3-D model.

The difficulty encountered in contact analysis is a highly non-linear behavior in any numerical approach. Especially, the inclusion of interfacial frictional force would easily result in a divergence during iterations. Contact elements in ANSYS/LS-DYNA include a pair of target (master) element and contact (slave) element. Automatic single side contact element (ASS2D) is applied to 2-D analysis and surface-to-surface contact element (STS) is applied to 3-D analysis. In this research, two contact pairs are needed to be carefully defined, i.e. (1) the region between capillary and the top rim of FAB, (2) the region between the bottom rim of FAB and the upper edge of pad. This can be seen in Fig. 20.

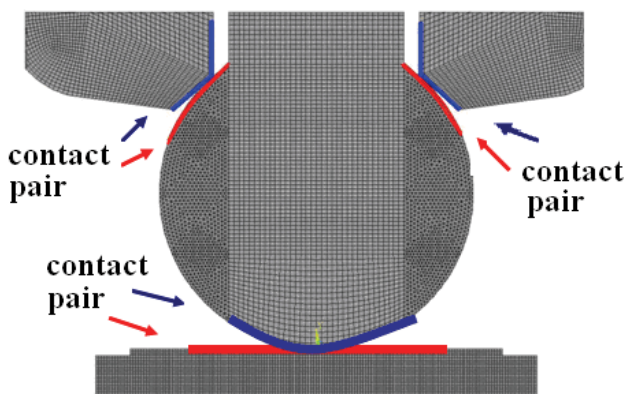


Fig. 20. Contact pairs in 2-D finite element model.

4.4 Loading

Fig. 21 schematically illustrates the complete mechanism of wirebonding process includes three stages, namely (1) Z-motion (vertical downward), (2) impact (welding) stage, and (3)

ultrasonic vibration stage, respectively. The loading sequence is (1) the contact speed is first defined as 0.5 mil/ms in Z-motion stage, (2) the capillary moves downward 10 μm within 0.79 ms, which simulates the bonding force in wirebonding process. Large plastic deformations on both sides of FAB are because the capillary continuously downward pushes the FAB. The contact face/length between FAB and pad became welded. (3) ultrasonic vibration provides 120 KHz frequency, 1 μm amplitude within 4 ms. The capillary travels downward 2 μm displacement to improve the bondability. Thus, the complete loading procedures in Fig.21 are: total traveling time is 4.79 ms, the vertical displacement is 12 μm and the horizontal displacement is 2 μm .

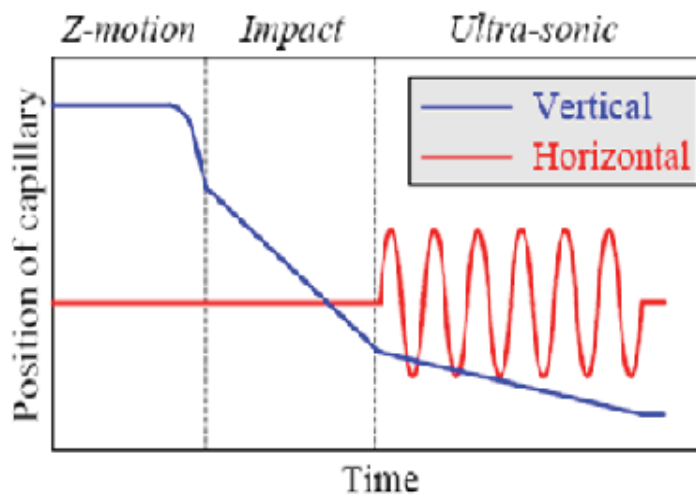


Fig. 21. Capillary position vs. bonding time in one cycle.

5. Parametric Study

A series of parametric study were performed to obtain the dynamic response on the underlay microstructure of Cu/low-k layer. Parametric studies were performed by (1) increase the number of via (2) change the height of via (3) enlarge the via (4) change the location of via (5) change amplitude of ultrasonic (6) increase the number of Cu metal layer (7) change the pattern of via in 3-D model.

6. Results and Discussion

As the shape of FAB is continuous formed by impact and ultrasonic vibration, the contact between FAB and pad provides a downward force, which implies only the compressive stress is transferred to passivation, USG and via. The stress in the metal layer could be tensile. Von Mises's stress in the microstructure once exceeds the yield stress will result in material failure in this region. Fig. 22 illustrates the equivalent stress wave propagation in one cycle (4.79ms). It reveals the stress propagates from contact FAB and pad to passivation, USG and Cu/low-k IMD through Cu via array. The predicted peak equivalent stress is 330.892 MPa in the Cu via and metal layer. It is of interesting to discover the peak stress was

originated in one of three stages. Fig. 23 demonstrates the equivalent stress wave propagation during impact stage (0.79 ms). It shows that the peak equivalent stress was generated at the end of impact stage. Fig. 24 illustrates the detail view of the equivalent stress distributions beneath the bond pad at 0.79 ms. It is reported that most of stressed via area is located in the center of underlay microstructure.

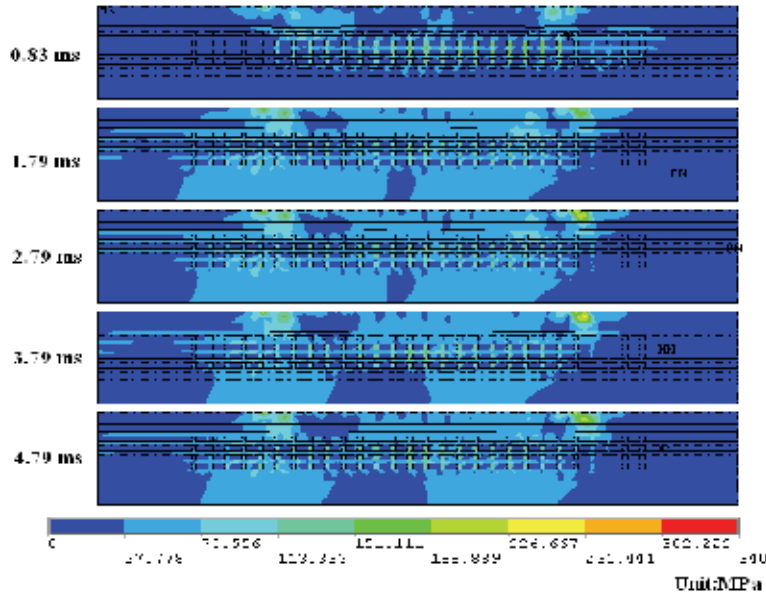


Fig. 22. Equivalent stress wave propagation in the underlay microstructure in one cycle (4.79ms).

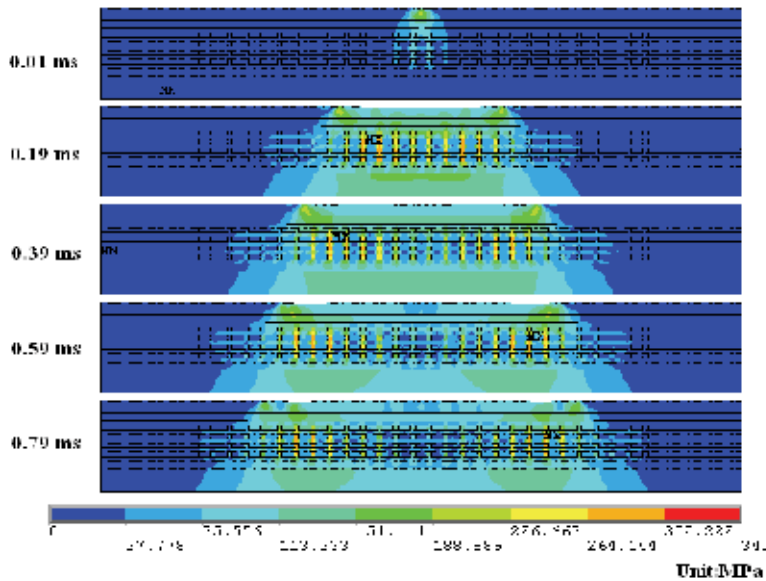
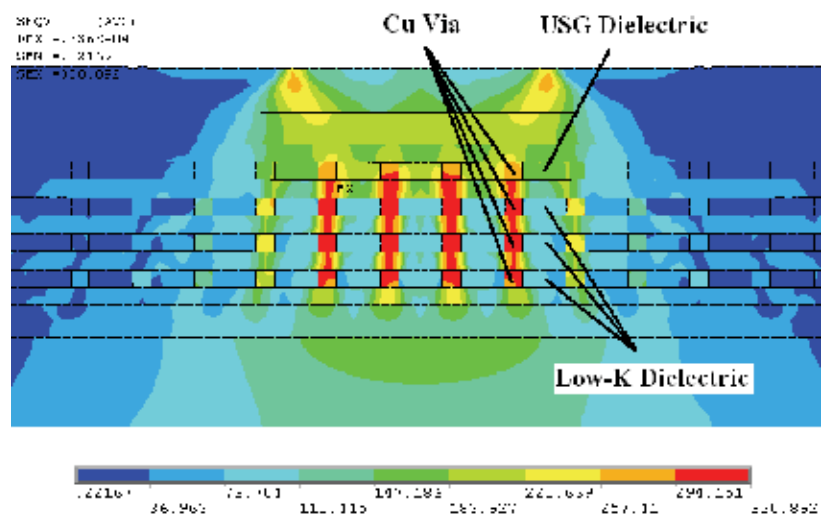


Fig. 23. Equivalent stress wave propagation in the underlay microstructure during impact stage (0.79ms).



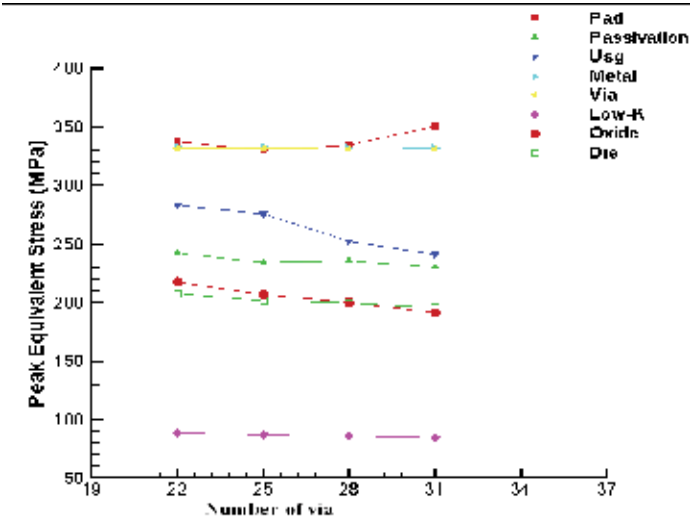


Fig. 25. Peak equivalent stress vs. number of via.

6.2 Height of via

The height of first via array was stretched from 0.3 μm to 0.4, 0.5 and 0.6 μm in Table 6. This would broaden the area of USG and the corresponding response was shown in Fig. 26. Predicted equivalent stress in USG was increased as the height was extended.

Height (μm)	Microstructure
0.3	
0.4	
0.5	
0.6	

Table 6. Expanding the height of via in microstructure.

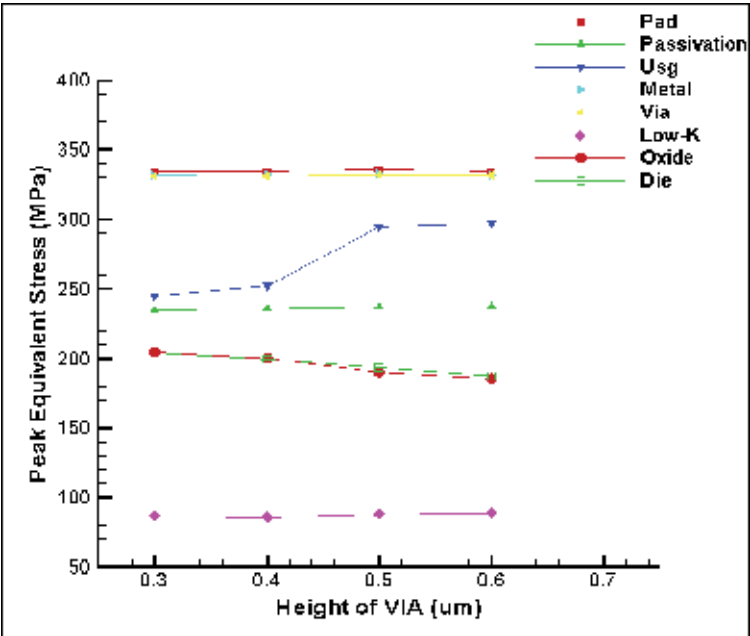


Fig. 26. Peak equivalent stress vs. via’s height.

6.3 Width of via

The width of via has been broadened from 0.3 μm to 0.4, 0.5 and 0.6 μm in Table 7. As the width of via is expanded, the volume of dielectric material is decreased, which dramatically reduces the equivalent stress in USG structure. Fig. 27 shows peak equivalent stress for all materials.

Width (μm)	Microstructure
0.3	
0.4	
0.5	
0.6	

Table 7. Broaden the width of via in the microstructure.

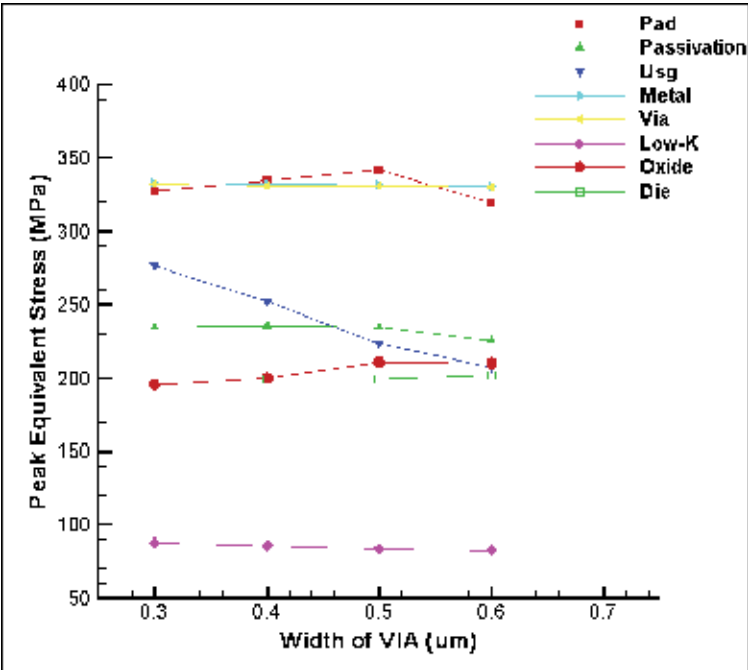


Fig. 27. Peak equivalent stress vs. via’s width.

6.4 Location of via

Distribution of via array has been designed the location of center, between and outsides shown in Table 8. As can be seen in Fig. 28, the predicted peak equivalent stress in via has been dramatically reduced when via array is located at outsides. This via structure is subjected to supporting the downward bonding force during impact stage.

Location	Microstructure
Center	
Between	
Outsides	

Table 8. Distribution of via array in the microstructure.

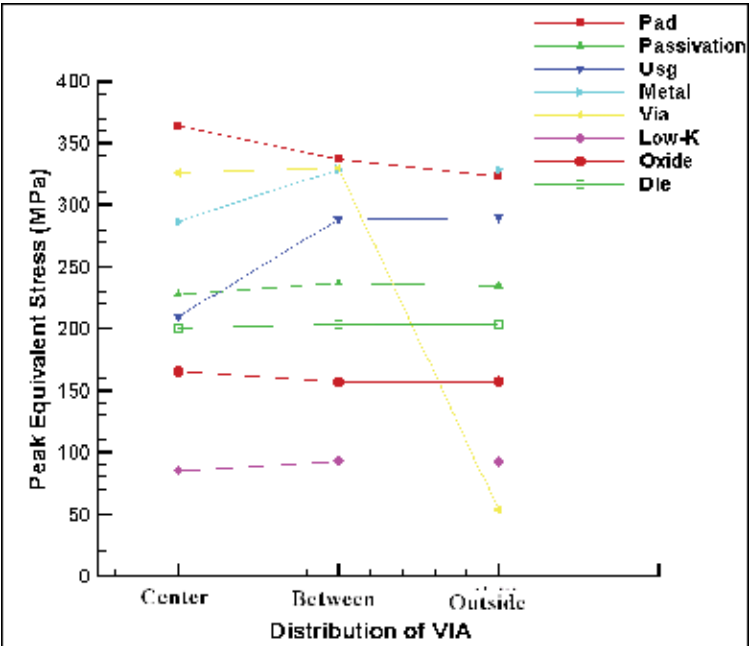


Fig. 28. Peak equivalent stress vs. distribution of via.

6.5 Amplitude of Ultrasonic

Because pad squeezing always observed during ultrasonic vibration stage, the amplitude of ultrasonic vibration in Table 9 is designed to study large plastic strain in pad. In Fig. 29, predicted plastic strain in pad for different amplitudes of ultrasonic vibration is 1.3% - 4.6%.

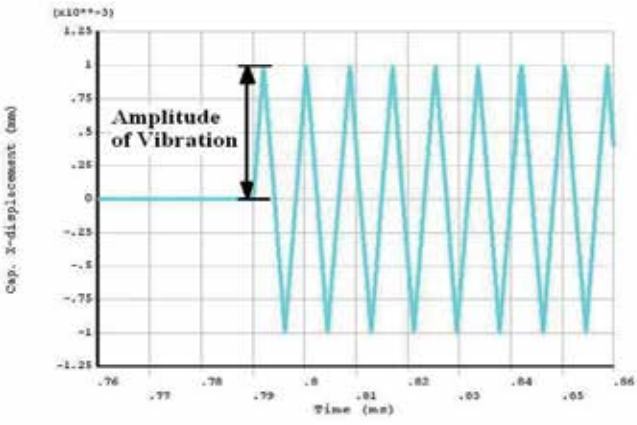
Amplitude (μm)	Microstructure
1.0	
1.2	
1.4	
1.6	

Table 9. Amplitude of ultrasonic vibration

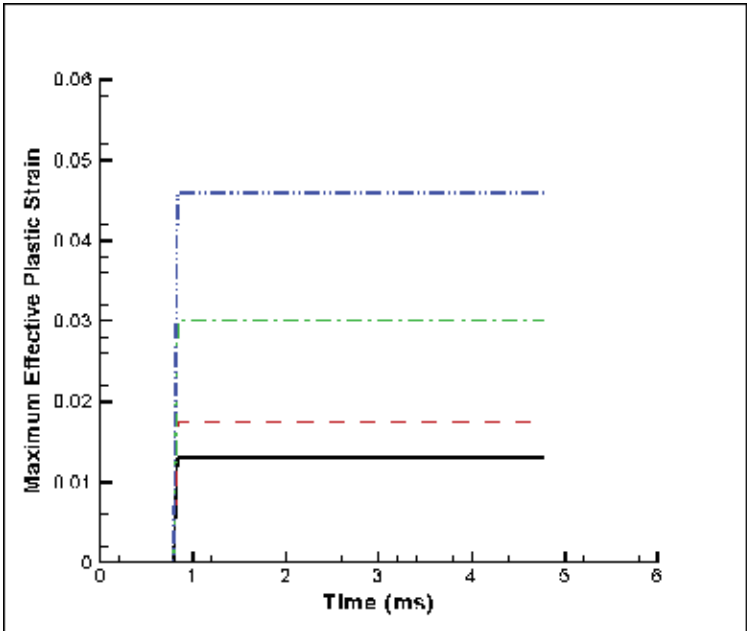


Fig. 29. Plastic strain in pad for different amplitudes of ultrasonic vibration

6.6 Number of Metal Layer

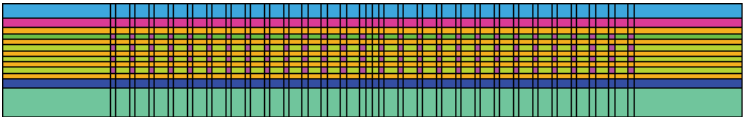
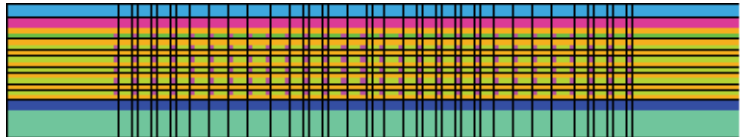
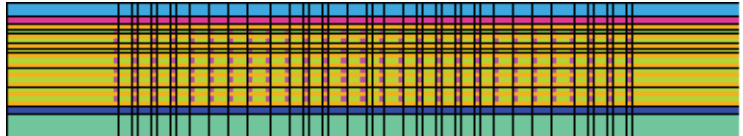
No. of Metal Layer	Microstructure
5	
7	
9	

Table 10. Number of metal layer in the microstructure.

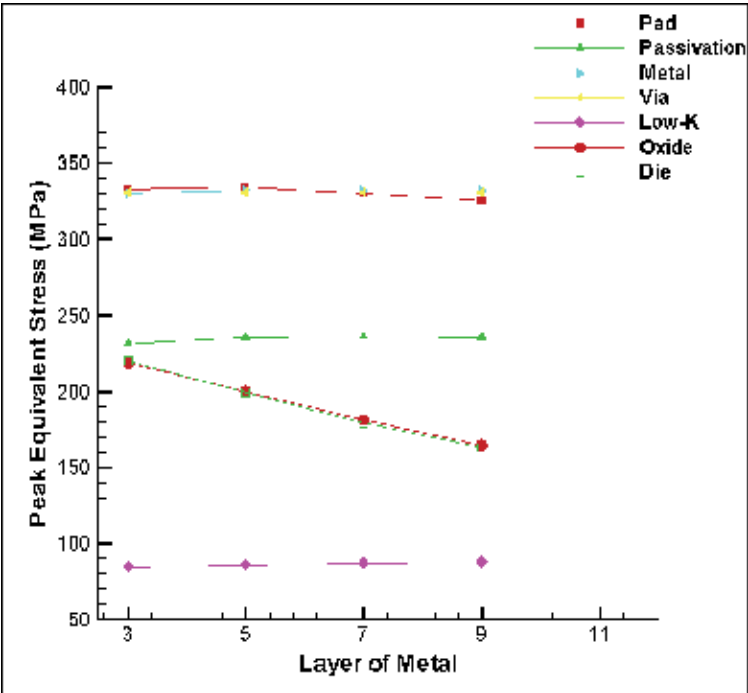


Fig. 30. Peak equivalent stress vs. number of metal layer.

6.7 Via Array in 3-D Model

Table 11 illustrates two type of 3D array for via. From Fig. 31, the loop array via has a better stress relief due to larger contact area under FAB and pad. A small peak equivalent stress in low-k IMD material was found for loop array via.

Via Array	Microstructure	
	Top View	ISO View
Loop		
Straight		

Table 11. Layout for 3D array via in the microstructure.

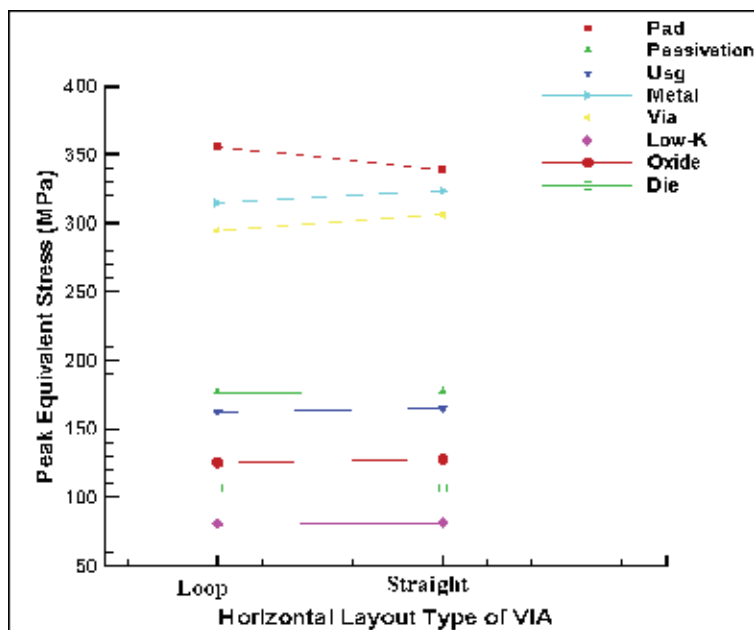


Fig. 31. Peak equivalent stress for different layout of via.

7. Conclusion

Although there still have some limitations in this study, material properties of EFO wire and underlay microstructure have been experimental determined to improve numerical analysis. Both 2D and 3D FE models based on ANSYS/LS-DYNA are developed to evaluate the dynamic responses (stress distribution and the failure mode) on the underlay microstructure of Cu/low-K wafer.

The complete mechanism of wirebonding process has been successfully predicted by FEA. During impact and ultrasonic vibration stages, stress wave propagation is found from top to down and from inside to outside.

Equivalent stress distributions are mostly concentrated on via area and it is found that peak equivalent stress in low-k area is larger than those in USG area.

An increase in the number of via will result in a decrease in the peak equivalent stress. An increase in the height of via will result in an increase in peak stress for USG material. An increase in the width of via will results in a decrease in the equivalent stress in USG structure.

For the distribution of via array, lower peak equivalent stress for all materials is observed in loop array.

For bondability study, large amplitude of ultrasonic vibration will induce higher stressed area in microstructure. Plastic strains in pad are increased as the amplitude is increased.

8. References

- Chang, W.Y., Hsu, H.C., Fu, S.L., Yeh, C.L. and Lai, Y.S. (2008) Characteristic of Heat Affected Zone for Ultra Thin Gold Wire/Copper Wire and Advanced Finite Element Wirebonding Model, *Proceeding of 10th Electronic Packaging and Technology Conference (EPTC 2008)*, pp.419-423, Singapore, December 2008.
- Degryse, D. and Vandevelde, B. (2004) Mechanical FEM Simulation of Bonding Process on Cu Low K Wafer, *IEEE Trans. on Component and Packaging Technology*, Vol. 27, 2004, pp.643-650.
- van Gils, M.A.J., van der Sluis, O., Zhang, G.Q., Janssen, J.H.J. and Voncken, R.M.J. (2005) Analysis of Cu/Low-K Bond Pad Delamination by Using a Novel Failure Index, *Proceeding of EuroSimE*, pp.190-196, 2005.
- Hsu, H.C., Chang, W.Y., Fu, S.L., Yeh, C.L. and Lai, Y.S. (2007) Dynamic Finite Element Analysis on Underlay Microstructure of Cu/low-K Wafer during Bonding Process, *Proceeding of 9th International Symposium on Electronic Materials and Packaging (EMAP 2007)*, pp. 231-234, Korea, November 2007.
- Kripesh, V., Sivakumar, M., Lim, L.A., Jumar, R. and Iyer, M.K. (2002) Wire Bonding Process Impact on Low-K Dielectric Material in Damascene Copper Integrated Circuits, *IEEE Electronic Component and Technology Conference*, pp. 873-880, 2002.
- Lim, Y.K., Lim, Y.H., Kamat, N.R., See, A., Lee, T.J. and Pey, K.L. (2005) Stress-induced Voiding beneath Vias with Wide Copper Metal Leads, *IEEE Physics and Failure Analysis Symposium*, pp. 103-106, 2005.
- Liu, Y., Irving, S. and Luk, T. (2004) Thermosonic Wire Bonding Process Simulation and Bond Pad over Active Stress Analysis, *Electronic Component and Technology Conference (ECTC 2004)*, pp.383-391, 2004.
- Wang, R.C.J., Lee, C.C., Chen, L.D., Wu, K. and Chang-Liao, K.S. (2006) .A Study of Cu/Low-k Stress-induced Voiding at Via Bottom and Its Microstructure Effect, *Microelectronic Reliability*, Vol.46, 2006, pp.1673-1678.
- Yeh, C.L. and Lai, Y.S. (2006) Comprehensive Dynamic Analysis of Wirebonding on Cu/Low-K Wafers, *IEEE Trans. on Advanced Packaging*, Vol. 29, 2006, pp.264-270.

Finite element analysis of deformation and fracture of cylindrical tubes under internal moving pressures

Majid Mirzaei

*Department of Mechanical Engineering, Tarbiat Modares University
Iran*

1. Introduction

This Chapter presents a review of the basic formulations, solution methods, and simulation procedures that are essential in the finite element analysis of deformation and fracture of cylindrical tubes under internal moving pressures. Such analyses are applicable to a number of theoretical and practical problems, like deformation and rupture of arteries due to blood flow, cyclic fracture of gas pipelines, vibrational behavior of pulse detonation engines, etc. However, the scope of this chapter is mostly concerned with *high-speed* moving pressures that can cause significant *fluctuating* stresses in the tube wall. Based on their relative magnitudes, these stresses can result in various types of mechanical failure such as: high cycle fatigue cracking (Mirzaei et al., 2006c), low cycle fatigue & dynamic tearing (Mirzaei, 2008a; 2009), and severe dynamic fragmentation (Price, 2006; Goto et al., 2008; Soto et al., 2010).

It should be emphasized that finite element simulations of the above mentioned processes are rather difficult and not only involve judicious implementation of advanced finite element (FE) techniques but also require a clear understanding of the nature of the underlying deformation and fracture mechanisms. Hence, this chapter starts with a brief review of the basics of deformation and fracture mechanisms that are specific to this type of problems. In continuation, the pertinent FE formulations and solution methods are reviewed and a treatise of the relevant fracture mechanics parameters is presented. The final sections of this chapter are devoted to practical examples including the FE analyses of deformation and fracture of an experimental aluminum tube and an exploded steel cylinder.

2. Structural Response of Tubes to High-speed Moving Pressures

The structural response of a cylindrical tube to internal moving pressure depends on several parameters; including the dimensions and mechanical properties of the tube as well as the magnitude and speed of the moving pressure. Classical examples include *high-speed* traveling pressure waves with specific profile caused by detonation of gaseous mixtures or traveling shock waves (see Fig.1).

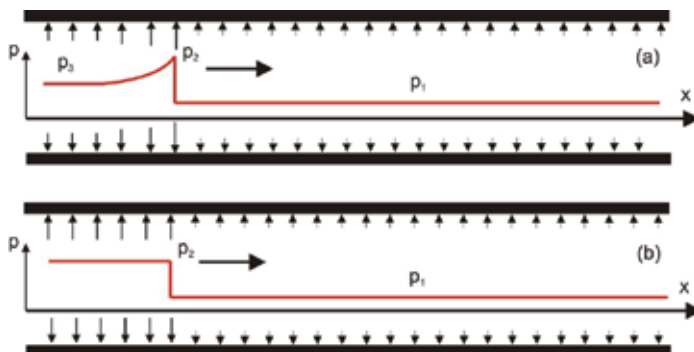


Fig. 1. Schematic of internal moving pressures a) gaseous detonation b) shock.

Grossly, we may define three pressure levels for which three different types of structural response can be expected. These pressures can be designated as low, medium, and high in relation with the magnitude of the resultant stresses which can be less than, equal to, or higher than the dynamic ultimate tensile strength of the material.

2.1 Low-Pressure Moving Front

For relatively *low pressures* (level 1) the passage of pressure front results in a pattern of fluctuating *elastic* strains which can exist even after the moving load dies out. For this case a variety of analytical, experimental, and FE solutions is available for the structural response (Tang, 1965; Reismann, 1965; Beltman et al., 1999; Beltman & Shepherd, 2002; Mirzaei et al., 2005, 2006a; Mirzaei, 2008b). An interesting application for this type of moving pressure is the pulse detonation engine (PDE) which is currently regarded as a promising candidate for providing very efficient propulsion systems for both subsonic and supersonic aviation. Since the PDE generates quasi steady thrust by high cycling of gaseous detonations, the interactions between the stress waves caused by successive detonations can be troublesome and cause high-cycle fatigue cracking in the tube wall.

2.2 Medium-Pressure Moving Front

Traveling of higher pressures (level 2), that can cause stresses in the order of the ultimate tensile strength of the material, may initiate axial cracks which with further propagation can result in *limited* fragmentation of the tube wall. There was no report in the open literature about confined crack growth in tubes under gaseous detonation, prior to the experimental study reported by Chao and Shepherd (2005). Their study was primarily motivated by accidents like those occurred in the nuclear power plants in Japan (Naitoh et al., 2003) and Germany in 2001. In both incidents, sections of steel steam pipes were fragmented due to combustion of hydrogen-oxygen mixtures created by radiolysis. One of the most important questions that arose during the accident investigation was whether the type of accidental combustion can be deduced from the fracture patterns. Recently it was shown that at this pressure level a major portion of crack growth is periodic and under the influence of the fluctuating stresses. This cyclic crack growth can result in formation of special markings on the crack surfaces and has specific features (Mirzaei, 2008a). According to the schematic depicted in Fig. 2, the passage of detonation front results in a pattern of fluctuating hoop

strains which exist even after the detonation loading dies out. In presence of an axial through-thickness crack, the points on the crack surface can continue to oscillate radially. However, because the crack surface is traction free, these points also tend to displace circumferentially under the influence of stresses imposed by the neighboring material. The outcome is that the points on the crack surface are forced to displace in the resultant direction. The continuation of this effect, which is maximum at the crack center and minimum at the crack tips, results in the bulging of crack flaps. With the extension of the bulged area, large tensile stresses (equivalent to yield stress) develop in the bulged region in the axial direction of the tube. Since cracks usually tend to grow perpendicular to the largest principal stress directions, the initial self-similar growth changes into a circumferential growth by curving around the bulged region. The occurrence of branching at this point is also possible for the cases that the energy release rate of the crack is high enough to support two crack fronts (Mirzaei, 2008a; Mirzaei et al., 2009).

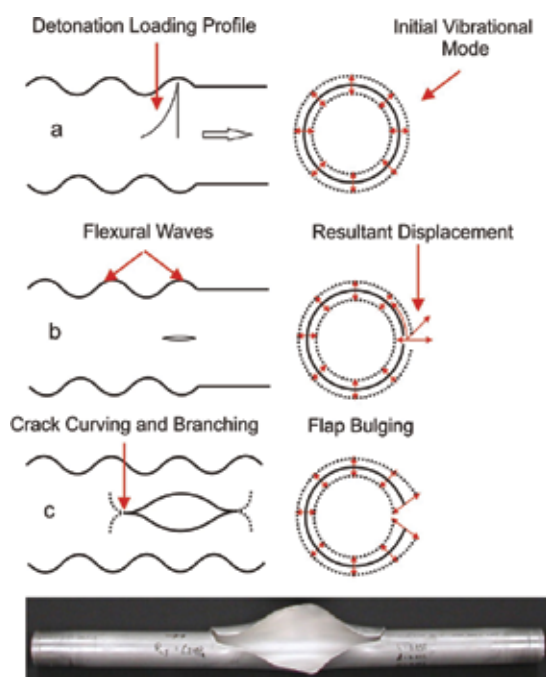


Fig. 2. Top: formation of flexural waves by a moving detonation front and the resulting flap bulging, curving, and branching (Mirzaei, 2008a). Bottom: a confined rupture of an experimental aluminum tube under internal gaseous detonation (Chao, 2004).

Unfortunately, reports on the FE simulation of detonation-driven fracture of tubes at this pressure level are not only scarce but also controversial. Due to the complexity of the process, researchers have followed two different simulation approaches. In the first approach the emphasis is essentially placed on the effects of the structural waves and the usage of dynamic fracture properties, while the probable effect of the structural behavior on the gas dynamics is ignored (Mirzaei & Karimi, 2006b; Mirzaei et al., 2009). The second approach involves coupled fluid dynamic and structural analysis using static fracture

resistance values (Cirak et al., 2007). We will discuss these issues in more detail at the end of this chapter.

2.3 High-Pressure Moving Front

For higher pressures (level 3) several cracks can initiate simultaneously and their rapid growth and branching can cause severe dynamic fragmentation. This procedure can result from both gaseous detonation (Price, 2006) and high-explosive detonation (Goto et al., 2007) (see Fig.3).

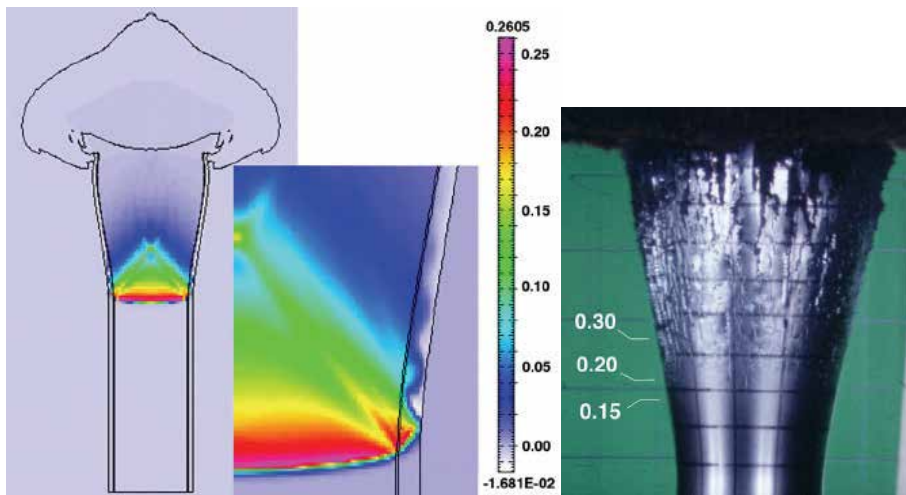


Fig. 3. Left: Computed pressure (in Mbar) in a cylinder (15 μ s after detonation). Right: optical image of fragmentation 25 μ s after detonation (Goto et al., 2007).

A very recent FE study of this procedure was reported by Soto et al. (2010) who proposed an efficient fluid–solid coupled finite element scheme for weapon fragmentation simulations.

3. Theoretical Aspects (an overview)

Successful simulations of deformation and fracture of cylindrical tubes not only involve a clear understanding of the underlying mechanisms but also require implementation of *transient dynamic elasto-plastic* FE formulation and specific modeling techniques. In sequel, a brief treatise of FE formulation and related fracture parameters is presented.

3.1 Finite Element Formulation

Scientists and engineers usually formulate mathematical models of physical phenomena in terms of differential equations. These *equations* along with the associated initial and boundary conditions form the “initial and boundary value *problems*”. In practice, finding analytical solutions for these problems is not possible for a majority of applications. The basic reason is that the solution must possess a certain degree of continuity in differentiation which is dictated by the differential equation. In other words, the differential equation places *strong* conditions on the solution. These strong conditions can be *weakened* by

recasting the differential equation in the form of certain integrals. The evaluation of these integrals can be further facilitated if the domain of the problem is subdivided into simple parts (elements). In practice, evaluation of the solution for the assemblage of these elements involves a large amount of numerical computations which can only be carried out using digital computers. The above mentioned activities constitute solution procedures which are called *Finite Element Methods*.

3.2 A Lagrangian Nonlinear Dynamic Formulation

In this section we will review some general definitions and formulations concerning finite element equations for time-dependent and non-linear continuum problems. For analyses with large displacements, rotations, and/or strains, we need to track the position and shape of the body throughout the analysis. The Lagrangian meshes conform to and move with the material points. Thus, they are most suitable for solid mechanics problems as they can handle the complicated boundaries with relative ease. Moreover, as they have the ability to follow the material points, they are most suitable to handle the history-dependent materials. Moreover, in the Updated Lagrangian (UL) formulation, for which Eulerian measures of stress and strain are used, the derivatives and integrals are taken with respect to the spatial coordinates and the only limitation is the element capabilities to deal with large distortions. Fig. 4 shows the undeformed configuration of a material continuum at time $t = t_0$ together with the deformed configuration at a later time $t = t$.

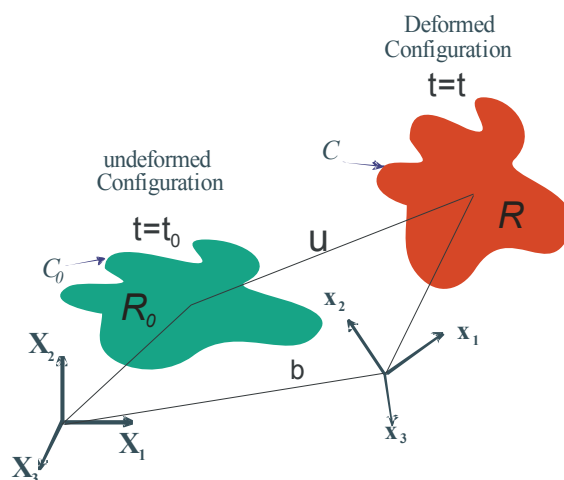


Fig. 4. Undeformed and deformed configurations of a material continuum.

We may describe the motion of the body by:

$$\mathbf{x} = \Phi(\mathbf{X}, t) \quad \text{or} \quad x_i = \phi_i(\mathbf{X}, t) \quad (1)$$

in which \mathbf{x} denotes the spatial (Eulerian) coordinates, \mathbf{X} is the material (Lagrangian) coordinates, t is time (or a time-like parameter), and Φ defines a mapping between the *initial* (undeformed) and the *current* (deformed) configurations (In the TL formulation, the initial configuration is also called the *reference* configuration since all the equations are referred to it).

Accordingly, the displacement, velocity, and acceleration can be defined as:

$$\begin{aligned}\mathbf{u} &= \mathbf{x} - \mathbf{X} \quad \text{or} \quad \mathbf{u}(\mathbf{X}, t) = \Phi(\mathbf{X}, t) - \Phi(\mathbf{X}, 0) \\ \mathbf{v}(\mathbf{X}, t) &= \frac{\partial \Phi(\mathbf{X}, t)}{\partial t} = \frac{\partial \mathbf{u}(\mathbf{X}, t)}{\partial t} \equiv \dot{\mathbf{u}} \\ \mathbf{a}(\mathbf{X}, t) &= \frac{\partial \mathbf{v}(\mathbf{X}, t)}{\partial t} = \frac{\partial^2 \mathbf{u}(\mathbf{X}, t)}{\partial t^2} \equiv \dot{\mathbf{v}}\end{aligned}\quad (2)$$

For a nonlinear dynamic problem for which we assume the *infinitesimal strain* condition the strong form is:

$$\begin{aligned}\sigma_{ji,j} + \rho b_i &= \rho \dot{v}_i \equiv \rho \ddot{u}_i \quad \text{in } \Omega \\ \text{B.C.} \begin{cases} n_i \sigma_{ji} = \bar{t}_i^*(\mathbf{x}, t) & \text{on } \Gamma_t^* \\ u_i = \bar{u}_i^*(\mathbf{x}, t) & \text{on } \Gamma_u^* \end{cases} \\ \text{I.C.} \begin{cases} u_i(t=0) = u_i^0 \\ v_i(t=0) = v_i^0 \end{cases} \\ \varepsilon_{ij} &= \frac{1}{2}(u_{i,j} + u_{j,i}) \\ \sigma_{ij} &= \sigma_{ij}(\boldsymbol{\varepsilon}(\mathbf{u}))\end{aligned}\quad (3)$$

Note that in general the stress tensor is a function of the strain tensor, which in turn is a function of displacements. The test and trial functions belong to the following spaces:

$$\begin{aligned}\delta v_i(\mathbf{X}) &\in \mathcal{U}_0, \quad \mathcal{U}_0 = \{ \delta v_i | \delta v_i \in C^0(\mathbf{X}), \delta v_i = 0 \quad \text{on } \Gamma_{v_i} \} \\ v_i(\mathbf{X}, t) &\in \mathcal{U}, \quad \mathcal{U} = \{ v_i | v_i \in C^0(\mathbf{X}), v_i = \bar{v}_i \quad \text{on } \Gamma_{v_i} \}\end{aligned}\quad (4)$$

Taking the product of the momentum equation with a test function δv_i and integrating over the current configuration, we have,

$$\int_{\Omega} \delta v_i (\sigma_{ji,j} + \rho b_i - \rho \dot{v}_i) d\Omega = 0 \quad (5)$$

and the integration by parts results in:

$$\begin{aligned}\int_{\Omega} \left[(\delta v_i \sigma_{ji})_{,j} - \delta v_{i,j} \sigma_{ji} \right] d\Omega + \int_{\Omega} \delta v_i \rho b_i d\Omega - \int_{\Omega} \delta v_i \rho \dot{v}_i d\Omega &= 0 \\ \int_{\Gamma_t^*} \delta v_i n_i \sigma_{ji} d\Gamma - \int_{\Omega} \delta v_{i,j} \sigma_{ji} d\Omega + \int_{\Omega} \delta v_i \rho b_i d\Omega - \int_{\Omega} \delta v_i \rho \dot{v}_i d\Omega &= 0 \\ \int_{\Gamma_t^*} \delta v_i \bar{t}_i^* d\Gamma - \int_{\Omega} \delta v_{i,j} \sigma_{ji} d\Omega + \int_{\Omega} \delta v_i \rho b_i d\Omega - \int_{\Omega} \delta v_i \rho \dot{v}_i d\Omega &= 0\end{aligned}\quad (6)$$

The above is known as the *principle of virtual power*, and is in fact the weak form of the momentum equation with the traction boundary conditions. Now we may consider the current domain as an assemblage of finite elements. We consider the coordinates of nodes by x_i^f , where

the superscript ranges from 1 to n_N (number of nodes). We will then use the shape functions $N^I(\mathbf{X})$ to interpolate the displacement and the virtual velocities between the nodal values as:

$$\begin{aligned} u_i(\mathbf{X}, t) &= \sum_{I=1}^{n_N} N^I(\mathbf{X}) u_i^I(t) \\ \delta v_i(\mathbf{X}, t) &= \sum_{I=1}^{n_N} N^I(\mathbf{X}) \delta v_i^I(t) \end{aligned} \quad (7)$$

Accordingly, the discrete form of the principle of virtual power can be written as:

$$\begin{aligned} &\sum_e \int_{\Omega_e} \sigma_{ij}(\boldsymbol{\varepsilon}(\mathbf{u})) N_{,j}^I \delta v_i^I d\Omega + \sum_e \int_{\Omega_e} \rho N^J N^I \ddot{u}_i^J \delta v_i^I d\Omega \\ &- \sum_e \int_{\Omega_e} \rho b_i N^I \delta v_i^I d\Omega - \sum_e \int_{\Gamma_{ie}^*} \bar{t}_i^* N^I \delta v_i^I d\Gamma = 0 \end{aligned} \quad (8)$$

The subscript 'e' denotes an individual element. Since the virtual displacements (test functions) are arbitrary for all nodes (except for the boundary) we may write:

$$\begin{aligned} &\sum_e \int_{\Omega_e} \sigma_{ij}(\boldsymbol{\varepsilon}(\mathbf{u})) N_{,j}^I d\Omega + \sum_e \int_{\Omega_e} \rho N^J N^I \ddot{u}_i^J d\Omega \\ &- \sum_e \int_{\Omega_e} \rho b_i N^I d\Omega - \sum_e \int_{\Gamma_{ie}^*} \bar{t}_i^* N^I d\Gamma = 0 \quad \forall (I, i) \notin \Gamma_{v_i} \end{aligned} \quad (9)$$

Since the stress tensor is a function of the strain tensor (which in turn is a function of displacements) the above is a system of nonlinear equations which can be solved using the Newton-Raphson method (Belytschko et al., 2000 ; Gullerud et al., 2004).

$$\begin{aligned} &\sum_e \int_{\Omega_e} \sigma_{ij}(\boldsymbol{\varepsilon}(u_i^I + du_i^I)) N_{,j}^I d\Omega - \sum_e \int_{\Omega_e} \rho N^I b_i d\Omega \\ &+ \sum_e \int_{\Omega_e} \rho N^J N^I \ddot{u}_i^J d\Omega - \sum_e \int_{\Gamma_{ie}^*} \bar{t}_i^* N^I d\Gamma = 0 \\ &\sum_e \int_{\Omega_e} \left\{ \sigma_{ij}(\boldsymbol{\varepsilon}(u_i^I)) + \frac{\partial \sigma_{ij}}{\partial \varepsilon_{kl}} \frac{\partial \varepsilon_{kl}}{\partial u_m^J} du_m^J \right\} N_{,j}^I d\Omega \\ &+ \sum_e \int_{\Omega_e} \rho N^J N^I \ddot{u}_i^J d\Omega - \sum_e \int_{\Omega_e} \rho N^I b_i d\Omega - \sum_e \int_{\Gamma_{ie}^*} \bar{t}_i^* N^I d\Gamma = 0 \end{aligned} \quad (10)$$

In order to carry out the differentiation of the strain tensor we use the kinematic equations to give,

$$\begin{aligned}
\varepsilon_{kl} &= \frac{1}{2}(u_{k,l} + u_{l,k}), \quad u_i(\mathbf{x}) = \sum_{I=1}^n N^I(\mathbf{x})u_i^I \\
\Rightarrow \varepsilon_{kl} &= \frac{1}{2} \sum_{I=1}^n \left(\frac{\partial N^I}{\partial x_l} u_k^I + \frac{\partial N^I}{\partial x_k} u_l^I \right) \\
\frac{\partial \varepsilon_{kl}}{\partial u_m^J} &= \frac{1}{2} \sum_{I=1}^n \left(\frac{\partial N^I}{\partial x_l} \delta_{IJ} \delta_{km} + \frac{\partial N^I}{\partial x_k} \delta_{IJ} \delta_{lm} \right) \equiv \frac{1}{2} \sum_{I=1}^n \left(\frac{\partial N^J}{\partial x_l} \delta_{km} + \frac{\partial N^J}{\partial x_k} \delta_{lm} \right)
\end{aligned} \tag{11}$$

Hence, we may rewrite Eq. (10) as follows,

$$\begin{aligned}
&\sum_e \int_{\Omega_e} \left\{ \sigma_{ij}(\boldsymbol{\varepsilon}(u_i^J)) + \frac{\partial \sigma_{ij}}{\partial \varepsilon_{ml}} \left[\frac{1}{2} \left(\frac{\partial N^J}{\partial x_l} \right) \right] du_m^J + \frac{\partial \sigma_{ij}}{\partial \varepsilon_{km}} \left[\frac{1}{2} \left(\frac{\partial N^J}{\partial x_k} \right) \right] du_m^J \right\} N_{,j}^I d\Omega \\
&+ \sum_e \int_{\Omega_e} \rho N^J N^I \ddot{u}_i^J d\Omega - \sum_e \int_{\Omega_e} \rho N^I b_i d\Omega - \sum_e \int_{\Gamma_{ie}^*} \bar{t}_i^* N^I d\Gamma = 0
\end{aligned} \tag{12}$$

in which we used the properties of the Kroneker Delta to change the indices of *epsilon* and *N*. Now we change the dummy index *k* to *l* and rearrange the terms as follows:

$$\begin{aligned}
&\sum_e \int_{\Omega_e} \left\{ \sigma_{ij}(\boldsymbol{\varepsilon}(u_i^J)) + \frac{\partial \sigma_{ij}}{\partial \varepsilon_{ml}} \frac{\partial N^J}{\partial x_l} du_m^J \right\} N_{,j}^I d\Omega \\
&+ \sum_e \int_{\Omega_e} \rho N^J N^I \ddot{u}_i^J d\Omega - \sum_e \int_{\Omega_e} \rho N^I b_i d\Omega - \sum_e \int_{\Gamma_{ie}^*} \bar{t}_i^* N^I d\Gamma = 0
\end{aligned} \tag{13}$$

Another changing of the dummy index *m* to *k* for convenience, results in,

$$\begin{aligned}
&\sum_e \int_{\Omega_e} \left\{ \sigma_{ij}(\boldsymbol{\varepsilon}(u_i^J)) + \frac{\partial \sigma_{ij}}{\partial \varepsilon_{kl}} N_{,l}^J du_k^J \right\} N_{,j}^I d\Omega \\
&+ \sum_e \int_{\Omega_e} \rho N^J N^I \ddot{u}_i^J d\Omega - \sum_e \int_{\Omega_e} \rho N^I b_i d\Omega - \sum_e \int_{\Gamma_{ie}^*} \bar{t}_i^* N^I d\Gamma = 0
\end{aligned} \tag{14}$$

With the definition of the *tangent stiffness matrix* and internal and external nodal forces, we may write the above as:

$$\begin{aligned}
M^{JI} \ddot{u}_i^J + K_{ik}^{IJ} du_k^J + \text{int } F_i^I - \text{ext } F_i^I &= 0 \quad \forall (I, i) \notin \Gamma_{u_i} \\
\mathbf{Ma} + \mathbf{Kdu} + \mathbf{f}^{\text{int}} - \mathbf{f}^{\text{ext}} &= 0
\end{aligned} \tag{15}$$

For which we have,

$$\begin{aligned}
 M^{IJ} &= \sum_e \int_{\Omega_e} \rho N^J N^I d\Omega \\
 K_{ik}^{IJ} &= \sum_e \int_{\Omega_e} \frac{\partial \sigma_{ij}}{\partial \varepsilon_{kl}} N_{,j}^I N_{,l}^J d\Omega \\
 {}^{\text{int}}F_i^I &= \sum_e \int_{\Omega_e} \sigma_{ij} \left(\varepsilon(u_i^I) \right) N_{,j}^I d\Omega \\
 {}^{\text{ext}}F_i^I &= \sum_e \int_{\Omega_e} \rho b_i N^I d\Omega + \sum_e \int_{\Gamma_{ie}^*} \bar{t}_i^* N^I d\Gamma
 \end{aligned} \tag{16}$$

A key element in the above expressions is the derivative of the stress tensor with respect to the strain tensor. If plastic deformation is considered as the source of nonlinearity, the stress tensor would become a function of the *strain-tensor increment* (which in turn is a function of the *displacement increment*), defined over a time increment Δt , for which we have:

$$\begin{aligned}
 K_{ik}^{IJ} &= \int_{\Omega} C_{ijkl} N_{,j}^I N_{,l}^J d\Omega \equiv \int_{\Omega} \frac{\partial \sigma_{ij}}{\partial \Delta \varepsilon_{kl}} N_{,j}^I N_{,l}^J d\Omega \\
 {}^{\text{int}}F_i^I &= \int_{\Omega} \sigma_{ij} \left(\Delta \varepsilon(u_i^I) \right) N_{,j}^I d\Omega \\
 {}^{\text{ext}}F_i^I &= \int_{\Omega} \rho b_i N^I d\Omega + \int_{\Gamma_i^*} \bar{t}_i^* N^I d\Gamma
 \end{aligned} \tag{17}$$

Here a main difficulty is to compute the stress caused by the incremental strain through the integration of the plastic stress-strain equations. On the other hand, the dynamic nature of the problem requires progression of the solution in time domain, for which we have explicit and implicit time integration methods. The explicit algorithms remove the need to invert the stiffness matrix in every time step by de-coupling the equilibrium equations. However, they are *conditionally* stable and the length of the time step must be smaller than a critical value to ensure that the solution converges. Hence, explicit algorithms are generally used for problems which by nature require small time steps, such as wave propagations, shock responses, and explosive or impact loadings.

The implicit algorithms are unconditionally stable and the maximum time step is governed by accuracy considerations. However, they require inversion of the stiffness matrix at every time step and this significantly increases the cost of computations. However, the time step length can be substantially larger thereby reducing the overall number of time steps. Thus, implicit algorithms are suitable for dynamic problems where the responses involve low frequency components, such as seismic responses or large deformations of elasto-plastic components under ramp or step loadings (Gullerud et al., 2004).

3.3 A Treatise of Related Fracture Parameters

Fracture is the separation of a component into, at least, two parts. This separation can also occur locally due to formation and growth of cracks. The crack propagation can occur in many ways, for instance we may have fast-unstable and slow-stable crack growth under monotonic loading or a cyclic growth under alternating loads. Based on Fracture Mechanics principles an energy criterion for the onset of crack growth can be defined in the following general form:

$$G = -\frac{d\Pi}{dA} \geq R \quad (18)$$

in which G is called the *energy release rate* (also known as the *crack driving force*), A is the cracked area, and R is the resistance of the material to crack growth. The energy release rate, G , can be considered as the energy *source* for the crack growth and may be obtained from the stress analysis of the cracked geometry. On the other hand, the resistance to crack growth (R) can be considered as the energy *sink* and depends on the operating *fracture mechanism*. It should be mentioned that the latter depends on many factors including: the chemical composition and microstructure of the material, temperature, environment, loading rate, and the state of stress. The resistance to crack growth can be defined by a special term called the *toughness* of the material. We should also distinguish between *ductile* and *brittle* as two different types of fracture. By definition, ductile fracture is always accompanied by a *significant* amount of plastic deformation, while brittle fracture is characterized by very little plastic deformation. It should be noted that due to finite strength of materials, there is always a small *damaged* zone around the crack tip. For metals, this damaged zone is referred to as the *crack tip plastic zone*. If the size of the plastic zone is small we may use Linear Elastic Fracture Mechanics (LEFM) parameters like G and K (energy release rate and stress intensity factor respectively). Nevertheless, for large plastic zone conditions the linear elastic assumptions are not correct, i.e., LEFM is not applicable and nonlinear models must be used. If this nonlinearity is not very significant, it can be handled with a *non-linear elastic* model, for which a non-linear-elastic energy release rate called J (usually known as the J -Integral) is used. However, we should note that similar to the LEFM, there is a limit to the validity of J with regard to the size of the plastic zone. Thus for situations where the crack tip plasticity is wide-spread the appropriate parameter is the crack-tip opening displacement or angle (CTOD and CTOA respectively).

3.3.1 CTOA Methodology for FE Simulations of Ductile Crack Growth

Both CTOD and CTOA can be used for determination of the onset of crack growth in ductile fracture problems. However, for problems that involve both initiation and growth phases, it has been shown that modeling approaches based on CTOA can provide a viable growth criterion for thin materials (Gullerud et al., 1999). Using the FE method, the CTOA is defined by nodal displacements normal to the crack plane (Gullerud et al., 1999, 2004). For 2D planar problems the CTOA can simply be considered as the angle defined by the first nodal displacement, one element back from the crack tip. Note that the CTOA (as defined for WARP3D FE code and used here) is the full angle from one crack face to the other. The CTOA-controlled crack growth operates by advancing the crack front a prescribed distance when the CTOA reaches a specified *critical* value.

The critical CTOA is calculated from the critical CTOD, which is a measure of fracture toughness. Fracture toughness for a specific material can vary with the thickness, strain rate, temperature, etc. The general methodology for using a calibrated value of critical CTOA for an arbitrary component made of a specific material can be summarized as: (a) obtain well characterized experimental fracture test data with crack tearing up to and beyond maximum load; (b) analyze the tests with elastic-plastic CTOA controlled crack-tearing finite element analysis with crack front modeling, mesh size, and material model of sufficient fidelity to

represent the tests; (c) iterate the FEA with an assumed CTOA value until the results are close to the experimental data in predicting deformation, strains, maximum load and also are reasonably close in predicting the trends and amount of crack growth; (d) use this calibrated CTOA value, material data, and mesh size to predict crack tearing fracture behavior for other types of specimens or applications (Gullerud et al., 1999).

3.3.2 Crack Growth Simulation using Cohesive Elements

In simulations of ductile crack growth there are instances that a single-parameter treatment of cracked bodies may not yield accurate results in presence of widespread yielding. The remedy is to use either two-parameter fracture mechanics methods (like the J-Q approach) or other solutions based on damage mechanics models (Anderson, 1994; Cornec et al., 2003). Among the latter, the cohesive model with interface elements seems very attractive because it is computationally efficient and its two parameters can be determined with relative ease. In this approach, the crack extension evolves as a direct outcome of the computations by creation of new traction free crack faces. This method has been formulated in WARP3D to provide it with the capability to simulate mixed mode crack growth, and also phenomena such as crack turning and branching (Gullerud et al., 2004). In general, the interface elements consist of two surfaces which connect the faces of two adjacent solid elements. The crack propagation occurs when the energy release rate reaches a critical value (G_c). The implementation of interface elements requires a *cohesive constitutive relation* which in WARP3D is defined as a traction-separation relationship of the form:

$$t = \exp(1.0)\sigma_c \frac{\delta}{\delta_c} \exp\left(-\frac{\delta}{\delta_c}\right) \quad (19)$$

In the above, t is the effective traction between the two surfaces of an interface element, σ_c is the peak traction which shows the local material strength, δ is the effective opening displacement, and δ_c denotes the value of δ at the point of peak traction. The *cohesive fracture energy* can be described by:

$$\Gamma_c = \exp(1.0)\sigma_c\delta_c \quad (20)$$

Using the procedure outlined by Li and Siegmund (2002) for thin-sheet ductile material in plane stress condition, the cohesive strength can be selected as $\sigma_c = 2\sigma_y$, where σ_y is the yield stress of the bulk material. Another key parameter is the length of the cohesive zone which is defined as the distance from the crack tip to the point where the maximum cohesive traction is attained. In order to obtain accurate results using the cohesive zone model (CZM), the tractions in the cohesive zone must be represented properly by the FE spatial discretization, i.e., a minimum number of elements is required to resolve the variations within the cohesive zone (Falk et al., 2001).

4. Practical Examples of FE Simulation

The following sections present two practical examples of FE simulations of deformation and fracture of thin tubes under internal gaseous detonation. The simulations are carried out for an experimental detonation tube and an exploded gas cylinder.

4.1 Elastodynamic Response of an Aluminum Tube (Mirzaei & Karimi, 2006b)

This section presents the FE analysis of the elastodynamic response of an aluminum tube *without* flaw using WARP3D (a 3D research code for dynamic nonlinear fracture analysis of solids). In this code the nonlinear dynamic equilibrium equations are solved using an incremental-iterative *implicit* formulation with full Newton iterations to eliminate residual nodal forces (Gullerud et al., 2004). Although for this case study axisymmetric elements and *explicit* time integration scheme are the most appropriate choices, the aim of this initial analysis was to obtain a *validated 3D model* before attempting the crack growth simulation. The model consisted of 29400 8-node brick elements. The element distribution in the radial, longitudinal, and circumferential directions was 5, 280, and 21 respectively. A major part of the FE simulation was the modeling and preconditioning of a moving pressure load with the required profile. The detonation loading history was preconditioned as a discrete version of the exponential approximation to the Taylor-Zeldovich model and was prescribed as a function of time at each node. The results of the FE simulation included the initial detonation loading and the subsequent *reflections of the flexural waves* at the flanges. In Fig.5 the FE results are compared with the experimental results of Chao (2004). It is clear that there is a very good agreement between the results over this time interval.

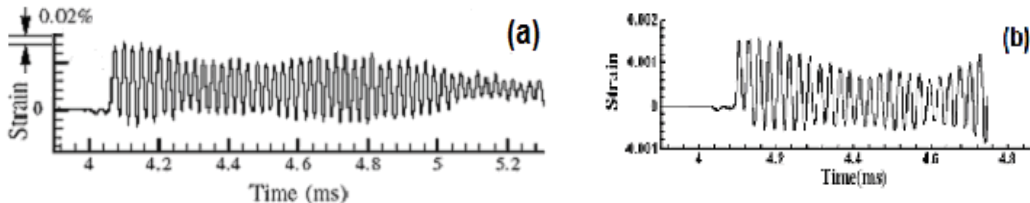


Fig. 5. (a): Experimental result (Chao, 2004) (b): WARP3D solution (Mirzaei & Karimi, 2006b).

4.2 Crack Growth Simulation for an Aluminum Tube (Mirzaei & Karimi, 2006b)

In the FE model of the pre-cracked tube the size of the far-field mesh was set to allow the simulation of the flexural (breathing mode) waves (in the form of hoop strains) so that they can be compared with the experimental strain signals. An upper bound of the mesh size was estimated using flexural wavelength. Because of the restrictions imposed by the CTOA approach, the crack front could not contain convenient collapsed elements that are used in computational fracture mechanics. Thus, the size of the elements near the crack tip was set small enough to capture the crack front plasticity and to maintain the accuracy of CTOA approach. In his experimental study, Chao (2004) used an axially oriented external notch, shaped as a semi elliptical surface crack. However, the finite element modeling of a semi elliptical surface crack, 35.2 mm long and 0.64 mm deep within a 0.89 mm wall thickness, could cause severe problems with regards to the shape and aspect ratios of the elements.

Therefore, the initial flaw was modeled by considering an equivalent through thickness crack. In practice the remote loading process was controlled so that the loading increments acting at the crack tip were not so large to cause large deformations in the crack tip elements in one loading step. This phenomenon is referred to as *overshoot* (Gullerud et al., 2004). Options are provided in the WARP3D code to control overshoot by restricting the size of one loading step based on projecting the expected CTOA from the previous load step. A control for general reduction of the step loading schedule is also available. But these approaches were not applicable in this study because they interfere with the simulation of detonation as a moving pressure load. In general, the "overshoot control" and the "automatic load reduction" change the load step, thereby affecting the detonation simulation. Thus, the load step size was substantially decreased to avoid the overshoot, resulting in prolonged runtimes.

4.2.1 Estimation of critical CTOA

Unfortunately, a CTOA calibration for the material in question is not available for such *high strain rates*. Hence, the critical CTOA was inferred from the experimental results as follows. Table 1 shows the experimental conditions and results for the pre-flawed aluminum tubes.

<i>Shot</i>	<i>d</i>	<i>2a</i>	<i>h</i>	<i>V_{cj}</i>	<i>P_{cj}</i>	<i>L</i>	<i>Fracture behavior</i>
	<i>mm</i>	<i>mm</i>	<i>mm</i>	<i>m/s</i>	<i>MPa</i>	<i>mm</i>	
30	0.64	35.2	0.89	2384	4.0	610	Crack confined in notch
34	0.64	35.2	0.89	2404	6.1	610	Forward and Backward growth

Table 1. Experimental conditions and fracture behavior for a pre-flawed aluminum tube (Chao, 2004).

The results indicate that in shot 30 the crack slightly advanced but remained confined within the notch, which indicates initiation plus a little propagation. Thus, shot 30 was recognized suitable for predicting the critical CTOA under actual dynamic condition that satisfies all of the loading rate and component thickness expectations for fracture toughness. The experimental results indicated that the crack remained confined in the notch exactly 0.2 ms after arrival of detonation loading. Thus, an initial FE simulation was carried out for this time period and the resulting maximum value of CTOA (4.61 degrees) was used as "dynamic critical CTOA" in the crack growth simulation for Shot 34 (see Table 1). Fig. 6 shows the results obtained from transient modeling of the moving pressure with full-dynamic effects.

As depicted in Fig. 6, the detonation front travels from left to right and causes a pattern of flexural structural waves as it proceeds toward the left crack tip. Also, it is clear the crack grows more in the forward direction, i.e., in the direction of movement of the detonation front. The same phenomenon was observed by Chao and shepherd in their experiments (Chao & Shepherd, 2005).

This asymmetric growth, which is one of the specific features of detonation-driven fracture of tubes and occurs quite naturally as a result of the difference between the relative speeds of the structural waves with respect to the two crack tips, will be discussed in sequel.

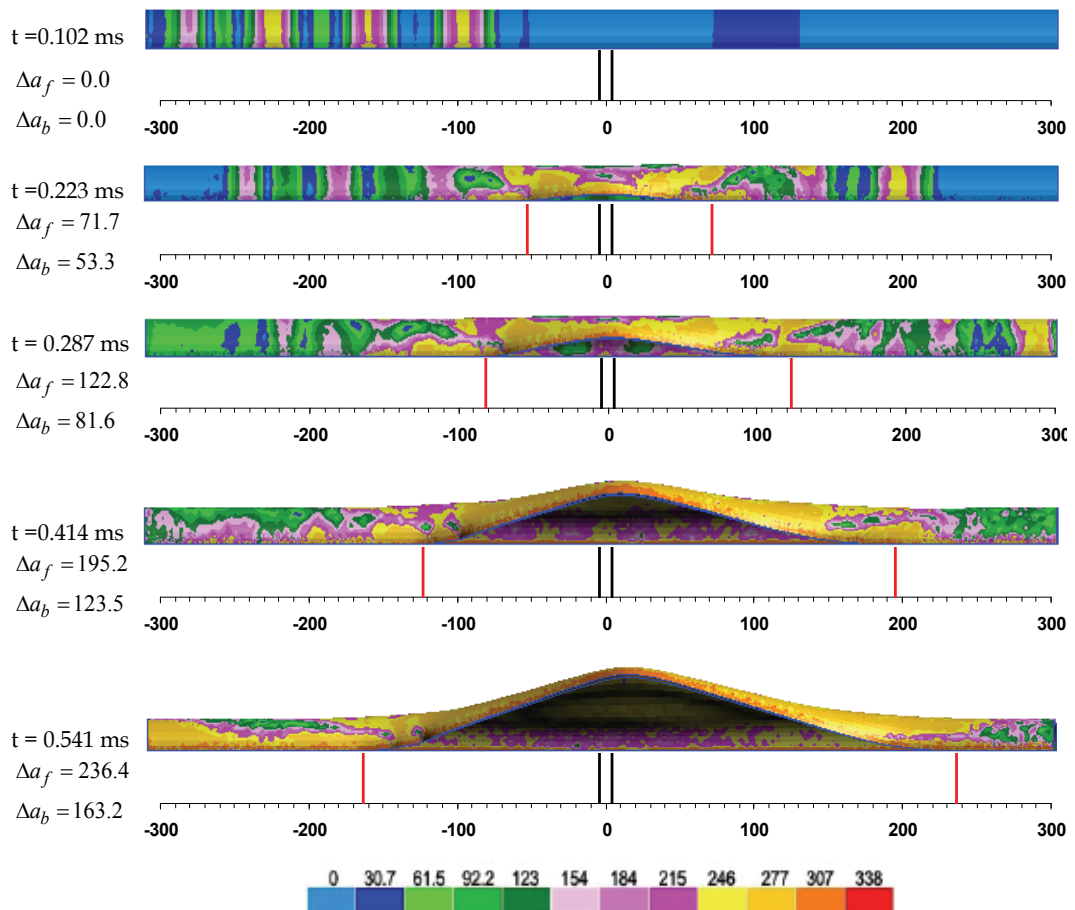


Fig. 6. Snapshots of simulated crack growth for an aluminum tube at different time intervals including the contours of the first principal stress component.

The simulations presented in Fig. 6 seem to be representative of a number of actual growth characteristics including flap bulging. Nevertheless, there are two major inconsistencies between the simulation results and the actual crack growth features. The first one is that the simulated crack speed is 2 times the actual average speed during the self-similar growth. The reason is that the retardation effects of cyclic crack growth were not considered in this analysis (Mirzaei et al., 2009). The second problem is that crack curving and branching could not be simulated using the CTOA approach implemented in WARP3D.

4.3 Crack Growth Simulation for an Aluminum Tube (Cirak et al., 2007)

Cirak et al. (2007) performed large-scale fluid-structure interaction simulation of viscoplastic deformation and fracturing of the experimental aluminum tube of Chao & Shepherd (2005). In practice, they used the static plain strain fracture toughness K_{Ic} to obtain the properties of their cohesive elements.

In general, their results showed some inconsistencies with the experiments, particularly with respect to the crack growth magnitude, speed, and premature cracking of the flap edges (see Fig.7). They also reported that their computational crack speeds were 2-3 times higher than experimentally observed (Deiterding et al., 2006).

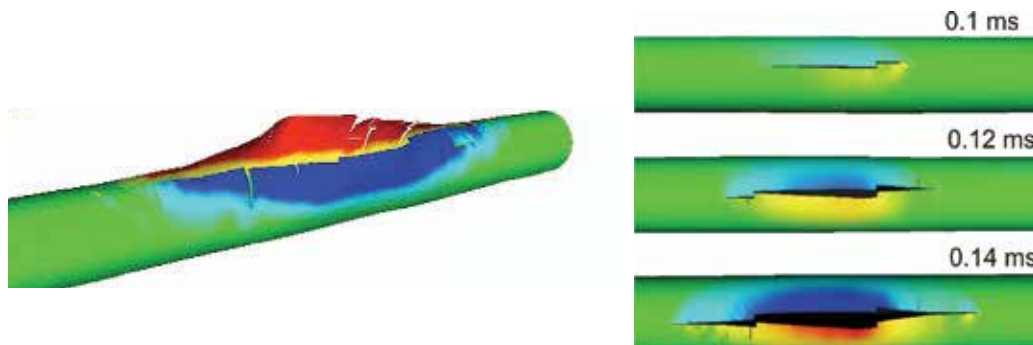


Fig.7. Crack growth simulation for an aluminum tube under internal gaseous detonation. Left: (Cirak et al., 2007). Right: (Deiterding et al., 2006)

4.4 FE Analysis of an Exploded Gas Cylinder (Mirzaei et al., 2009)

This section presents the finite element simulations of deformation and fracture of a gas cylinder which catastrophically failed as a result of an accidental explosion (see Fig.8). The results of a previous detailed investigation of this incident indicated that detonation of a low-pressure oxygen-rich mixture of hydrogen and oxygen was the cause of the cylinder failure (Mirzaei, 2008). The overall transient dynamic response of the cylinder to gaseous detonation loading was studied using the ANSYS/LS-DYNA V10 package and the crack growth simulations were performed using the WARP3D-R15 research code. The crack growth analyses were performed using interface cohesive elements. The accuracy of the finite element results was verified using analytical solutions and data collected from the remains of the cylinder.

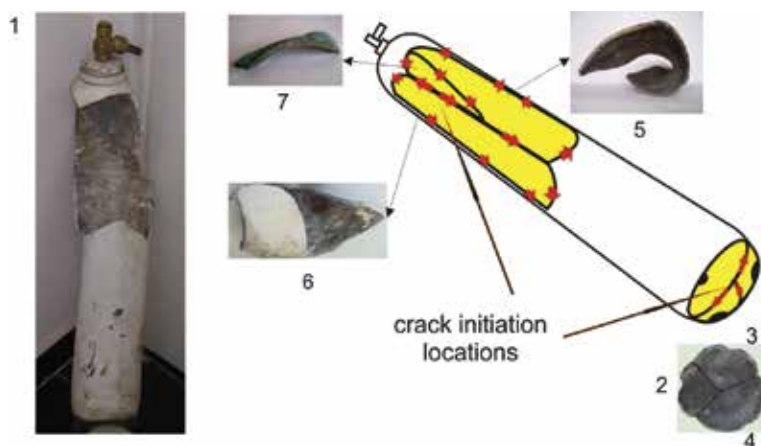


Fig. 8. Remains of an exploded gas cylinder along with a schematic showing the original locations of the collected fragments and the crack growth directions. The sketch is not to scale.

4.4.1 Elasto-plastic Dynamic Response

A transient elasto-plastic dynamic analysis was carried out to obtain the structural response of the cylinder to a moving pressure load with specific profile using 2011 rotary symmetric solid elements. The preconditioning and application of the loading history was similar to the procedure described for the aluminum tube before. The simulation was carried out for the detonation speed of 2300 m/s, initial pressure of 1.9 MPa, maximum pressure of 37 MPa, and exponential decay factor of $T = 0.15$ ms (Mirzaei, 2008). In order to validate the FE model, the results were compared with the analytical solutions reported in Refs (Mirzaei et al 2005; Mirzaei 2008b). To avoid the effects of the boundary conditions on the main signal, the results were compared at a distance of 200 mm from the upper neck. Fig.9 shows the results for a time period of 1ms, which includes the initial detonation loading and the subsequent reflections of the flexural waves.

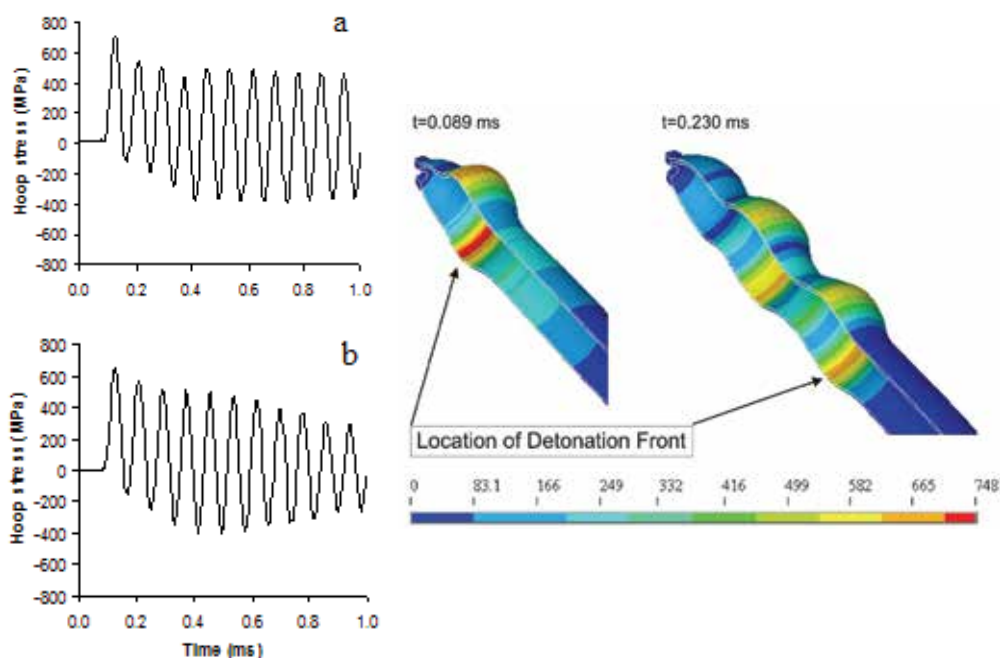


Fig. 9. Left: hoop stress vs. time at a distance of 200 mm from the upper neck. The results include the main signal (up to 0.35 ms) plus the reflected waves, a): elasto-dynamic solution (analytical), b): elasto-plastic dynamic solution (FE). Right: two snapshots showing the development of flexural waves with time and the related distributions of the Von-Mises stress (elasto-plastic dynamic FE). The stresses are in MPa.

The very good agreement between the two results for the main signal (the first 4 stress waves) is quite clear. However, the FE solution shows a stronger decay in the amplitude of the reflected waves. This can be attributed to the restraining effects of the bottom cap which was modeled as a simple support in the analytical solution. Fig. 9 also shows two snapshots of the development of flexural waves with time and the related distributions of the Von-Mises stress (elasto-plastic dynamic FE). It is clear that the maximum stress occurs at a

distance of 100 mm from the upper neck. This is the same location that the initiation of a 6-mm axial crack was recognized on the fragments Nos. 5 and 7 (see Fig.8).

4.4.2 Crack Growth Simulation

The dynamic mechanical properties of the cylinder material are shown in Table 2. The details of the experimental procedures were reported in (Mirzaei, 2008).

ρ	E	ν	σ_y	σ_u	K_c (h= 6 mm)
kg/m ³	GPa		MPa	MPa	MPa \sqrt{m}
8000	206	0.33	598	700	72

Table 2. Mechanical properties for the cylinder material (carbon steel).

Accordingly, the cohesive energy for crack growth was computed as

$$\Gamma_c = G_c = \frac{K_c^2}{E} = 25.16 \text{ kJ/m}^2 \quad (21)$$

Although different models have been proposed to estimate the length of the cohesive zone, the two most commonly used models in the literature are proposed by Hillerborg et al. (1976) and Rice (1979). The simulations reported here are based on the Hillerborg's model in which the cohesive zone length is expressed by:

$$l_{cz} = E \frac{G_c}{\sigma_c^2} \quad (22)$$

In this study 7-8 elements were used over the cohesive zone length to be able to fully resolve the variations within this zone. Fig.10 depicts the snapshots of simulated crack growth at different time intervals showing the contours of the first principal stress component. *The successive opening and closing of the crack flaps is a clear indication of a cyclic crack growth governed by the flexural waves.* Also, it is clear the crack grows more in the forward direction, i.e., in the direction of movement of the detonation front. The same phenomenon was observed by Chao and shepherd in their experiments (Chao, 2004; Chao & Shepherd 2005). This asymmetric growth occurs quite naturally as a result of the difference between the relative speeds of the structural waves with respect to the two crack tips. Let us denote the speed of the structural waves, the forward tip, and the backward tip by V_{sw} , V_{ft} , and V_{bt} respectively. In practice, the speed at which each stress cycle outdistances from the forward tip is; $V_{sw} - V_{ft}$, which is naturally less than $V_{sw} + V_{bt}$ (the speed of separation between the flexural wave and the backward tip). Hence, the *effective* time that each stress cycle acts on crack tips is longer for the forward tip.

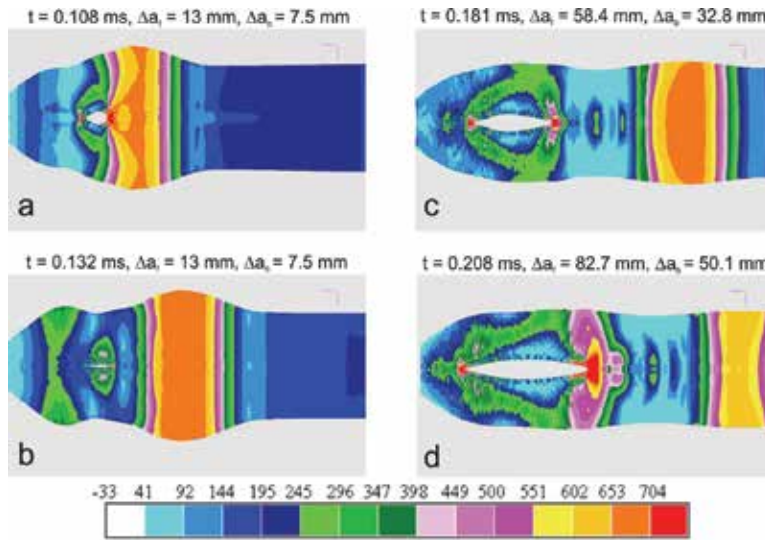


Fig. 10. Snapshots of simulated crack growth at different time intervals showing the contours of the first principal stress component. The movement of the detonation front is from left to right. The stresses are in MPa.

Although the simulations presented in Fig. 10 seem to be representative of a number of actual growth characteristics, there are major inconsistencies between the simulation results and the actual crack growth features. The most important one is that the simulated crack speed is 3 times the actual average speed during the self-similar growth. Another problem is that the crack branching does not occur at the actual location. These inconsistencies show that the implemented approach has not been able to model all the features of the cyclic crack growth mechanism. The remedy is that the growth retardations imposed by various *fatigue crack closure* mechanisms should be incorporated in the FE simulations.

4.4.3 Crack Growth Simulation with Retardation

In practice, the high amplitude stress cycles (caused by high-speed pressure) create an extreme low cycle fatigue (LCF) crack growth problem which can be represented by the upper portions of the Region III on the classical fatigue crack growth rate diagrams. In this Region the static modes of crack propagation are also operative, and because of high amounts of plastic deformation at the crack tip and the resultant plastic stretches on the crack wake the main retardation mechanism is the *plasticity-induced crack closure* (Anderson, 1995). Under these conditions, a cycle by cycle FE simulation of crack closure is the best practice. However, such simulation requires specific FE formulation for the cohesive elements. Thus, as an initial step, the closure effects were artificially incorporated into the WARP3D model by increasing the critical energy release rate in every increment. Accordingly, the cohesive parameters of the elements in front of the crack tip were proportionally increased to maintain the same stiffness in the linear part of traction-separation model after every incremental growth. Fig. 11 shows 4 snapshots of simulated crack growth with retardation at different time intervals including the contours of the first principal stress.

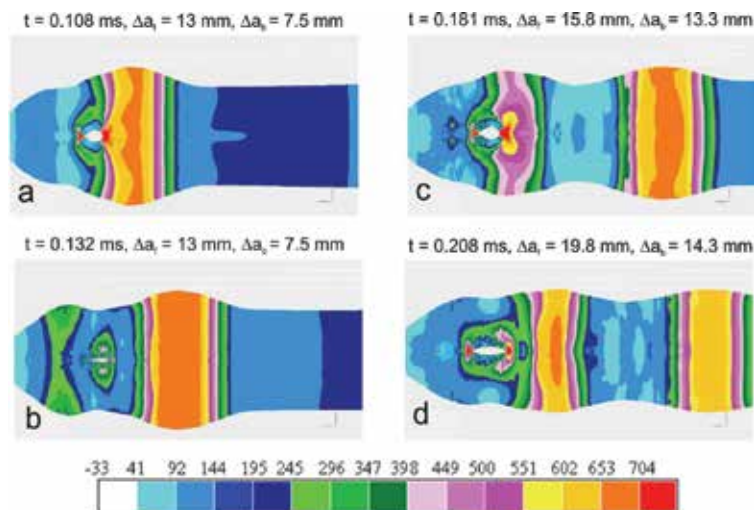


Fig. 11. Snapshots of simulated crack growth with retardation at different time intervals showing the contours of the first principal stress. The movement of the detonation front is from left to right. The stresses are in MPa.

In practice, the incorporation of retardation effects not only adjusted the overall crack speed but also resulted in crack curving and branching at the expected location. It is interesting to note that because of the growth retardation caused by closure mechanisms the crack experiences a relatively larger number of cycles to grow to a certain length. Accordingly, the cyclic bulging of the crack flaps can expand to the extent that the resultant axial stresses are significant enough to cause curving and branching. Fig. 12 also shows the onset of a second branching similar to the one that actually occurred in the exploded cylinder (see Fig.8).

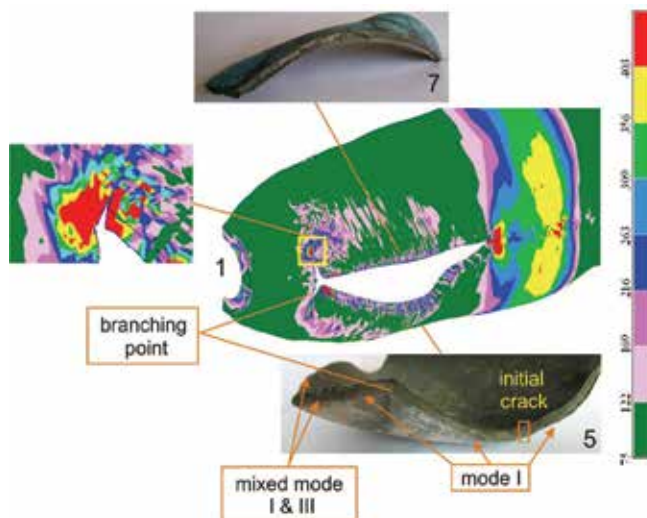


Fig. 12. Simulation of crack flap bulging and the resultant crack branching. The onset of the second branching is also visible in the magnified image of the upper branch. The stresses are in MPa.

5. Conclusions and Some Avenues for Future Research

The structural response of a cylindrical tube to internal moving pressure depends on several parameters including the dimensions and mechanical properties of the tube as well as the magnitude and speed of the moving pressure. For instance, high-speed traveling of relatively *low pressures* (level 1) can result in a pattern of fluctuating high frequency *elastic* strains. For the FE analysis of this problem an elastodynamic formulation with axisymmetric 2D elements and explicit time integration is quite appropriate. An interesting research activity in this area will be the investigation of the effects of successive gaseous detonations on the vibrational behavior of Pulse Detonation Engines.

Traveling of higher pressures (level 2), that can cause stresses in the order of the ultimate tensile strength of the material, may initiate axial cracks which with further propagation can result in *limited* fragmentation of the tube wall. Usually the initial cracks are created by local excessive shear and the self-similar crack propagation is governed by the structural waves. The FE analysis of this problem requires nonlinear dynamic elasto-plastic FE formulation with 3D elements, along with cohesive elements for crack initiation and growth simulation. It should be noted that a successful FE simulation of such crack growth should capture a number of specific features of detonation-driven fracture of cylindrical tubes. These features are:

- a) Cyclic crack propagation during the self-similar growth of the axial crack
- b) Cyclic bulging of the crack flaps
- c) Crack curving and branching adjacent to the bulged area

It should be emphasized that, since a major portion of detonation-driven fracture of cylindrical tubes can involve *cyclic* crack growth, successful simulation of the fragmentation and accurate assessments of quantities like crack speed require specific procedures that suit this particular type of growth. This involves the implementation of specialized techniques in the modeling process plus the inclusion of retardation phenomena like fatigue crack closure. Unfortunately, the current commercial FE codes do not have built-in capabilities for such simulations.

Finally, traveling of *very* high pressures (level 3), that can cause stresses much higher than the ultimate tensile strength of the material, can simultaneously initiate several cracks in the tube wall. The rapid growth and branching of these cracks result in severe dynamic fragmentation. For this class of problems usually Lagrangian FE codes with explicit time integration along with simple failure criteria (like the equivalent plastic strain) are used. However, a main difficulty is that the problem can involve hundreds of thousands of time steps and millions of elements. Moreover, numerical instabilities caused by separation of several fragment of elements must be taken into account. In practice, nonlinear-explicit 3D finite element codes like DYNA3D and LS-DYNA provide many capabilities for such simulations.

6. References

- Anderson, T.L. (1995). *Fracture Mechanics: Fundamentals and Applications*. 2nd ed. CRC Press. ISBN 0-8493-4260-0.

- Beltman, W., Burcsu, E., Shepherd, J., & Zuhail, L. (1999). The structural response of cylindrical shells to internal shock loading. *Journal of Pressure Vessel Technology*, 121, pp. 315–322.
- Beltman, W., & Shepherd, J. (2002). Linear elastic response of tubes to internal detonation loading. *Journal of Sound and Vibration*, 252 (4), pp. 617–655.
- Belytschko, T., Liu, W.K., & Moran B. (2000). *Nonlinear finite element for continua and structures*. Wiley.
- Chao, T.W., & Shepherd, J. E. (2005). Fracture response of externally flawed aluminum cylindrical shells under internal gaseous detonation loading. *International Journal of Fracture* ;134(1):59–90.
- Chao, T.W., (2004) *Gaseous detonation-driven fracture of tubes*. PhD thesis, California Institute of Technology, Pasadena, California.
- Cirak, F., Deiterding, R., & Mauch, SP. (2007). Large-scale fluid-structure interaction simulation of viscoplastic and fracturing thin-shells subjected to shocks and detonations. *Computers & Structures*; 85(11-14):1049–65.
- Cornec, A., Scheider, I., & Schwalbe, K.H. (2003). On the practical application of the cohesive model. *Engineering Fracture Mechanics*. 70:1963-1987.
- Deiterding, R., Cirak, F., & Meiron, D. (2006). Computational results for the fracturing tube validation experiment. *ASCI/ASAP Research review Caltech*
<http://www.cacr.caltech.edu/asc/wiki/pub/FracturingTube/WebHome/DetFracSim.ppt>
- Falk, M.L., Needleman, A., & Rice, J.R. (2001). A critical evaluation of cohesive zone models of dynamic fracture. *J Phys IV Proc* 2001:543–50.
- Goto, D.M., Becker, R., Orzechowski, T.J., Springer, H.K., Sunwoo, A.J., & Syn, C.K.R. (2008). Investigation of the fracture and fragmentation of explosively driven rings and cylinders. *International Journal of Impact Engineering*, Vol.35, Issue 12, 1547-1556.
- Gullerud, A.S., Koppenhoefer, K.C., Roy, A., & Dodd, Jr R.H. (2004). WARP3D: 3-D dynamic nonlinear fracture analysis of solids using parallel computers and workstations. Structural Research Series (SRS) 607, UILU-ENG-95-2012, University of Illinois at Urbana-Champaign.
- Gullerud, A.S., Dodds, Jr R.H., Hampton, R.W., & Dawicke, D.S. (1999). Three-dimensional modeling of ductile crack growth in thin sheet metals: computational aspects and validation. *Engineering Fracture Mechanics*, 63:347-74.
- Hillerborg, A., Modeer, M., & Petersson, P.E. (1976). Analysis of crack formation and crack growth in concrete by means of fracture mechanics and finite elements. *Cement and Concrete Research*, 6, 1976 :773-82
- Li, W. & Siegmund, T. (2002). An analysis of crack growth in thin-sheet metal via a cohesive zone model. *Engineering Fracture Mechanics*, 69:2073-93.
- Mirzaei, M., Mazaheri, K., & Biglari, H. (2005). Analytical modeling of the elastic response of tubes to internal detonation loading. *International Journal of Pressure Vessels and Piping*, Vol.82, No. 12, 883-895.
- Mirzaei, M., Biglari, H., & Salavatian, M. (2006a). Analytical and numerical modeling of the elastodynamic response of a cylindrical tube to internal gaseous detonation. *International Journal of Pressure Vessels and Piping*, Vol.83, No.7, 531-539.

- Mirzaei, M., Karimi, R. (2006b). Crack growth analysis for a cylindrical shell under dynamic loading., *Proceedings of ASME PVP 2006 / 11th International Conference on Pressure Vessel Technology, ICPVT-11*, Vancouver, Canada.
- Mirzaei, M., Salavatian, M., & Biglari, H. (2006c). Simulation of fatigue crack growth in a detonation tube. *Proceedings of ASME PVP 2006 / 11th International Conference on Pressure Vessel Technology, ICPVT-11*, Vancouver, Canada.
- Mirzaei, M. (2008a). Failure analysis of an exploded gas cylinder. *Engineering Failure Analysis*, Vol.15, No.7, 820-34.
- Mirzaei, M. (2008b). On amplification of stress waves in cylindrical tubes under internal dynamic pressures. *International Journal of Mechanical Sciences* Vol.50, No.8, 1292-1303.
- Mirzaei, M., Harandi, A., & Karimi, R. (2009). Finite element analysis of deformation and fracture of an exploded gas cylinder. *Engineering Failure Analysis* Vol.16, 1607-1615.
- Naitoh, M., Kasahara, F., Kubota, R., & Ohshima, I. (2003). Analysis of pipe rupture of steam condensation line at Hamoaka-1, (I) accumulation of non-condensable gas in a pipe. *Journal of Nuclear Science and Technology*, Vol.40, No.12, 1032.
- Price, W.H. (2006). An acetylene cylinder explosion: A most probable cause analysis. *Engineering Failure Analysis*, Vol.13, No., 705-715.
- Reismann, H. (1965). Response of a pre-stressed cylindrical shell to moving pressure load. *Proceedings of Eighth Midwest Mechanics Conference*, S. Ostrach and R. Scanlon, Eds., Pergamon Press, 1965, pp. 349-363.
- Rice, JR. (1979). The mechanics of earthquake rupture. *Proceedings of the international school of physics "Enrico Fermi"*, Course 78, 1979. Italian Physical Society/North-Holland; 1980: 555-649.
- Soto, O., Baum, J., & Löhner, R. (2010). An efficient fluid-solid coupled finite element scheme for weapon fragmentation simulations. *Engineering Fracture Mechanics*. Vol.77, 549-564.
- Tang, S. (1965). Dynamic response of a tube under moving pressure. *Proceedings of the American Society of Civil Engineers*, vol. 5, Engineering Mechanics Division, 1965, pp. 97-122.

FE Analysis of Evolution of Defects during Rolling

Dr. YU Hai-liang^{1,2}

¹Key Laboratory for Advanced Materials Processing Technology of Ministry of Education,
Department of Mechanical Engineering, Tsinghua University,
Beijing 100084
China

² State Key Laboratory of Rolling and Automation, Northeastern University,
Shenyang 110004
China

Finite element method (FEM) has been widely employed for simulation of rolling problems, such as thermal field [1,2], stress-strain field [3], microstructure distribution [4], rolling force [5], rolling pressure distribution [6], plate view shape [7,8], roll deflection [9,10], tracing the evolution of macrosegregation [11]. As the development of FEM, there are three important ways for the application of FEM in analysis of rolling problems recently. Firstly, rapid FEM, and its on-line application [12-14]; secondly, coupled multi-scale FEM, macro - FEM & micro - FEM & crystal - FEM, coupled the macro-deformation and microstructure evolution and texture distribution [15-16]; third, tracing the defects in whole rolling processes [11, 18-21]. Defects might appear in materials, such as cracks, inclusions. The evolution behavior of defects in steels during rolling severely affects the rolled products quality. In this chapter, the author summed up the previous researches on the evolution of cracks and inclusions during rolling, mainly contains the researches on the evolution of surface cracks by 2D thermo-mechanical FEM, the evolution of surface cracks during vertical-horizontal (V-H) rolling by 3D FEM, the evolution of internal cracks during V-H rolling, the evolution of inclusions during flat rolling.

1. Evolution of surface cracks by 2D thermo-mechanical FEM

Since the brittle fracture model proposed by Griffith last century, a large number of researchers and engineers have carried out a lot of work that focuses on the behavior of cracks. The research reports on the appearance and the propagation of cracks in rolled steels during rolling were investigated [22~25].

Considerable investigations have been carried out on applying FEM for simulation of the propagation and closure of cracks in materials during rolling. A self-healing shape memory alloy (SMA) composite was simulated via a finite element approach that allows crack to propagate in a brittle matrix material by Burton, et al [26]. The SMA wires were carefully

modeled using a one-dimensional SMA constitutive model and implemented into user-defined truss elements. Loading of the composite allowed a crack to propagate from an initiation site and the wires bridge the crack as detwinned martensite forms with the applied loading. Awais [27] and Son et al [28] employed the two-dimension FEM and Processing Map to analyze the closure and growth of surface crack in bars in the rolling process. Ervasti, et al [29, 30] simulated the closure and growth of longitudinal and transversal cracks in flat rolling process, and analyzed the closure and growth of cracks under a variety of the crack sizes, roll radii, friction coefficients, etc. Yukawa, et al [31, 32] analyzed the deformation of the micro-cracks, the foreign bodies pressing and so on in rolling process by a two-dimensional (2D) rigid-plastic FE code developed by them. In this part, the researches on the evolution of surface cracks by 2D thermo-mechanical FEM are carried out.

1.1 Finite element method

1.1.1 Basic theory

Eq. (1) [33] shows the calculation equation of thermal distribution in an isotropic element with an internal heat source.

$$\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2} + \frac{\dot{q}}{k} = \frac{\rho c}{k} \frac{\partial T}{\partial t} \quad (1)$$

where ρ is density; c is specific heat; k is coefficient of heat conductivity; \dot{q} is internal heat source strength, which is the plastic work done of slab deformation during rolling.

$$\dot{q} = m \int_0^{\dot{\epsilon}} \bar{\sigma} d\dot{\epsilon} \quad (2)$$

where m is coefficient of heat transform by plastic work done, $m = 0.7 \sim 0.95$ [34].

During hot rolling, the thermal transfers between the slab surface and the external environment contain two ways: heat emission and convection current. When the slab is not in the deformation zone, the heat radiation is much larger than the heat convection which could be neglected. According to the Stefan-Boltzman equation,

$$Q = H_r A (T - T_\infty) \quad (3)$$

where $H_r = S_B B_S (T^2 + T_\infty^2)(T + T_\infty)$, S_B is the coefficient of Stefan-Boltzman; B_S is the blackness on slab surface, T is the slab surface temperature; T_∞ is the temperature of environment.

The heat transfers between the slab surface and the work roll is calculated by the Eq. (4).

$$Q = h_i A (T - T_R) \quad (4)$$

where h_i is the coefficient of convective heat transfer between the slab and the roll; T_R is the temperature of roll.

The coupled thermal-mechanical method is solved by dynamic explicit approach [35].

The heat transfer equations are integrated using the explicit forward-difference time integration rule:

$$\theta_{i+1}^N = \theta_i^N + \Delta t_{i+1} \dot{\theta}_i^N \quad (5)$$

where θ^N is the temperature at node N and the subscript i refers to the increment number in an explicit dynamic step. The values of $\dot{\theta}_i^N$ are computed at the beginning of the increment by

$$\dot{\theta}_i^N = (C^{NJ})^{-1} (P_i^J - F_i^J) \quad (6)$$

where C^{NJ} is the lumped capacitance matrix, P^J is the applied nodal source vector, and F^I is the internal flux vector.

The equations of motion for the body are integrated using the explicit central-difference integration rule

$$\dot{u}_{i+\frac{1}{2}}^N = \dot{u}_{i-\frac{1}{2}}^N + \frac{\Delta t_{i+1} + \Delta t_i}{2} \ddot{u}_i^N \quad (7)$$

$$u_{i+1}^N = u_i^N + \Delta t_{i+1} \dot{u}_{i+\frac{1}{2}}^N \quad (8)$$

where u^N is a degree of freedom (a displacement or rotation component) at node N. The accelerations at the beginning of the increment are computed by

$$\ddot{u}_i^N = (M^{NJ})^{-1} (P_i^J - F_i^J) \quad (9)$$

where M^{NJ} is the mass matrix, P^J is the applied load vector, and F^J is the internal force vector.

1.1.2 Parameters and FE model

Fig.1 shows the schematic drawing of slab rolling with a surface crack. In the models, the work roll diameter (D_W) is 1 150 mm, the slab thickness before rolling (H_0) is 250 mm, and the slab thickness after rolling (H_1) is 230 mm. The cracks are assumed to be V-shaped. There are two important parameters for definition of the crack: crack height (H_C), and crack open-angle (θ), whose effects on the behavior of cracks during rolling are analyzed. The values of H_C are assumed to be 2 mm, 4 mm, 6 mm, 8 mm, and 10 mm respectively. The values of θ are assumed to be 5°, 10°, 20°, 30°, 40° separately.

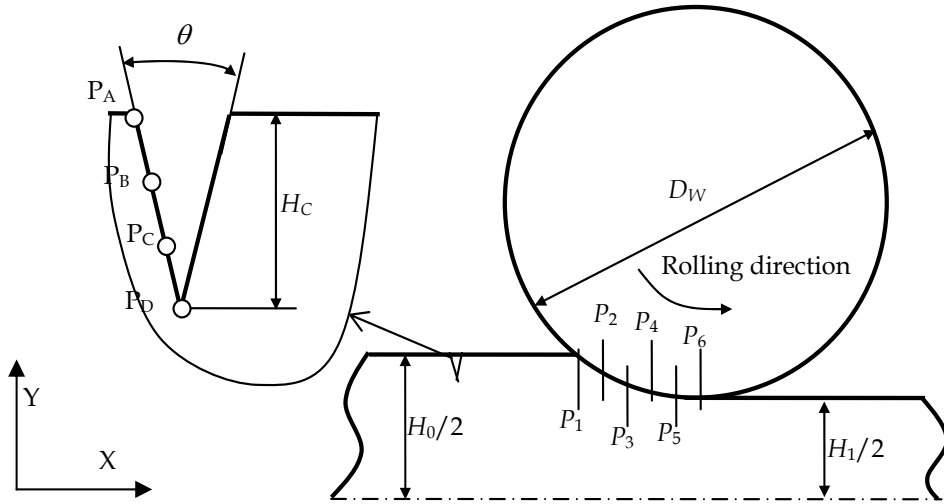


Fig. 1. Schematic drawing of slab rolling with a surface crack

The Coulomb friction model is used for solving the pressure between the crack surfaces and that between the slab and the roll. The influence of friction factor between the slab and the roll on the behavior of cracks during rolling are analyzed. The friction coefficients are assumed to be 0.2, 0.25, 0.35, 0.45, and 0.55.

The heat transfer between the slab and the roll, and between the crack surfaces are appeared, where the heat conductivity is assumed to be constant 46 W/m·K. During rolling, the deformation resistance of slab is related with the strain, strain rate, and temperature. The roll is assumed to be rigid, and the slab is isotropic bilinear model. The main material parameters are shown in **Table 1**.

Parameters	Roll	Slab
Density, kg/m ³	7830	7830
Young's modulus, GPa	210	117
Poisson's ratio	0.3	0.36
Deformation resistance, MPa	-	$\sigma = A\epsilon^B \dot{\epsilon}^{CT+D} e^{FT}$
Heat capacity, J/(kg·K)	460	460

Table 1. Material parameters (σ is stress; ϵ is true strain; $\dot{\epsilon}$ strain rate; A, B, C, D, and F are constants)

Owing to the symmetry of the rolls and the slab, a half of rolling geometrical model is employed. During rolling, the width spread of slab in a local zone is small, where the behavior of cracks could be approximate as the 2D deformation, so a 2D geometrical model is used. With the parameters above, the models of slab rolling with a surface crack are established, which are meshed by the quadrilateral elements. The FE meshing of slab with a surface crack during rolling is shown in **Fig.2**. The nodes on the bottom of slab are constrained for displacement along slab thickness direction, $U_Y = 0$. During rolling, the slab enters the roll with an initial velocity, and exits under the action of friction force between the roll and the slab.

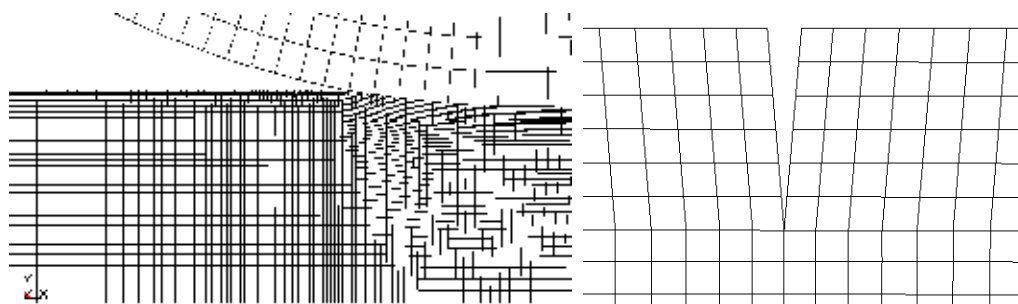


Fig. 2. Meshing of slab with a surface crack during rolling

1.2. Results and discussion [36]

1.2.1 Stress around crack tip during rolling

Fig. 3 shows the stress (σ_x) distribution around the crack at different positions in the deformation zone during rolling when the initial crack height is 10 mm, the initial crack open-angle is 10° and the friction coefficient is 0.35. In the bite zone, the compressive stress appears around the crack, and the crack gradually close, as shown in **Fig.3** (a) ~ (c). In **Fig.3** (d), the stress around crack is small, and the crack keeps its shape. When the crack is in the

forward zone, the tensile stress appears around cracks, and when the crack is at P_6 , the maximum tensile stress around crack tip attaches to 900 MPa, as shown in Fig.3 (e) and (f). It is obvious that the larger the compressive stress is, the more easily the cracks close, and the larger the tensile stress is, the more easily the cracks propagate. The maximum compressive stress and the maximum tensile stress at crack tip during rolling are further analyzed as follows.

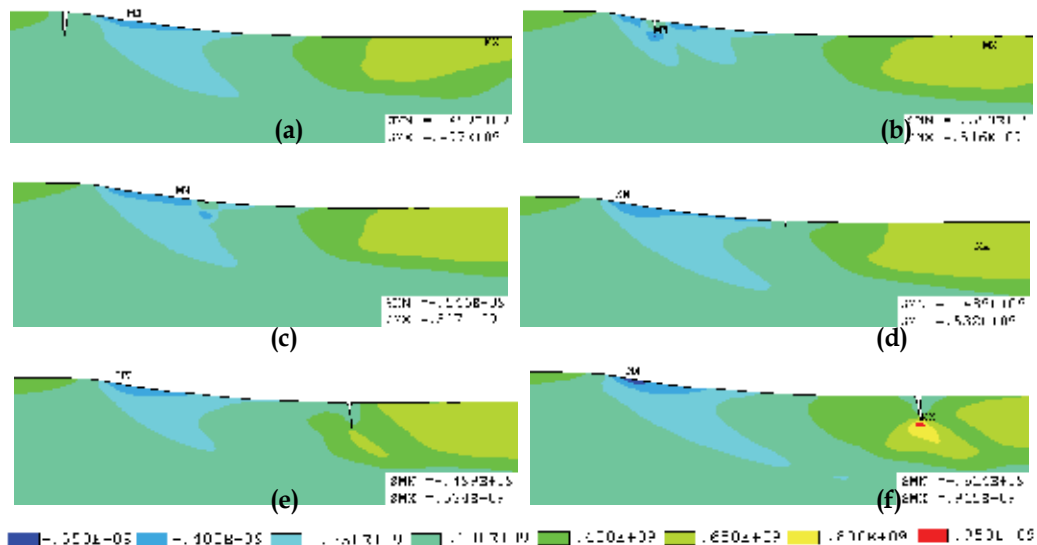


Fig. 3. Stress distribution around the crack at P_1 (a), P_2 (b), P_3 (c), P_4 (d), P_5 (e) and P_6 (f) in the deformation zone

Fig. 4 shows the maximum compressive and tensile stress at crack tip during rolling under various initial crack heights when the friction coefficient is 0.35 and the initial crack open-angle is 10° . As the initial crack height decreases, the maximum compressive stress increases. By contraries, the maximum tensile stress increases as the initial crack height increases.

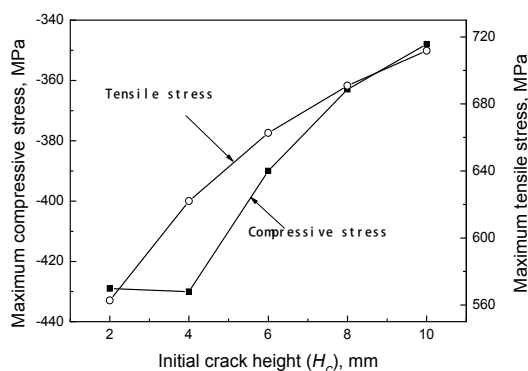


Fig. 4. Maximum compressive and tensile stress at crack tip during rolling for various H_c

Fig. 5 shows the maximum compressive and tensile stress at crack tip for a variety of initial crack open-angles when the friction coefficient is 0.35 and the initial crack height is 10 mm. With increasing the initial crack open-angle, the compressive stress increases. Under the rolling conditions, the maximum tensile stress decreases as the initial crack open-angle increases when the crack open-angle is 5~20°, which increases when the crack open-angle is between 20~40°.

Fig. 6 shows the maximum compressive and tensile stress at crack tip under different friction coefficients when the initial crack height is 10 mm and the initial crack open-angle is 10°. The maximum compressive stress increases with decreasing the friction coefficients. The maximum tensile stress increases with increasing the friction coefficients when the friction coefficient is from 0.2 to 0.45.

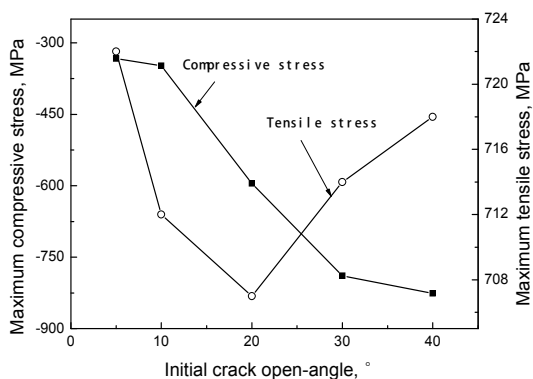


Fig. 5. Maximum compressive and tensile stress at crack tip during rolling for various θ

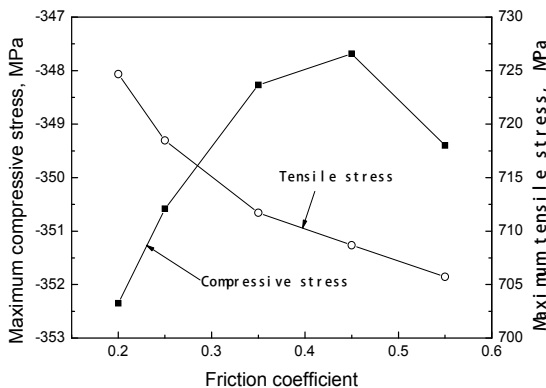


Fig. 6. Maximum compressive and tensile stress at crack tip during rolling for various friction coefficients

Compared the stress value in Figs.4, 5 and 6, the influence of the initial crack open-angle on the maximum compressive stress is the most obvious, whose difference is up to 450 MPa, and the influence of the initial crack height on the maximum tensile stress is the most obvious; the influence of the friction coefficient on the compressive stress is the slightest, and the influence of the initial crack open-angle on the maximum tensile stress is the slightest.

1.2.2 Crack open-angle and height after rolling

Fig. 7 shows the crack open-angle and the crack height after rolling under various the initial crack heights when the initial crack open-angle is 10° and the friction coefficient is 0.35. After rolling, the crack open-angle increases with reducing the initial crack height, and the crack height after rolling is direct ratio to the initial crack height.

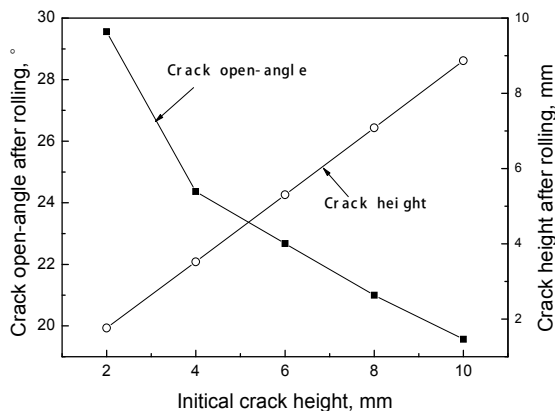


Fig. 7. Crack open-angle and height after rolling for various H_C

Fig. 8 shows the crack open-angle and height after rolling under a variety of initial crack open-angle when the initial crack height is 10 mm and the friction coefficient is 0.35. As the initial crack open-angle increases, the crack height after rolling decreases and the crack open-angle after rolling increases.

Fig. 9 shows the crack open-angle and height after rolling under various friction coefficients when the initial crack open-angle is 10° and the initial crack height is 10 mm. With increasing the friction coefficient, the crack open-angle after rolling increases and the crack height decreases.

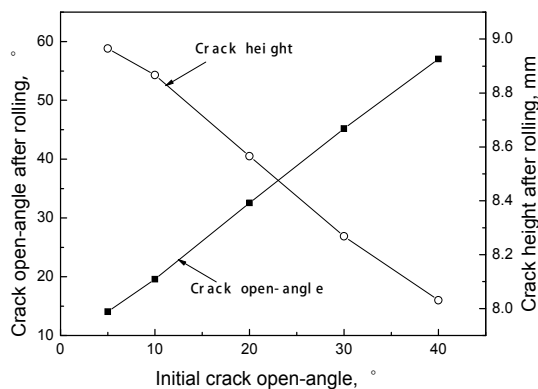


Fig. 8. Crack open-angle and height after rolling for various θ

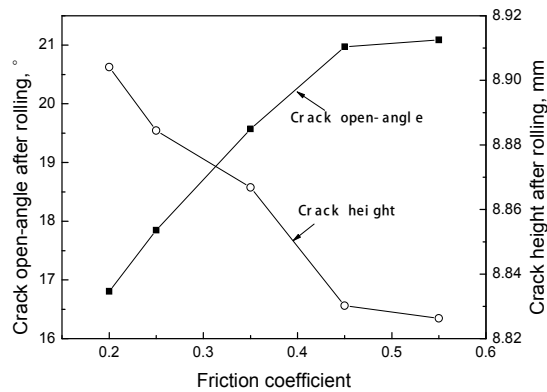


Fig. 9. Crack open-angle and height after rolling for various friction coefficients

1.2.3 Thermal distribution around crack surfaces during rolling

Fig. 10 shows the thermal distribution around the crack when the initial crack height is 10 mm, the initial crack open-angle is 10° and the friction coefficient is 0.35. After rolling, the temperature on crack surface is much less than that in slab matrix. It is obvious that the higher the temperature on crack surface, the more easily the behavior of crack metallurgical healing occurs under the compressive stress. So the thermal distribution around the crack will severely affect the behavior of cracks during rolling.

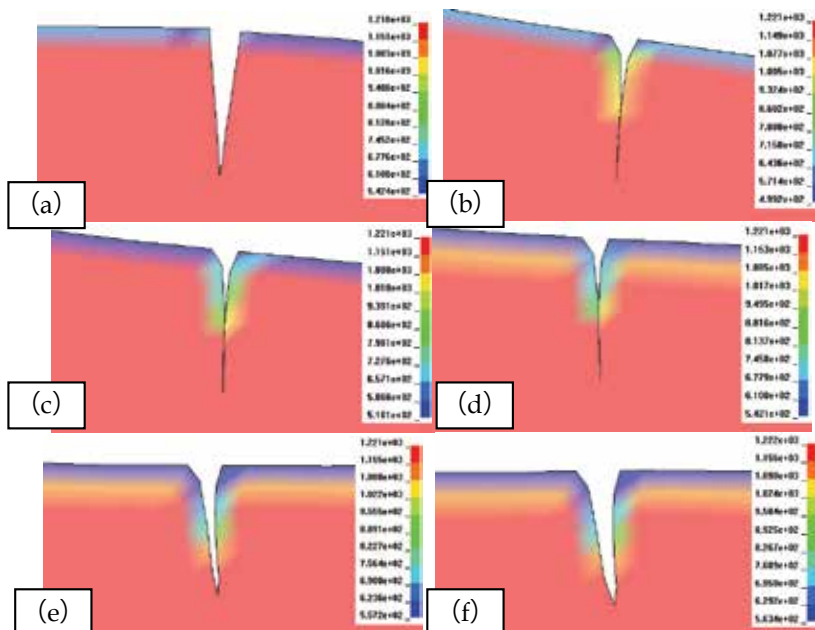


Fig. 10. Thermal distribution around crack at P_1 (a), P_2 (b), P_3 (c), P_4 (d), P_5 (e) and P_6 (f) during rolling ($^\circ\text{C}$)

Fig. 11 shows the temperature variation curve on crack surfaces during rolling under various initial crack heights when the initial crack open-angle is 10° and the friction coefficient is 0.35. In Fig.11 (a), the temperature in P_A changes slightly under various the crack height. As the position approaches to the crack tip, the difference between the temperatures increases. In Fig.11 (b), the temperature in P_B , when the crack height is larger than 8 mm, the temperature changes slightly, and when the crack height is less than 4 mm, the minimum temperature changes slightly and the difference increases as the rolling time increases.

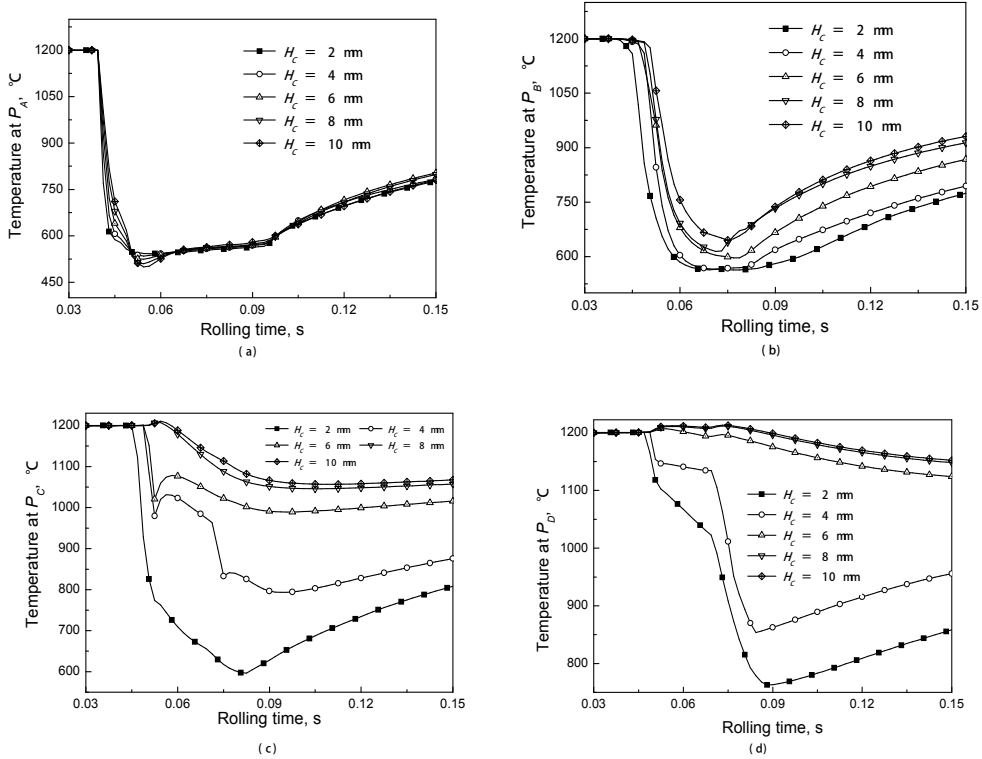


Fig. 11. Temperature at P_A (a), P_B (b), P_C (c) and P_D (d) during rolling for various H_c

In Fig.11(c), when the initial crack height is larger than 4 mm, the temperature increases in the deformation zone, and the temperature directly decreases to minimum temperature in the deformation zone when the initial crack height is 2 mm. In Fig.11 (d), there are two peaks of temperature in the deformation zone when the initial crack height is larger than 6 mm, and the temperature variation is less than 100 °C. When the initial crack height equals to 2 mm, the temperature variation approaches 450 °C.

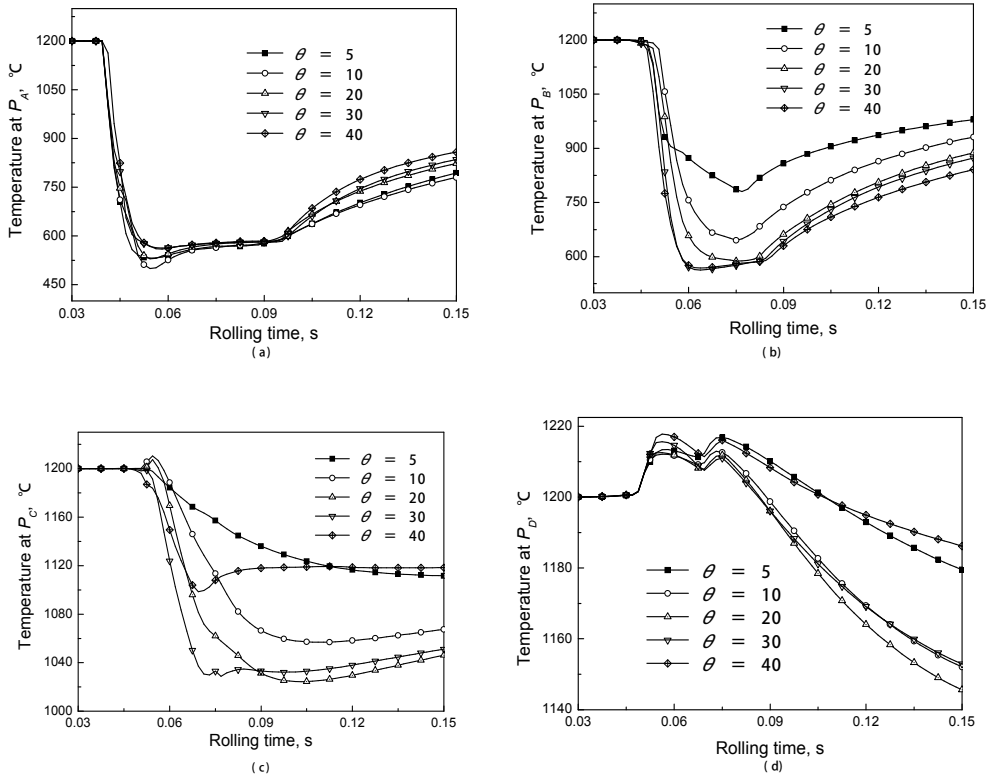


Fig. 12. Temperature at P_A (a), P_B (b), P_C (c) and P_D (d) during rolling for various θ

Fig. 12 shows the temperature variation curve on crack surface under a variety of initial crack open-angles when the initial crack height is 10 mm and the friction coefficient is 0.35. In Fig.12 (a), the temperature varies slightly in the deformation zone, and after rolling, the heating-up speed increases as the crack open-angle increases when it is greater than 10° . In Fig.12 (b), the temperature drop decreases with decreasing the initial crack open-angle. The minimum temperature is about 800°C when the initial crack open-angle is 5° , and the minimum temperature is about 550°C when the initial crack open-angle is 40° , the difference attaches to 250°C . In Fig.12(c), the temperature drop increases with increasing the initial crack open-angle when they are less than 20° , which decreases when they are larger than 20° , that is caused by the plastic work done. In Fig.12 (d), the temperature variation among them is less than 100°C , and they have the same regularity as that in Fig.12(c).

Fig. 13 shows the temperature variation curve on crack surfaces during rolling under various friction coefficients when the initial crack height is 10 mm and the initial crack open-angle is 10° . The temperature on crack surface varies slight during rolling except that the friction coefficient is 0.2. And the temperature rises with increasing the friction coefficient.

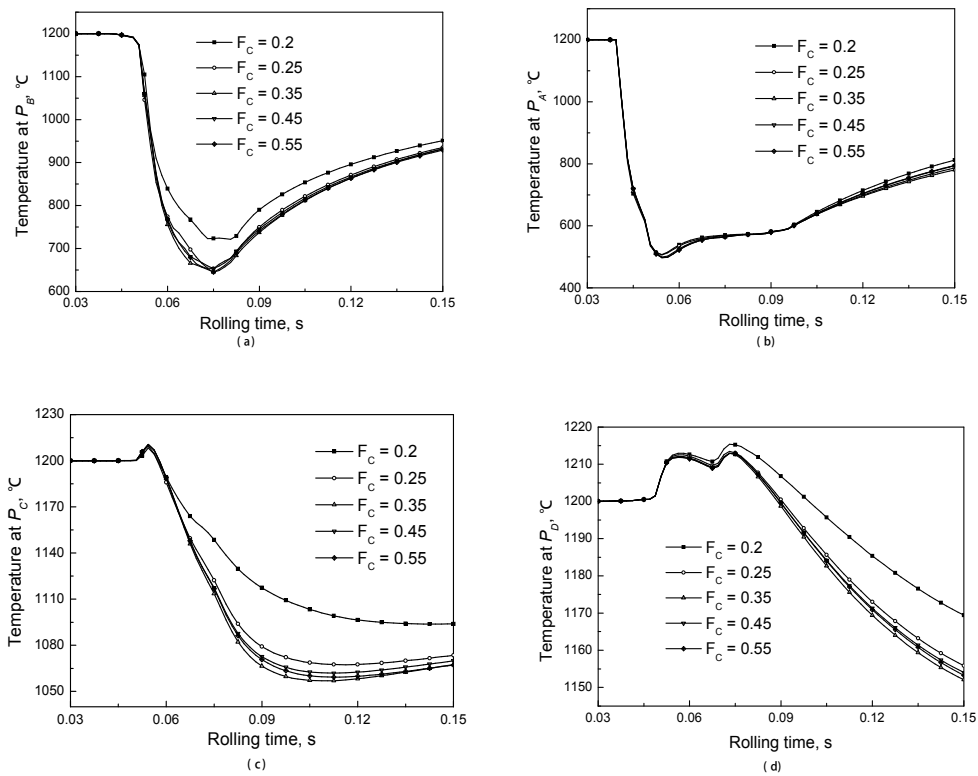


Fig. 13. Temperature at P_A (a), P_B (b), P_C (c) and P_D (d) during rolling for various friction coefficients

From the results above, the cracks close in the bite zone of deformation zone during rolling. The compressive stress appears on the crack surfaces. Then the crack might be healing under the conditions. In the forward slipping zone of deformation zone during rolling, the tensile stress appears around the cracks. When the tensile stress is less than the metallurgical bonding strength, the cracks will not open again, so after a pass, part of cracks heals, then as the rolling passes increase, the cracks gradually close after rolling, as shown in Fig.14(a). When the tensile stress makes part of healing zone open again, the position of P_X might move to the slab surface, after lots of rolling passes, part of cracks heals, and part of cracks becomes the slab surfaces, as shown in Fig.14 (b). When the tensile stress makes the healing part open again, and the crack open-angle is larger than the crack open-angle before rolling, and the crack height decreases after rolling. As the rolling pass increases, the crack surfaces become the slab surface wholly, as shown in Fig.14(c).

The cracks might heal when the temperature is enough high [37, 38], and the influence of thermal and stress distribution around cracks is very important. When the temperature of crack surface is quite low, the metallurgical bonding strength on crack surfaces will decrease severely. As shown in Fig. 11, the temperature distribution on the whole crack surface during rolling is lower than 800 °C when the initial crack height is 2 mm, so the cracks can not heal during rolling. Meanwhile, the crack open-angle after rolling is 29.5° whose initial

crack open-angle is 10° . As the rolling passes increase, the cracks open whole and become part of slab surfaces, as shown in Fig.14 (c). As shown in Fig. 12, the temperature of in P_C and P_D are larger than 1000°C , and the cracks might repair themselves under compressive stress, as shown in Fig.14 (b), even the whole cracks close as shown in Fig.14 (a).

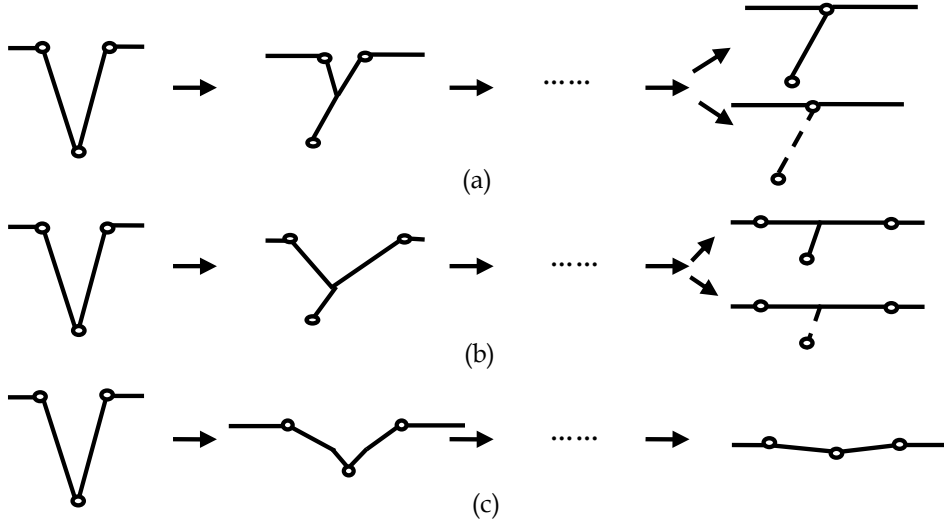


Fig. 14. Schematic of behavior of cracks on slab surface during rolling

2. Evolution of surface cracks on slab corner and edge by 3D FEM

The vertical-horizontal (V-H) rolling has been widely used for adjustment of the slab width for connection between the continuous casting and the finish rolling process. The deformation behavior of slab during rolling affects the quality of rolled steel. The FEM has been widely used for simulation of the behavior of slab during V-H rolling. The behavior of slab during V-H rolling was simulated by Xiong et al with the thermal-mechanical coupled rigid-plastic FEM [39], the thermal distribution and the strain distribution of slab during rolling were obtained, and the unsteady head and tail of slab were analyzed which were in good agreement with those of experimental ones. The three-dimensional rigid plastic/visco-plastic FEM was employed to simulate the behavior of slab during vertical rolling, the slab shape and the rolling force during rolling were obtained that were in good agreement with the experimental ones [40]. The behavior of slab during multi-pass V-H rolling was simulated by the author, and the shape of head and tail of slab during horizontal rolling with different dog-bone shape was also analyzed [41]. In this part, the researches on the evolution of surface cracks on slab corner and edge by 3D FEM were carried out.

2.1 Basic methods of analysis and assumptions

2.1.1 Explicit dynamic FEM

The fundamental equation of explicit dynamic element method [42] is

$$M\ddot{u} + C\dot{u} + F = R \quad (10)$$

Compared with static analysis, the mass matrix and damping matrix are introduced into equilibrium equation of dynamic analysis because of the existence of inertia force and damping force. The final solution equations are ordinary differential equation groups instead of algebraic ones. Generally, the Eq. (10) can be obtained by integrating central difference interpolation, which belongs to immediate integration method. In this method, the discrete velocity and acceleration at t_n can be expressed as:

$$\dot{u}_n = \frac{1}{2\Delta t}(u_{n+1} - u_{n-1}) \quad (11)$$

$$\ddot{u}_n = \frac{1}{\Delta t^2}(u_{n+1} - 2u_n + u_{n-1}) \quad (12)$$

Substituting the Eq.(11) and Eq.(12) into Eq.(10), the discrete-time recurrence equation is obtained as following:

$$u_{n+1} = [M + \frac{\Delta t}{2}C]^{-1}[\Delta t^2(R_n - F_n) + 2Mu_n - (M - \frac{\Delta t}{2}C)u_{n-1}] \quad (13)$$

The central difference method is stable under certain condition that the time step Δt must be lower than Δt_{\min} . Otherwise, the algorithm will be unstable. Generally, according to the Courant-Friedrichs-Levy stability criterion, the critical time step value is:

$$\Delta t_{\min} = \frac{2}{w_{\max}} = \frac{l}{c} \quad (14)$$

For 3D element, the sound velocity in the material is:

$$c = \sqrt{\frac{(1-\nu)E}{(1+\nu)(1-2\nu)\rho}}. \quad (15)$$

2.1.2 Contact problem of crack surfaces

In rolling process, the contact problems between the slab and the rolls, between the crack surfaces are the surface to surface problems, which are solved with the penalty function algorithms. The contact stiffness (k) is determined by the Eq. (16)[43].

$$k = \frac{f_s \times A_E^2 \times K}{V_E} \quad (16)$$

Where, f_s is the penalty factor, which equals 0.1 here;

A_E is the area of contact segment;

V_E is the element volumes;

K is the bulk modulus of contacted element. Where $K = \frac{E}{3(1-2\sigma)}$, E is Young's

modulus, and σ is the Poisson's ratio.

2.1.3 Multi-pass rolling process simulation

The updating geometrical method[44] is employed to simulate the multi-pass rolling processes with a single pass rolling process FE model, which is that adding displacements from the previous analysis results and updating the geometry of the finite element model to

the deformed configuration. The updating geometrical process of rolling processes is shown in Fig. 15.

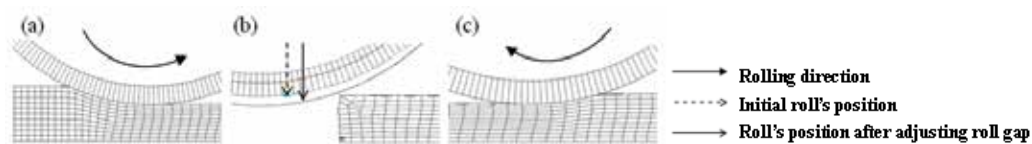


Fig. 15. Updating geometrical process of rolling process

2.1.4 Assumptions

In the research of the closure and growth of crack, there are two assumptions:

- (1) The crack is made up of two surfaces;
- (2) The crack is "V"-shape [29, 30] which exists on slab corner;
- (3) The height of the crack on top surface equals to that on side surface of slab.

2.2 Evolution of corner cracks during rolling

2.2.1 Basic parameters and rolling conditions

The closure and growth of crack during multi-pass V-H rolling process ($V_1-H_1-V_2-H_2-V_3-H_3$) [44] has been simulated. In the simulation, the diameter of horizontal roll is 1150mm. The edger roll of 980mm in diameter separately employs the flat edger roll and the grooved edger roll whose dimensions are shown in Fig.16, where the groove fillet radius R equals to 30, 50, 80 and 110mm. The initial dimension of slab profile is 1200×250mm. The "V"-shape cracks on slab corner are regarded as two surfaces. The crack size (crack height on slab top surface-crack height on slab side surface-crack width) employs 20-20-2 (Fig.17), 15-15-1.5, 10-10-1, 5-5-0.5mm respectively.

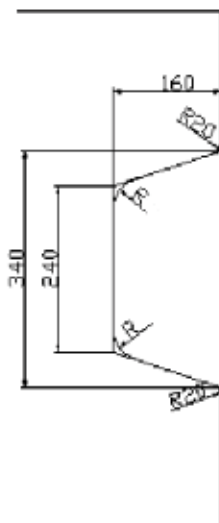


Fig. 16. Dimensions of grooved edger roll

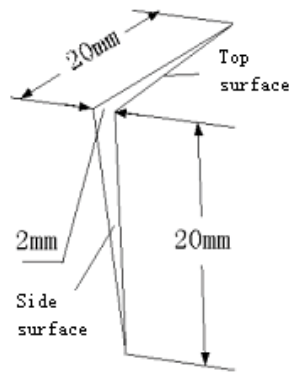


Fig. 17. Dimensions of one kinds of crack with “V”-shape

In this simulation, the rolls are assumed to be rigid and the slab is modeled as bilinear isotropic. The bilinear isotropic model is shown in Eq. (18).

$$\sigma_Y = \sigma_0 + E_p \varepsilon_p^{eff} \quad (18)$$

Where, σ_0 is the initial yield resistance;

ε_p^{eff} is the effective plastic strain;

E_p is the plastic hardening modulus.

The main chemical compositions of slab are (mass %): C 0.18, Si 0.32, Mn 0.82. The yield stress at the high temperature related to the true strain ε , the true strain velocity $\dot{\varepsilon}$, and the deformation temperature T . During this simulation, the initial yield stresses were computed by Eq. (19).

$$\sigma = A \varepsilon^B \dot{\varepsilon}^{CT+D} e^{FT} \quad (19)$$

where A, B, C, D, F are constants, as shown in Table 2.

Coefficient	A	B	C	D	F
Value	1715.706	0.17311	0.16952	0.05515	-0.267881

Table 2. Coefficients in the initial yield stress model

In the simulation, the main material parameters are shown in Table 3.

Parameters	Roll	Slab
Young's modulus, GPa	210	117
Poisson's ratio	0.3	0.36
Density, kg/m ³	7850	7850
Plastic hardening modulus, MPa	-	10.1
Initial yield stress, MPa	-	Eq. (19)

Table 3. Basic material parameters in rolling process

During rolling process, the draft of vertical rolling in every pass is 50mm, and the draft of horizontal rolling is 10mm. In this paper, there are 18 kinds of simulation schedules, which are shown in **Table 4**.

Schedule	E-roll shape	R/mm	Crack size /mm	f
1	Flat	—	20-20-2	0.25
2	Flat	—	20-20-2	0.35
3	Flat	—	20-20-2	0.45
4	Flat	—	20-20-2	0.55
5	Groove	50	20-20-2	0.25
6	Groove	50	20-20-2	0.35
7	Groove	50	20-20-2	0.40
8	Groove	50	20-20-2	0.43
9	Groove	50	20-20-2	0.45
10	Groove	50	20-20-2	0.47
11	Groove	50	20-20-2	0.50
12	Groove	50	20-20-2	0.55
13	Groove	50	15-15-1.5	0.35
14	Groove	50	10-10-1	0.35
15	Groove	50	5-5-0.5	0.35
16	Groove	30	20-20-2	0.35
17	Groove	80	20-20-2	0.35
18	Groove	110	20-20-2	0.35

Table 4. Simulation schedules and rolling conditions

2.2.2 Establishment of models

Owing to the symmetry of slab and rolls, 1/4 of slab and rolls were included in the geometric model. According to the parameters mentioned above, the rolling model of the first pass was built. The rolls were considered as rigid because their deformation could be neglected. The whole geometric model was meshed with 8 nodes and hexahedral elements. At the same time, it just refined the elements near the slab margin where the deformation assembled. The nodes on the bottom face of slab were constrained along Y direction, $U_y=0$; and the nodes on the center face of slab were constrained along Z direction, $U_z=0$. The geometrical model and meshing of slab with crack before rolling was shown in **Fig.18**. During simulation, the slab entered the rolls with an initial velocity and exited the rolls under the friction force. The rolling models of the 2nd and the 3rd passes were obtained by updating geometry, changing material attribution, boundary conditions and loads.

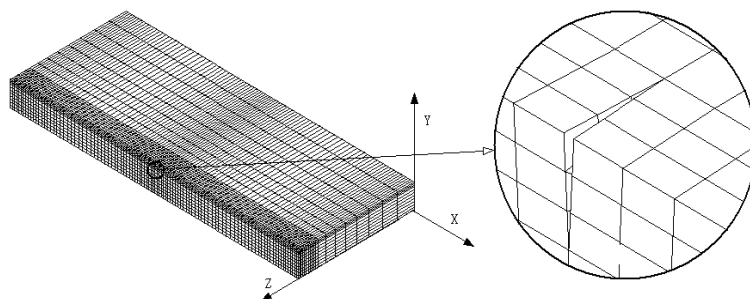


Fig. 18. Geometry and meshing of slab with crack

2.2.3 Results and analysis [45]

The deformation behavior of crack could be realized through the analysis of the change of the crack width during rolling process, which may be used to do the qualitative analysis of the closure and growth of crack. Meanwhile, when the crack closes, two surfaces of the crack will contact each other and appear contact pressure. The contact pressure between crack surfaces may be used to analyze the contact strength for crack surfaces which can be employed to analyze which method is more helpful for the crack closure. However, when the crack is growing or open, the contact pressure will become zero, then, it can't support help for researching. In this paper, the crack width and the contact pressure on crack surfaces at the exit stage of rolling are used to analyze the behavior of transversal crack on slab corner during multi-pass V-H rolling process.

(1) Influence of the friction coefficient on the crack's closure and growth

1> Flat edger roll

When the flat edger roll is used during multi-pass V-H rolling process, after horizontal rolling, the transversal crack on slab corner may close. During following it just analyzes the closure and growth of transversal crack during V_1 - H_1 rolling process. And the simulation schedules 1~4 as shown in Table 4 are adopted.

<1> Crack width

Fig.19 shows the influence of the friction coefficient on the crack width during V-H rolling process when the flat edger roll is used. In vertical rolling, the influence of friction coefficient on crack width is little, and when the friction coefficient is among 0.25~0.55, the cracks width all are less than 0.25mm. But during horizontal rolling process, the influence of the friction coefficient on crack width is very obvious. When the friction coefficient changes among 0.25~0.35, the cracks close, and when the friction coefficient changes among 0.35~0.55, the final crack width increases gradually with the increasing of the friction coefficient.

<2> Contact pressure

Fig.20 shows the influence of the friction coefficient on the contact pressure on crack surfaces on slab corner during V-H rolling process when the flat edger roll is used. It is very clear that the value of contact pressure in vertical rolling is much larger than that in horizontal rolling. During vertical rolling process, with the changing of friction coefficient, the contact pressure varies among 50~70MPa. And with the increasing of friction coefficient, the contact pressure decreases. During horizontal rolling process, with the increasing of friction coefficient, the contact pressure decreases acutely, from 45MPa to 0MPa, and the crack opens again.

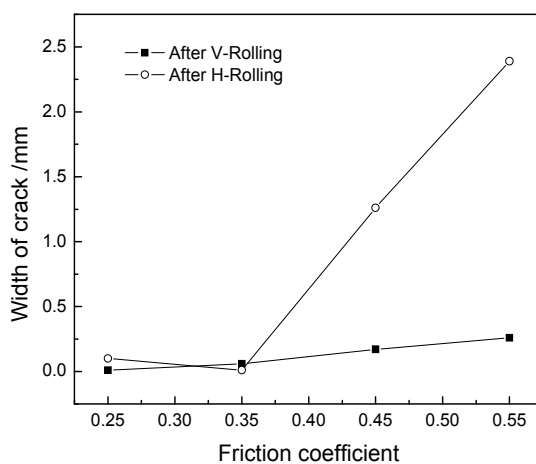


Fig. 19. Influence of the friction coefficient on the crack width after V1 and H1 for flat E-roll

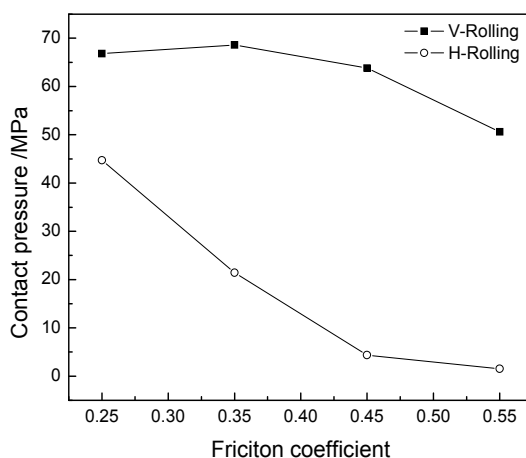


Fig. 20. Influence of the friction coefficient on the contact pressure during V1 and H1 for flat E-roll

2> Grooved edger roll

When the grooved edger roll is used, after horizontal rolling, the crack may be opened again. In the following, the closure and growth of crack is analyzed in the V_1 - H_1 - V_2 - H_2 - V_3 - H_3 rolling process; and the simulation schedules 5~12 as shown in Table 4 are adopted.

<1> Crack width

Fig.21 shows the influence of friction coefficient on crack width during V-H rolling process when the grooved edger roll is used. During vertical rolling process, the cracks close well, but after horizontal rolling process, the cracks all open again. Compared with the crack width when using the flat edger roll, the influence of friction coefficient on crack width is quite small.

<2> Contact pressure

In Fig.22, it could be found that the cracks will open again during horizontal rolling process when using the grooved edger roll. So in the following it just analyzes the influence of friction coefficient on contact pressure during vertical rolling process, as shown in Fig.22. In Fig.22, it could be found that with the increasing of the friction coefficient, the main tendency of contact pressure will decrease during V_1 , V_2 , V_3 rolling process, which isn't good for the closure of crack.

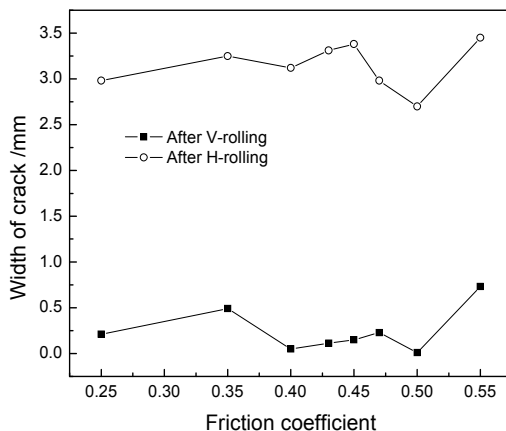


Fig. 21. Influence of the friction coefficient on the crack width after V1 and H1 for grooved E-roll

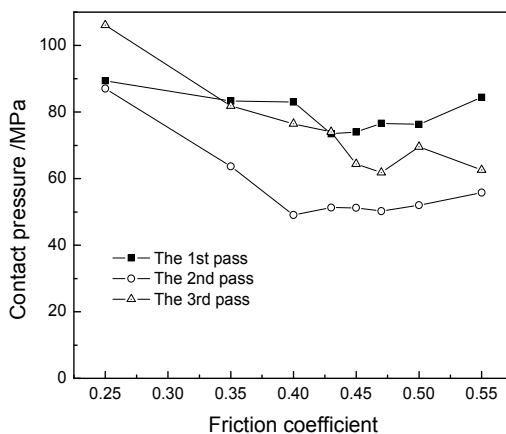


Fig. 22. Influence of the friction coefficient on the contact pressure for grooved E-roll

From the above analysis, it could be known that it is helpful for crack closure to decrease the friction coefficient between the slab and the rolls not only using flat edger roll but also grooved edger roll during multi-pass vertical rolling.

(2) Influence of the initial crack size on the closure and growth of crack

In this following the simulation schedules 6, 13, 14 and 15 are investigated to analyze the influence of crack size on the closure and growth of crack during multi-pass V-H rolling process.

1> Crack width

From above analysis, it is known that the transversal crack on slab corner will close during vertical rolling when the grooved edger roll is used. So in the following it just analyzes the crack width changing after horizontal rolling, which is shown in Fig.23. After H_1 rolling process, when the initial crack size is 20-20-2mm, the crack width is 3.25mm; when the initial crack size is 5-5-0.5mm, the crack width is 1.34mm. And when the crack size is between the former and the latter, the crack width will decrease as the initial crack size decreases after horizontal rolling process.

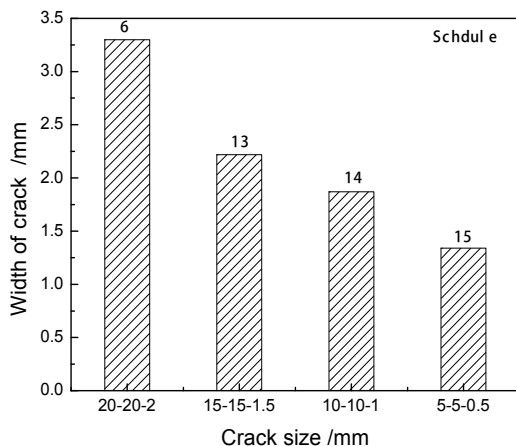


Fig. 23. Influence of the initial crack size on the crack width after H_1 for grooved E-roll and $f=0.35$

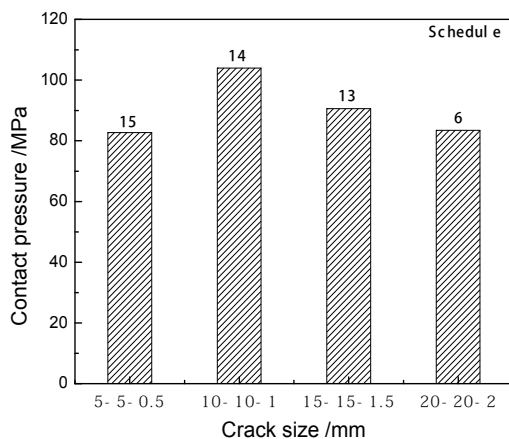


Fig. 24. Influence of the initial crack size on the contact pressure after V_1 for grooved E-roll and $f=0.35$

2> Contact pressure

The contact pressure becomes zero because the crack opens after horizontal rolling process

during V-H rolling process when the grooved edger roll is used. So it just analyzes the contact pressure during V_1 rolling process. **Fig.24** shows the influence of initial crack size on the contact pressure of transversal crack on slab corner. In Fig.24, it could be found that the contact pressure of transversal crack decreases as the initial crack size increases when the initial crack size is between 10-10-1mm and 20-20-2mm.

From above analysis it is known that the crack will be easy to close when initial crack size decreases during multi-pass V-H rolling process.

(3) Influence of edger roll shape on the closure and growth of crack

In the below, the multi-pass V-H rolling process (V_1 - H_1 - V_2 - H_2 - V_3 - H_3) were simulated according to the schedules 2 and 6 in Table 4, and the influence of edger roll shape on the closure and growth of crack was analyzed.

1> Crack width

Fig.25 shows the transversal crack width on slab corner after H_1 , H_2 , H_3 rolling process when using the flat edger roll and the grooved edger roll. In the figure, it could be found that the crack closes well for the flat edger roll and the crack opens again for the grooved edger roll. Meanwhile, the crack width decreases with the increasing of rolling pass.

2> Contact pressure

Fig.26 shows the contact pressure of transversal crack on slab corner in each pass both using the flat edger roll and using the grooved edger roll. From which it could be found that the contact pressure is among 70~80MPa during V_1 , V_2 , V_3 rolling process and 0MPa during H_1 , H_2 , H_3 when the grooved edger roll is employed. When employing the flat edger roll, there exists contact pressure in transversal crack on slab corner during V_1 , H_1 , V_2 , H_2 , V_3 , H_3 rolling process, but they all are less than 60MPa.

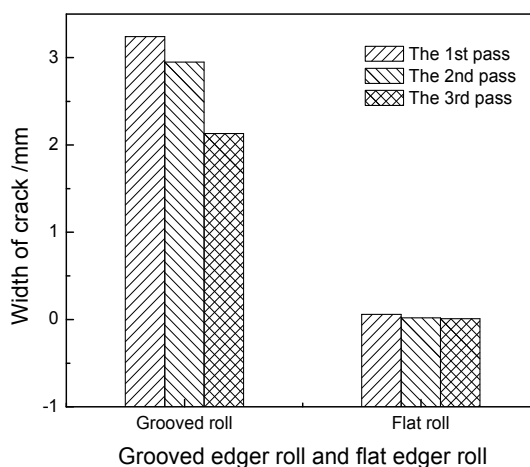


Fig. 25. Influence of the roll shape on crack width

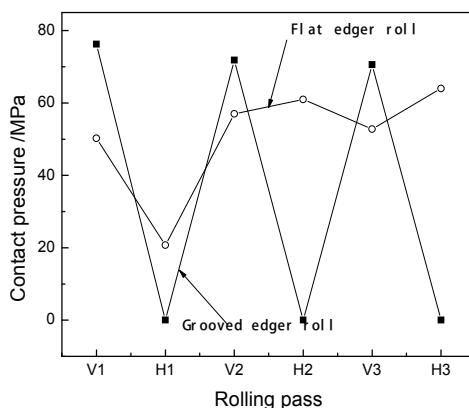


Fig. 26. Influence of the roll shape on contact pressure of crack

From above analysis it could be found that the flat edger roll is more helpful to crack closure than the grooved edger roll when the cracks on some steel grades are hard to heal.

(4) Influence of the groove fillet radius of grooved edger roll on the closure and growth of crack

In the following, the multi-pass V-H rolling process (V_1 , H_1 , V_2 , H_2 , V_3 , H_3) were simulated according to the schedules 6, 16, 17 and 18 in Table 4, and the influence of groove fillet radius on the closure and growth of crack was analyzed.

1> Crack width

Fig.27 shows the width of transversal crack on slab corner under different groove fillet radius of the grooved edger roll. From which it could be known that when the groove fillet radius $R=30\text{mm}$, the crack width is 2.55mm after H_1 rolling process, and it is 2.25mm after H_3 rolling process. With the increasing of R from 50mm to 110mm gradually, the crack width decreases after horizontal rolling process. And what is similar with the Fig.25 is that the crack width decreases as the rolling pass increases.

2> Contact pressure

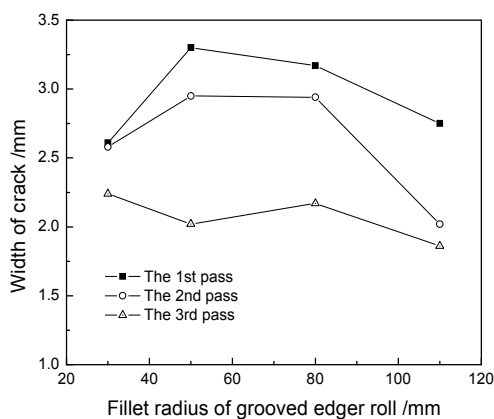


Fig. 27. Influence of the groove fillet radius of the grooved E-roll on crack width

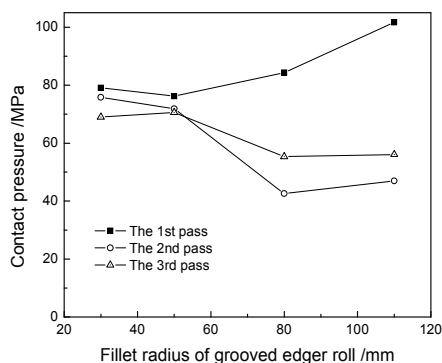


Fig. 28. Influence of the groove fillet radius of the grooved E-roll on the contact pressure

Fig.28 shows the values of contact pressure of transversal crack on slab corner during V_1 , V_2 , V_3 rolling process under different groove fillet radius of grooved edger roll. From the figure, it could be found that the value of contact pressure of transversal crack increases with the groove fillet radius increasing during V_1 rolling process (from 80MPa to 100MPa). But during V_2 and V_3 rolling process, the value of contact pressure of transversal crack decreases with the increasing of the groove fillet radius (from 80MPa to 50MPa). Meantime, if the value of groove fillet radius is relatively small ($R=30\sim50\text{mm}$), the value of contact pressure is quite stable during vertical rolling process. If the value of groove fillet radius is between 80 and 110mm, the change of the value of contact pressure is larger, from 100MPa to 50MPa. From the above analysis, it could be known that the crack width decreases as the groove fillet radius increases from $R=50$ to $R=110$ mm. Meantime, the value of contact pressure increases during V_1 rolling process and decreases during V_2 and V_3 rolling process as the groove fillet radius increases.

2.2.4 Experimental

(1) Experimental conditions

According to the schedule 1 (as shown in Table 4, the friction coefficient [45] between the lead and the rolls is about 0.2~0.28 in cold rolling), the experiment of the closure and growth of transversal crack during multi-pass V-H rolling process was done in our laboratory.

The $\Phi 300\text{mm}$ mill is adopted in the vertical rolling, and the $\Phi 180\text{mm}$ mill is adopted in the horizontal rolling. The pure lead of slab was used, and the profile of slab is $120\times 25\text{mm}$. The transversal crack with "V"-shape was curved on slab corner. It is shown in **Fig.29**. The height of crack on top surface is 5mm, which on side surface is 5mm, and the width of crack is 0.5mm.



Fig. 29. Experimental sample

During rolling process, according to the rolling schedule V_1 - H_1 - V_2 - H_2 - V_3 - H_3 the experiment is investigated. The rolling velocity is 200mm/s. The draft of vertical rolling in every pass is 5mm, and the draft of horizontal rolling is 1mm.


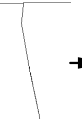
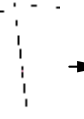
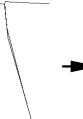
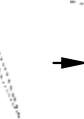
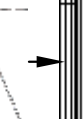






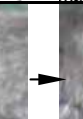



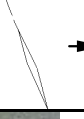
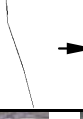





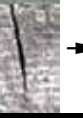
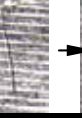
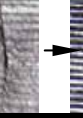
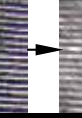

		Before rolling	V_1	H_1	V_2	H_2	V_3	H_3
Top surface	FEM							
	Experiment							
Side surface	FEM							
	Experiment							

Table 5. Behavior of crack between FEM results and experimental results

(2) Experimental results

The behavior of transversal crack on slab corner during multi-pass V-H rolling process through finite element simulation and experimental simulation is shown in **Table 5**. In the table, the closure and growth from the slab top surface and the slab side surface during every rolling pass could be obvious seen. It is clear that the calculated results have the same tendency of crack behavior with the experiment results using the pure lead.

2.3 Evolution of longitudinal crack on slab edge during rolling

2.3.1 Basic Parameters and Rolling Conditions

In the FE models, the following geometrical parameters are used: the radius of horizontal roll is 1 150 mm; the radius of edger roll is 980 mm employing the flat edger roll and the grooved edger roll respectively, and the fillet radii (R) of grooved edger roll are 30 mm, 50 mm, 80 mm, 110 mm separately, other dimensions of the grooved edger roll are shown in **Fig. 16**. The initial dimensions of slab profile are 250 mm \times 1 200 mm.

The V-shaped (crack width direction) cracks are carved on the upper surface of the slab, which are assumed to be disk-shaped along the crack length direction, as shown in **Fig.30** (X is the rolling direction, Y is the slab thickness direction, Z is the slab width direction). The main parameters of cracks (the crack width, the crack depth and the crack length (w - h - l)) are 2-5-20. In the simulation, it is assumed that the cracks do not weld when the cracks close, and they could be open again.

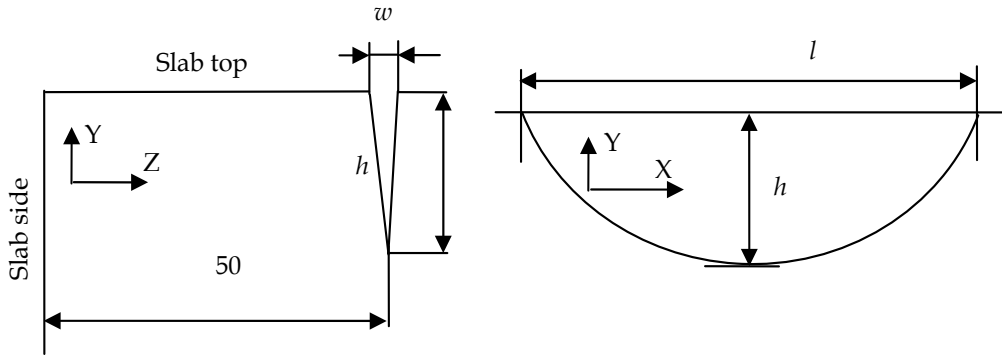


Fig. 30. Crack shape and size

2.3.2 FE Models

Owing to the symmetry of slab and rolls, a quarter of the slab and the rolls were considered in the geometrical model. The V-H rolling models of the first pass were built with the parameters above. The models were meshed with 8-noded hexahedral elements. It just refined the elements near the cracks and the slab margin where the deformation assembled. In rolling process, the edger roll and the horizontal roll rotate with a stable angular velocity, and the slab enters the rolls with an initial velocity and exits under the friction forces. The geometry and meshing of slab with a longitudinal crack on surface in slab edge before rolling is shown in Fig. 31. When building the rolling models of the subsequent passes (V_2-H_3), the step-by-step change in the slab was determined by updating geometry, modifying material attribution, boundary conditions, and loads according to the results of the earlier pass.

During rolling, the draft of vertical rolling is 50 mm each pass, and the outlet thickness after horizontal rolling is 240 mm. The roll rotating speed is 4 m/s. The friction coefficient (F_s) between the slab and the rolls employs three schemes: 0.2, 0.35 and 0.5 respectively.

The nodes on the surface where Y equals 0 are constrained with displacement,

$$U_{Y,Y=0} = 0 \quad (20)$$

The nodes on the surface where Z equals 0 are constrained with displacement,

$$U_{Z,Z=0} = 0 \quad (21)$$

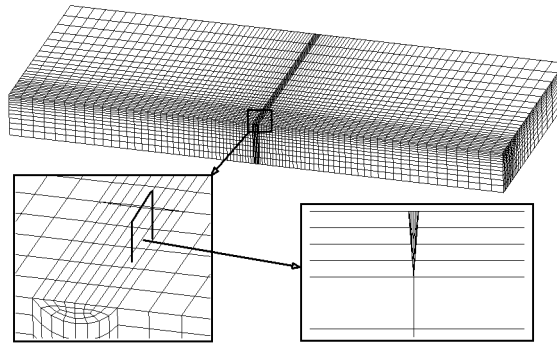


Fig. 31. The geometry and meshing of slab with a longitudinal crack

2.3.3 Results and discussion [46]

Table 6 shows the crack shape from top surface after each pass in multi-pass V-H rolling processes, where R equals 30 mm, 50 mm, 80 mm and 110 mm separately, and F_s equals 0.2, 0.35 and 0.5 separately.

During V-H rolling, the crack closes in vertical rolling process; however, the crack might open again after horizontal rolling process for some kinds of conditions. Meanwhile, with the increase of rolling pass, the crack gradually closes after horizontal rolling process.

When the fillet radii of grooved edger roll are same, the crack width increases as the friction coefficients between the slab and the rolls increase. Similarly, when the friction coefficients are identical, the crack width increases as the fillet radii of grooved edger roll increase.

Fig. 32 shows the influence of fillet radii of grooved edger roll on the crack shape in XY plane when F_s equals 0.35, (a) is after V_1 rolling process, and (b) is after H_1 rolling process. In the figure, the cracks shape in XY plane after rolling is similar to that before rolling process. After vertical rolling, the crack depth increases, meanwhile, which increases with the increase of the fillet radii of grooved edger roll. When R equals 30 mm, the crack depth is around 6 mm, and when R equals 110 mm, the crack depth is around 7.5 mm. Similarly, the crack depth increases as the fillet radii of grooved edger roll increase after the horizontal rolling.

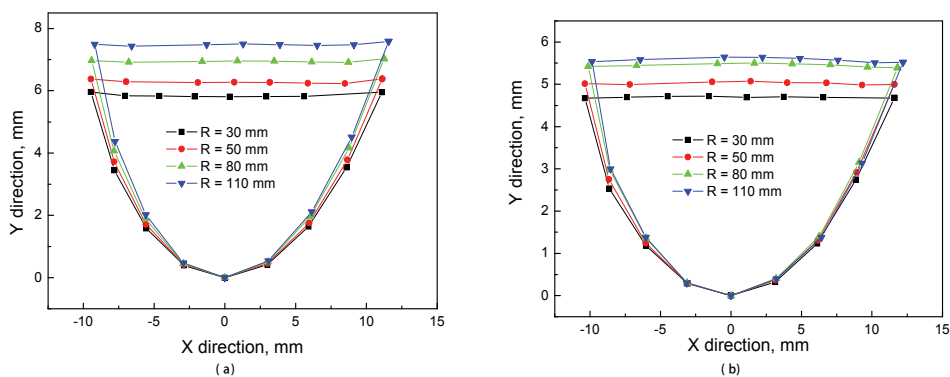


Fig. 32. Crack shape in XY plane for R when F_s equals 0.35 after V_1 rolling (a) and after H_1 rolling (b)

Fig. 33 shows the influence of the friction coefficient on the crack shape in XY plane when R equals 50 mm. The influence of friction coefficient on the crack shape in XY plane is little, and the crack depth increases with the increase of the friction coefficient.

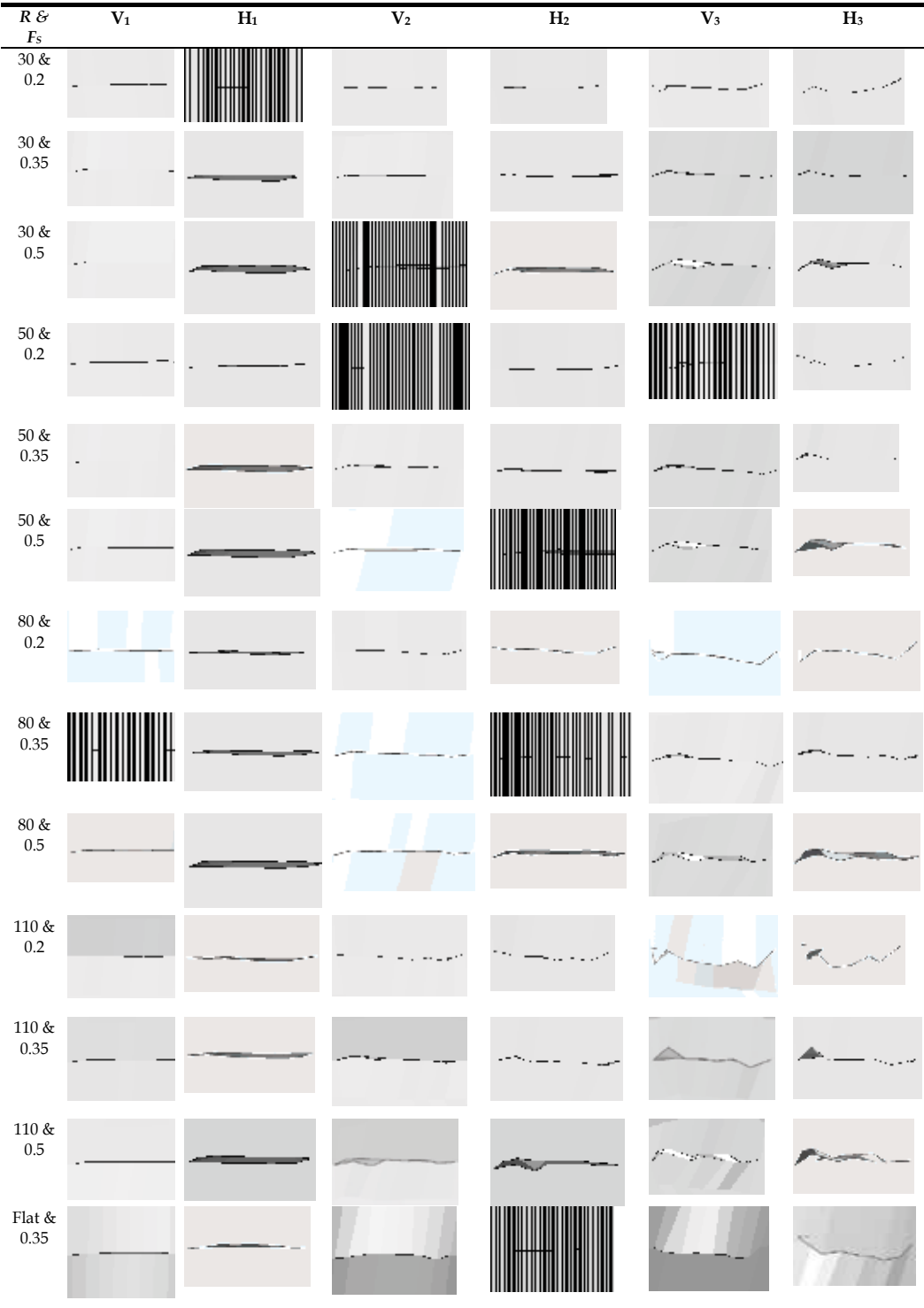


Table 6. Crack shape from top surface in multi-pass V-H rolling processes

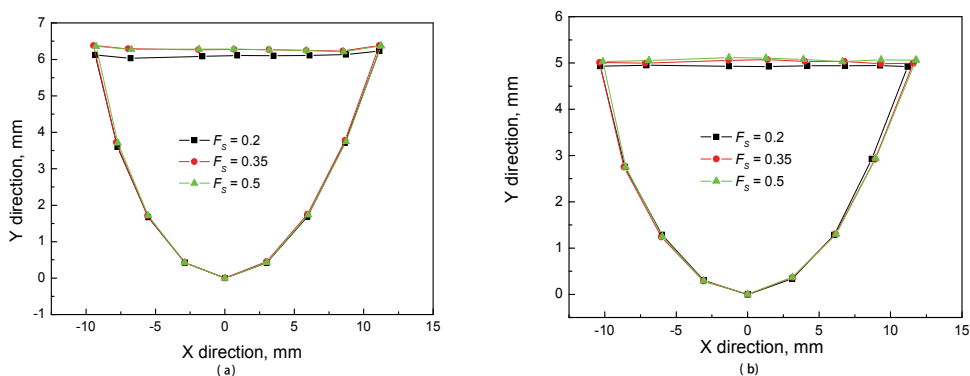


Fig. 33. Crack shape in XY plane for FS when R equals 50 mm after V1 rolling (a) and after H1 rolling (b)

Fig. 34 shows the influence of the fillet radii of grooved edger roll on the crack shape in YZ plane after H_1 rolling process when the F_s equals 0.35. The cracks close well near the crack tip, and they open near slab top surface. Meanwhile, the cracks directions slant to Y direction with a degree. The slant degree increases with the increase of the fillet radii of grooved edger roll during rolling process.

Fig. 35 shows the crack shape in YZ plane after horizontal rolling process when the R equals 50 mm under a variety of the friction coefficients. The cracks close in the zone near crack tip and open near the slab top surface. As the friction coefficient increases, the crack width increases.

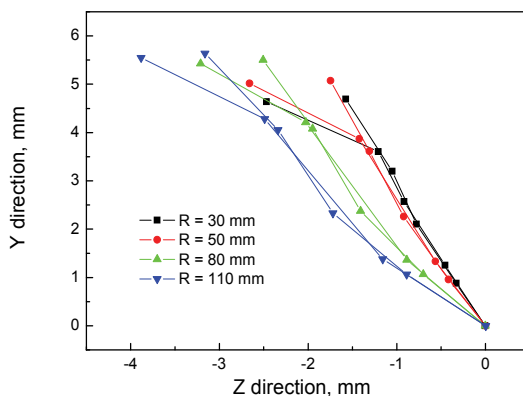


Fig. 34. Crack shape in YZ plane for R when F_s equals 0.35

Analysis of the crack length change regularities in rolling processes. **Fig. 36** shows the influence of fillet radii of grooved edger roll on the crack length elongation ratio (crack length after rolling/ crack length before rolling) in rolling process when the F_s equals 0.35. In V_1 and H_1 rolling process, the crack length elongation ratio increases with the increase of fillet radii, however, the crack length elongation ratio decreases with the increase of fillet

radii in $V_2 - H_3$ rolling processes. Meanwhile, the crack length elongation ratio when the grooved edger roll is used is larger than that when the flat edger roll is used.

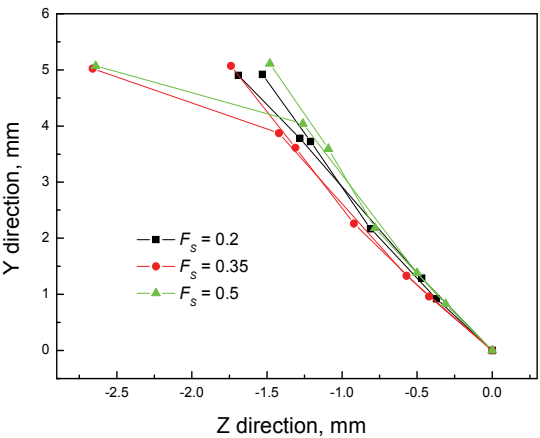


Fig. 35. Crack shape in YZ plane for FS when R equals 50 mm

Fig. 37 shows the influence of the friction coefficient between the slab and the rolls on the crack length elongation ratio when the fillet radius equals 50 mm. The crack length after rolling increases as the friction coefficient increases.

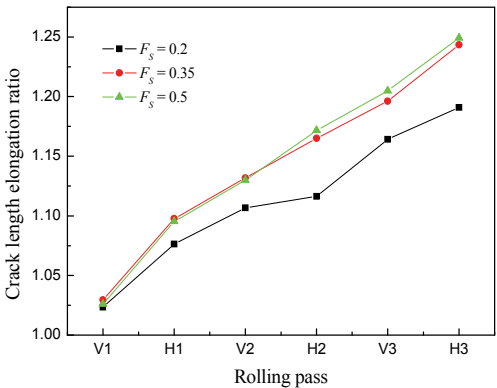


Fig. 36. Crack length elongation ratio for edger roll shape when FS equals 0.35

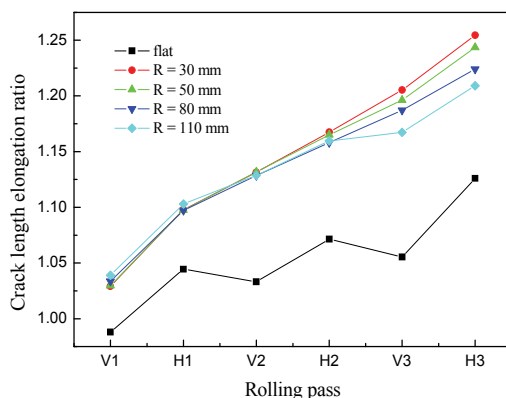


Fig. 37. Crack length elongation ratio for friction coefficient when R equals 50 mm

2.3.4 Experimental

Pure lead is often used for simulation of the slab deformation during hot rolling, because the mechanical characteristics of the pure lead at room temperature are similar to that of steel at high temperature. Meanwhile, the pure lead is also used to simulate the behavior of defects in steel hot deformation process.

In order to validate the calculated results, experiments of the closure and growth of longitudinal cracks on surface of slab edge in multi-pass V-H rolling processes were investigated in laboratory with the pure lead. The vertical rolling employs the $\Phi 300$ mm mill with grooved roll whose fillet radius is 11 mm; the horizontal rolling employs the $\Phi 180$ mm mill. The dimensions of slab profile are 25 mm \times 120 mm, and a V-shaped longitudinal crack is carved on the top surface in slab edge, the crack width is 0.6 mm, the crack depth is 1.5 mm, and the crack length is 6 mm. The rolling speed is 200 mm/s. In rolling process, according to the V₁-H₁-V₂-H₂-V₃-H₃ rolling schedule, the draft in vertical rolling is 5 mm, and the draft in horizontal rolling is 1 mm.

Table 7 shows the crack shapes after each pass between the calculated results and the experimental results (the friction coefficient between the pure lead and the steel is around 0.2^[1]). From the results, the calculated results have the similar deformation regularities with the experimental results. After the V₁ pass, the cracks close well, however, after the H₁ pass, the crack opens a little. In the subsequent passes, the crack gradually closes, after the V₃ pass, the crack nearly closes well.

Fig. 38 shows the crack length elongation ratio between the calculated results and the experimental results. In the figure, the differences between the experimental results and the calculated results are small. In the V₁~H₂ passes, the experimental results are around the calculated results; and in the V₃ and H₃ passes, the experimental results are few larger than the calculated results. About the differences between the experimental ones and the calculated ones, the main reason is the different roll radii between the experiment and the FE simulation. Meanwhile, the influence of the thermal distribution in slab during rolling on the behavior of cracks need to be further studied in future.


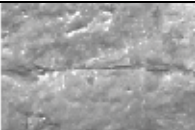
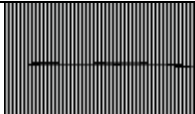



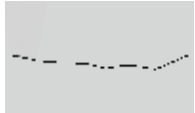





Rolling pass	Vertical rolling process		Horizontal rolling process	
	Calculated results	Experimental results	Calculated results	Experimental results
1				
2				
3				

Table 7. Compared the calculated results with the experimental ones of the crack shape from top surface

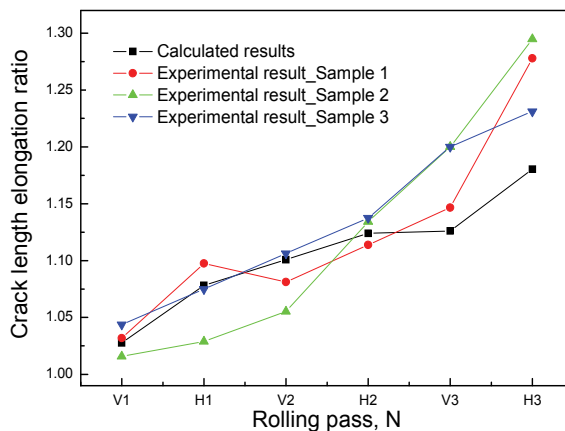


Fig. 38. Compared the calculated results with the experimental ones of the crack length elongation ratio

3. Evolution of internal cracks at slab edge by 3D FEM

The internal cracks might appear in slabs during continuous casting for the thermal stress, friction force, mechanical force, etc [47]. During hot rolling, the cracks might close, even heal. However, they might propagate for the concentrated force or other reasons. So analysis of the closure and growth of cracks during rolling is significant for improvement of the slab quality.

3.1 Parameters and boundary conditions

The diameter of horizontal roll is 1 150 mm. The diameter of grooved vertical roll is 980 mm, and other dimensions are shown in Fig. 16, where the R is 50 mm (30, 80, 110 mm used in

analysis of the behavior of cracks under various the fillet radius of grooved vertical roll). The dimension of slab profile before rolling is 1 200×250 mm. The cracks are assumed to be disk-shape cracks. **Fig.39** shows the schematic drawing of cracks shape, position, and size, where, X is rolling direction, Y is slab thickness direction, Z is slab width direction. In this simulation, the crack width and the crack height ($w \times h$) is 0.1×1 mm (0.2×2, 0.3×3, 0.4×4 mm used for analysis of the behavior of cracks under various crack size). And the crack position ($W \times H$) is 62.5×25 mm (0×25, 31.5×25, 93.75×25 mm and 62.5×125, 62.5×100, 62.5×75, 62.5×50 mm used in analysis of the behavior of cracks under various crack positions). During rolling, the draft in vertical rolling is 50 mm, and the draft in horizontal rolling is 10 mm (15, 20, 25 mm used in analysis of the behavior of cracks under various the draft in horizontal rolling). The friction coefficient between the slab and the rolls is 0.35. And the rotating speed of rolls is 4 m/s.

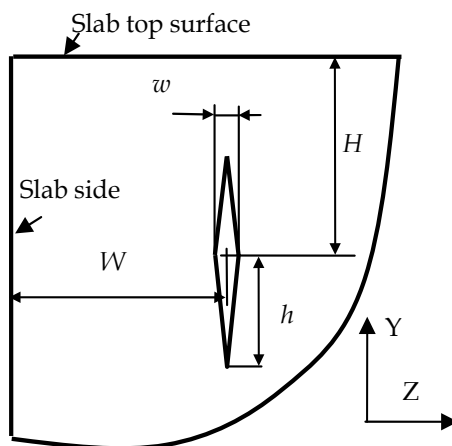


Fig. 39. Crack shape, size, position

The rigid material model is employed for rolls and the bilinear isotropic hardening material model for slab. In the simulation, the initial deformation resistance is calculated by Eq. 19. The main materials parameters are shown in **Table 3**.

3.2. FE models

Owing to the symmetry distribution of the rolls and slab, a quarter of rolling parts is employed. Because the crack size is much smaller than the slab size, an inlaid model is used. According to the parameters above, the FE model of V-H rolling process has been established. The whole model is discretized by the 8-noded hexahedral elements. The nodes on the symmetry plane of slab thickness are constrained, $U_Y = 0$; the nodes on the symmetry plane of slab width are constrained, $U_Z = 0$. During rolling process, the slab enters rolls with an initial speed and exits rolls under the action of friction force. The geometry and meshing of slab with an internal-longitudinal crack is shown in **Fig. 40**.

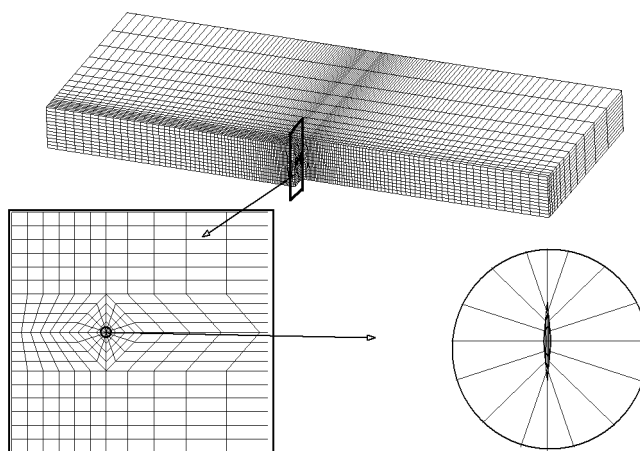


Fig. 40. Geometry and meshing of slab with an internal-longitudinal crack

3.3 Results and discussion [48]

Fig. 41 shows the crack shapes during rolling, where (a)~(d) are the crack shape during vertical rolling, and (e)~(f) during horizontal rolling. During vertical rolling, the crack gradually closes. As shown in Fig. 41 (d), the crack closes well. But during horizontal rolling, the crack gradually grows. As shown in Fig. 41 (g), the crack clearly appears in slab. Compared the crack size in Fig. 41 (g) with that in Fig.41 (a), the former is much less than the latter.

It is easy to understand that the crack surfaces will contact each other when the crack closes, and then an action force which is named the contact pressure will appear on crack surfaces.

Fig. 42 shows the curve of contact pressure on crack surfaces (next just says contact pressure) during rolling process. The contact pressure during vertical rolling is larger than that during horizontal rolling. In the bite zone of vertical rolling, the contact pressure gradually increases from 0 to 220 MPa, but which gradually decreases to 80 MPa in the exit zone of vertical rolling. In horizontal rolling, the contact pressure shapely decreases to 40 MPa, and exit about 20 MPa. In the below, the maximum value of contact pressure during vertical rolling and horizontal rolling are employed to analyze the influencing factors on the behavior of cracks.

Fig. 43 shows the maximum value of contact pressure during V-H rolling under a variety of fillet radii of grooved vertical roll. With increase of fillet radii, the contact pressure gradually increases from 110 MPa to 260 MPa during vertical rolling, but which gradually decreases from 50 MPa to 30 MPa during horizontal rolling. So the crack will be easily healed as the fillet radius increases.

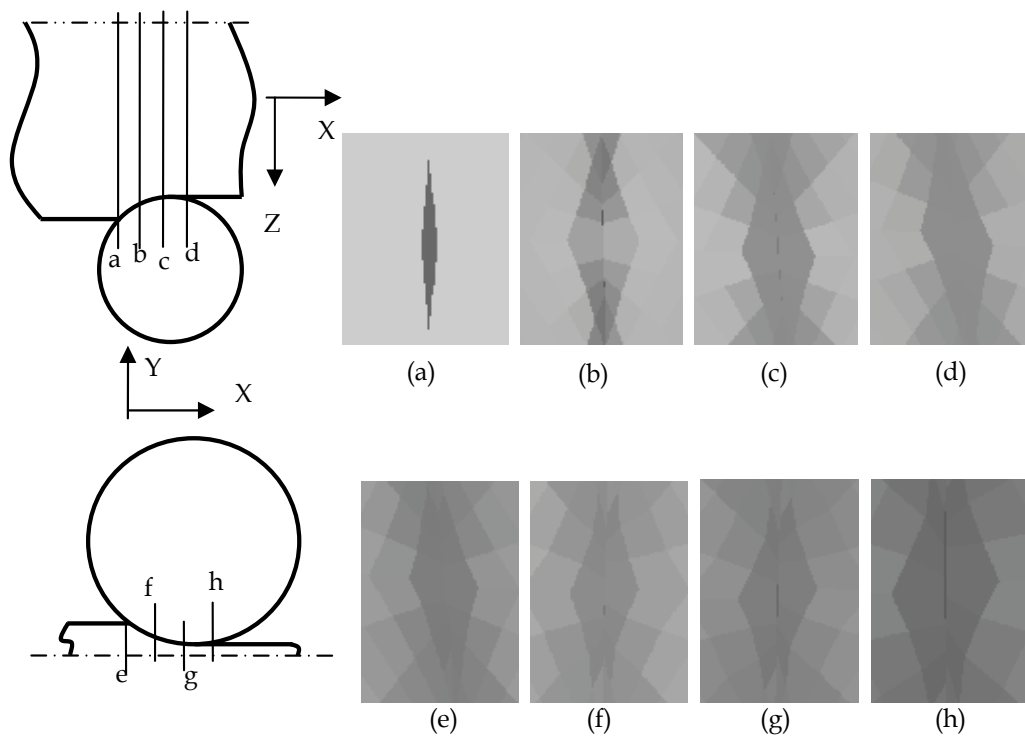


Fig. 41. Crack shape during V-H rolling

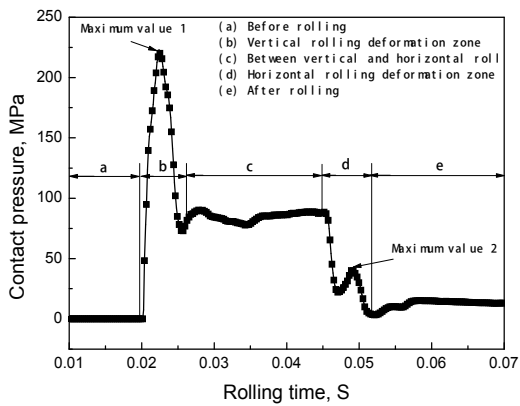


Fig. 42. Contact pressure on crack surfaces during V-H rolling

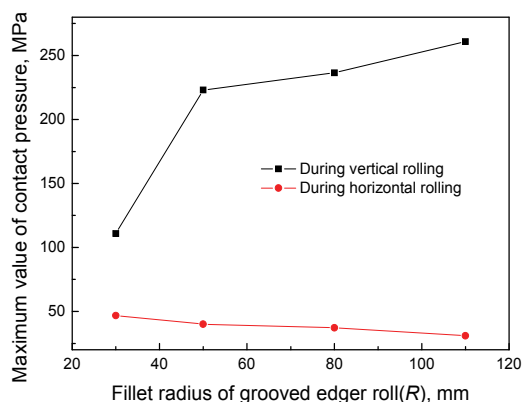


Fig. 43. Maximum value of contact pressure for various fillet radius of grooved vertical roll

Fig. 44 shows the maximum value of contact pressure during rolling under a variety of distance away slab central line along thickness direction ($125-H$). During vertical rolling, with increases of the distance away slab central line along thickness direction from 0 to 63 mm, the contact pressure increases, but the contact pressure decreases from 63 to 93 mm. During horizontal rolling, the contact pressure gradually increases as the distance increases.

Fig. 45 shows the maximum value of contact pressure during rolling under a variety of distance away slab central line along width direction ($600-W$). During vertical rolling, with increases of distance, the contact pressure increases from 0 to 221 MPa. During horizontal rolling, the contact pressure increases when the distance is between 450 and 500 mm but decreases when the distance away slab central line along width direction is between 500 and 575 mm.

Fig. 46 shows the maximum value of contact pressure during rolling under a variety of crack sizes. During rolling, when the crack width is less than 0.2 mm, the contact pressure changes slightly. However, with increases of crack size, the contact pressure shapely decreases. When the crack width is larger than 0.3 mm, the crack can close well under this rolling condition.

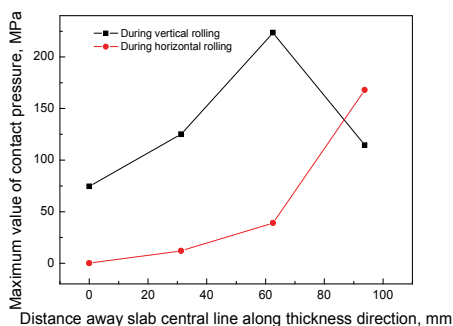


Fig. 44. Maximum value of contact pressure for various distance away slab central line along thickness direction

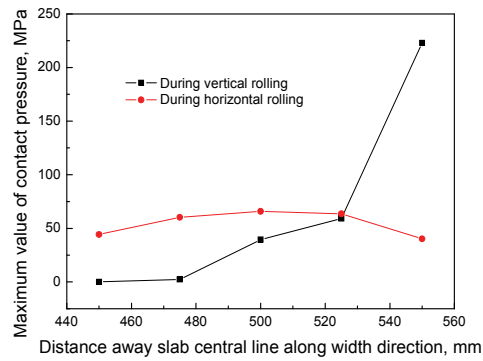


Fig. 45. Maximum value of contact pressure for various distance away slab central line along width direction

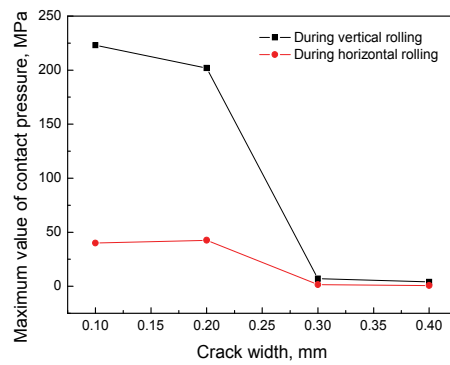


Fig. 46. Maximum value of contact pressure for various crack sizes

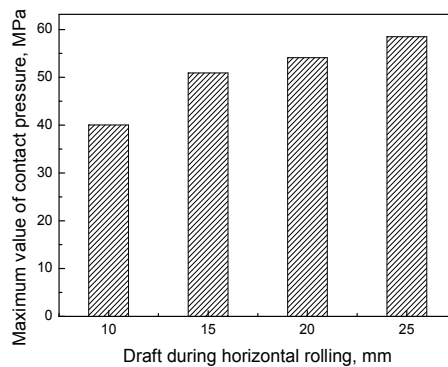


Fig. 47. Maximum value of contact pressure for various drafts during horizontal rolling

Fig. 47 shows the maximum value of contact pressure during rolling under a variety of draft in horizontal rolling process. During rolling, with increases of the draft, the contact pressure increases.

During hot rolling, the rolling temperature is high which generally is about 800~1200 °C. At this time, the ions near the crack surfaces are very activity. Meanwhile, there are many influencing factor which might make the cracks heal, such as power supplement, matter supplement. So the cracks might heal during rolling. **Fig. 48 [49]** shows the condition of crack healing under various slab reductions. As shown in the figure, the crack size decreases clearly with increase of slab reduction ratio. And when the rolling conditions are certain, the contact pressure will decrease as the crack size increases, as shown in Fig. 46.

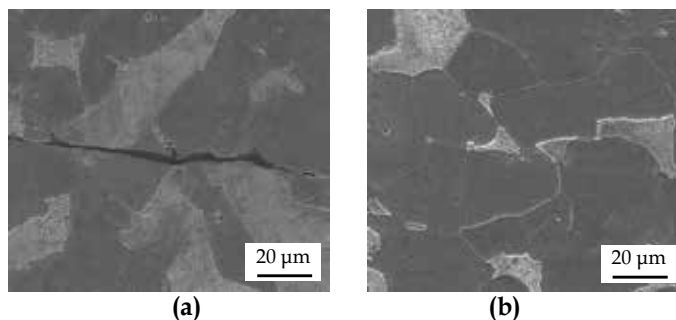


Fig. 48. Crack shape under 5%(a) and 50% (b) slab reductions during hot rolling[48]

It is easy to understand that the compression stress in slab increases with increase of slab reduction ratio, which means that the contact pressure will increase. So as the contact pressure increases, the cracks will be easier to heal. From the results above, we know the contact pressure during vertical rolling is much larger than that during horizontal rolling. So when the internal-longitudinal cracks appear in slab, it is significant to control the slab vertical rolling. As the fillet radii increase, the equivalent reduction ratio of slab increases, so the contact pressure increases. Meanwhile, the contact pressure also increases with increase of reduction ratio of slab during horizontal rolling, as shown in Fig. 47.

Fig. 49 [50] shows the strain distribution during rolling process. During vertical rolling, the deformation of slab mainly assembled near slab edge, so the contact pressure decreases when cracks approaching the slab central line along slab width direction. Meanwhile, the strain also decreases when approaching the slab central line along slab thickness direction, so do the contact pressure, as shown in Figs. 44 and 45.



Fig. 49. Strain distribution after vertical (a) and horizontal (b) rolling

4. Evolution of inclusions in strip during rolling by FEM

The non-metallic inclusions inevitably exist in steel strips for the deoxygenation, etc, which is still one of the most important topics during continuous casting although their quantity, size, shape, distribution and composition are controlled at a lower level [51]. The chemical compositions of

non-metallic inclusions in type 304 stainless steel strips mainly are the oxide of aluminum, silicon, chromium, manganese, etc [52]. The distribution and the size of inclusions severely affect the quality of strips. Analysis of the deformation of non-metallic inclusions in rolling process is significant for improvement of the strip quality.

The FEM has been used for analyzing the deformation of inclusions in a slab during rolling. A rigid-viscoplastic 2D FE code was developed to analyze the shape of inclusions during rolling along rolling direction of oxidized inclusions [53] and MnS inclusions [54] in flat rolling process under different rolling temperature, friction and rolling schedules. The start of void deformation close to a hypothetical single circle macro-inclusion in the hot rolling of steel slabs which was affected by the reduction, roll radius and inclusion position was simulated through the FE-code LS-Dyna3D [55]. It pointed out that the large roll radii should be chosen for preventing void initiation, heavy draughts resulting in large contact lengths are favorable and inclusion close to the surface are less dangerous than those in the centre of the slab. The void generation and development around a rigid circle inclusion simulated by FEM [56] showed that the void length in front of inclusion was larger than that in rear of inclusion under different rolling conditions. The behavior of short cracks between the inclusions and the matrix under rolling contact fatigue load was simulated by Melander [57] who analyzed the effects of the inclusion shape on the behavior of cracks. In this chapter, the application of FEM in simulation of evolution of cracks and inclusions in rolled steel during rolling was carried out. In this part, the researches on the evolution of inclusions during flat rolling were carried out.

4.1 Evolution of inclusions combined with matrix

4.1.1 Basic parameters and rolling conditions

Fig. 50 shows the schematic drawing of rolling strip with a spherical inclusion, where X is rolling direction, Y is the strip thickness direction, and Z is the strip width direction.

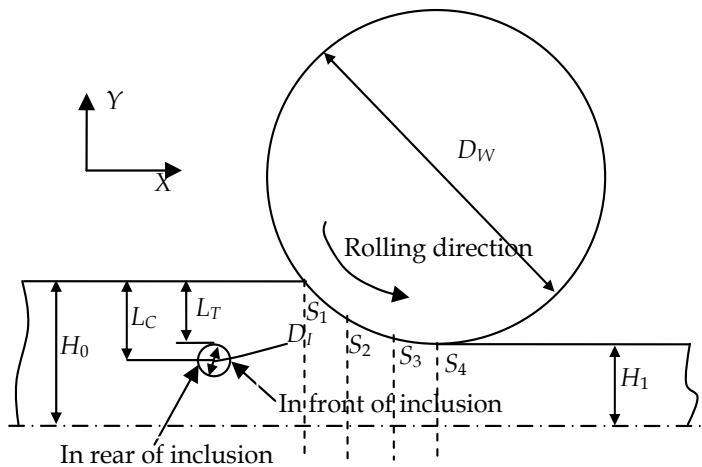


Fig. 50. Schematic drawing of rolling strip with a spherical inclusion

In the simulation, the work roll diameter (D_W) is 400 mm. The diameters of inclusion (D_i) are 10, 20, 30, 40, 50 μm respectively which are assumed to be spherical and in the position of $1/8$, $1/4$, $3/8$ and $1/2$ of strip thickness ($L_C/(2H_0)$). A part of strip is employed whose width

is 30 mm for saving computing time. The friction coefficient between the strip and the roll is 0.15. And the rolling schedule is shown in **Table 8**.

Rolling pass, N	1	2	3
Strip thickness before rolling ($2H_0$), mm	3.0	2.0	1.4
Strip thickness after rolling ($2H_1$), mm	2.0	1.4	1.0

Table 8. Rolling schedule

During rolling, the strip matrix is the type 304 stainless steel, and the inclusion is assumed to be Al_2O_3 (As shown in **Fig. 51**). The main material parameters of rolls, strip matrix [58] and inclusion [59] used in FE models are shown in **Table 9**. The deformation resistance of strip in the i th pass is calculated by Eq. (20).

$$\sigma = C[0.001 + \ln(\frac{h_i}{H_0})]^n \quad (20)$$

where, C is the constant related to the strip temperature, h_i is the strip thickness in the i th pass, H_0 is the initial strip thickness, n is the strain hardening coefficient.

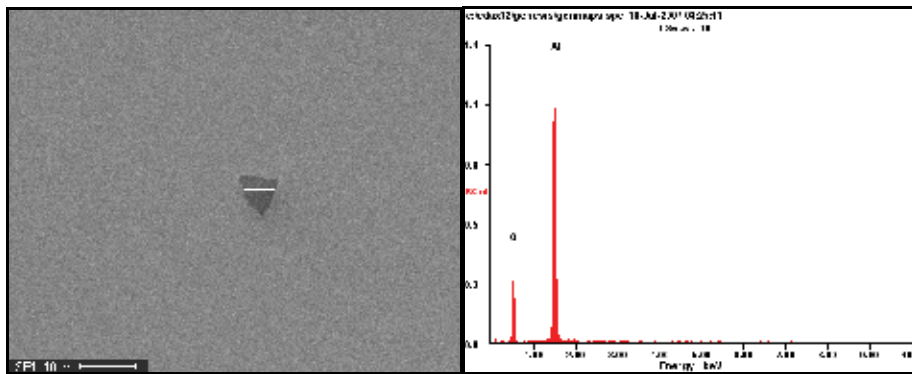


Fig. 51. Inclusion in Type 304 stainless steel strip

Parameters	Roll	Strip	Inclusion
Density, kg/m^3	7850	7830	3800
Young's modulus, GPa	210	193	352
Poisson's ratio	0.3	0.36	0.24
Initial deformation resistance, MPa	-	205	263.9
Strain hardening coefficient in Eq.(1)	-	0.112	-

Table 9. Material parameters in rolling process

4.1.2 FE model

Owing to the symmetry of strip and rolls, a quarter of the strip and the rolls are considered in the geometrical model. The rigid material model is employed for the rolls and the bilinear isotropic material models for the strips and the inclusions. The geometrical models of the strip with an inclusion before rolling are built which are meshed with 8-noded hexahedral elements. There are 6912 elements in the inclusion and 43520 elements in the strip matrix. The FE meshing of strip

with an inclusion before rolling is shown in Fig. 52. In rolling process, the roll rotates with a stable angular velocity, and the strip enters the roll with an initial velocity and exits under the action of friction force. The nodes in the middle cross section of strip thickness are constrained, $U_Y=0$; and the nodes in the middle cross section of strip width are constrained, $U_Z=0$. The FE models of the second pass and the third are obtained by updating geometrical method.

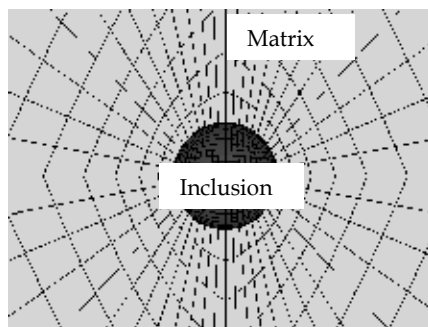


Fig. 52. FE meshing of strip with an inclusion

4.1.3 Results and discussion

(1) Deformation of inclusions during rolling [60]

Table 10 shows the inclusion shape after deformation in the rolling processes. The deformation of inclusion gradually increases from S_1 to S_4 as well as that from Pass 1 to Pass 3. When the inclusion is bond-limited with the strip matrix, after multi-pass cold rolling, the inclusion shape calculated by simulation is in good agreement with the experimental result as shown in Fig. 53. The deformation process of inclusions in detail under various rolling conditions is shown as follows.

Fig. 54 shows the projection of the inclusion shape after every pass in cold rolling process when the inclusion diameter is 20 μm . The inclusion spreads along the rolling direction and decreases along the strip thickness direction but slightly changes along the strip width direction. Meanwhile, the unequal deformation of inclusion gradually increases as the rolling pass increases, and the deformation in the front of inclusion is larger than that in the rear of inclusion.

	Rolling direction			
	S_1	S_2	S_3	S_4
During pass 1				
Inclusion shape				
During multi-pass	Before rolling	Pass 1	Pass 2	Pass 3
Inclusion shape				

Table 10. Inclusion deformation in rolling process when inclusion size is 20 μm

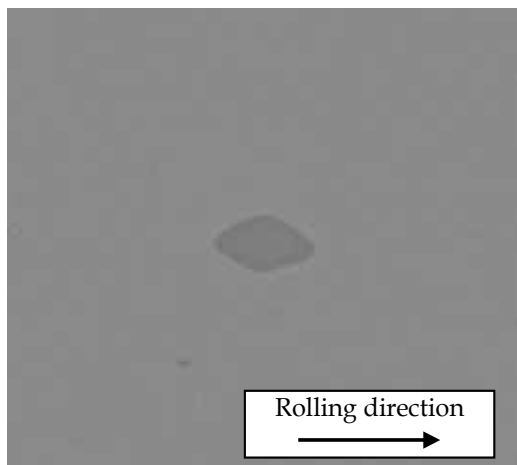


Fig. 53. Inclusion shape after cold rolling

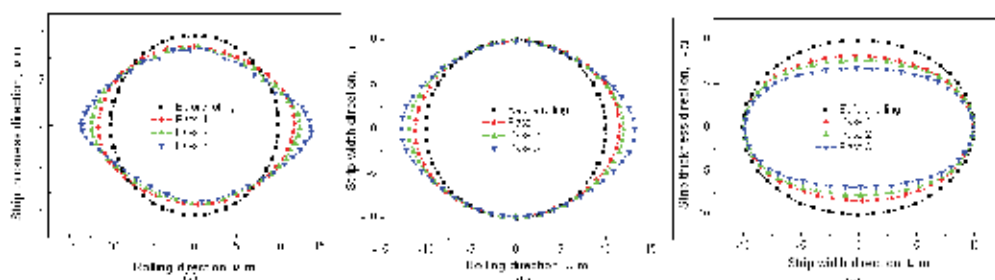


Fig. 54. Inclusion shape in XY (a), XZ (b), and YZ (c) profiles after multi-pass rolling for 20-micron inclusion

In order to better analyze the influence of inclusion sizes on the deformation of inclusions, the scaled profile of inclusions (generating the scaled inclusion profiles with the inclusion diameter) is employed. **Fig. 55** shows the inclusion shape in the projection of XY profile for a variety of inclusion sizes in cold rolling process. In the first pass, the inclusions have similar shape when the inclusion diameter is larger than 20 μm , but the deformation of inclusion is much less when the inclusion diameter is less than 10 μm . As for different inclusion size, the difference of inclusion shape gradually increases with increase of rolling pass. And the smaller the inclusion diameter is, the less the inclusion deforms.

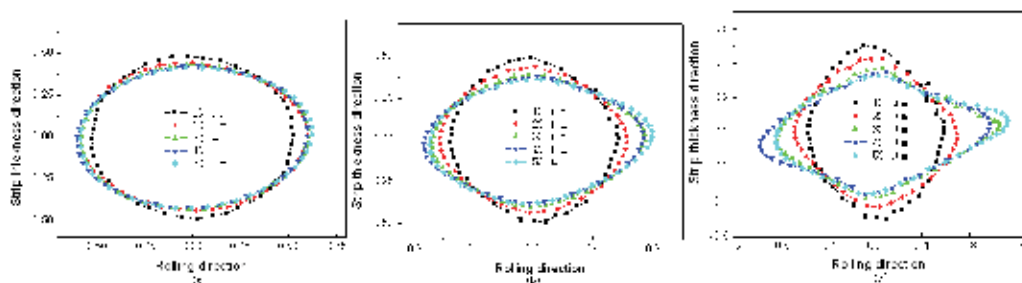


Fig. 55. Inclusion deformation for inclusion sizes after the first (a), second (b), and third (c) pass

Fig. 56 shows the inclusion shape in the projection of XY profile for a variety of the inclusion positions in cold rolling process. The inclusions in different positions of strip have similar deformations in the first pass except the inclusion in position of 1/8 of strip thickness. Meanwhile, the shorter the distance between the inclusion and the strip surface is, the larger the inclusion deforms.

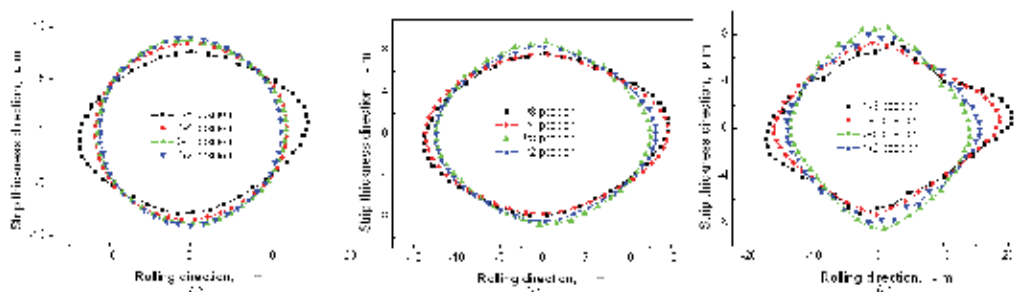


Fig. 56. Inclusion deformation for inclusion positions after the first (a), second (b), and third (c) pass

(2) Movement of inclusion in strip [61]

In rolling process, the influence of the distribution of inclusions on the strip quality is very large. For some kinds of inclusions whose deformation resistances are larger than that of the strip matrix, the deformation of inclusion is less than that of the strip matrix during rolling, which might move to the strip subsurface and then affect the quality of strip surface. Sometimes, the inclusion size is large, the distance between the inclusion and the strip surface will decrease clearly with reduction of the strip thickness; even the inclusion exposes to the strip surface and worsens the strip quality. Meanwhile, the surface cracks of slab might be also caused for the enriched inclusions near the subsurface of strip. In the following, the absolute distance between the inclusion and the strip surface (L_T') and the relative distance between the inclusion and the strip surface (L_T' / H_i) are employed respectively.

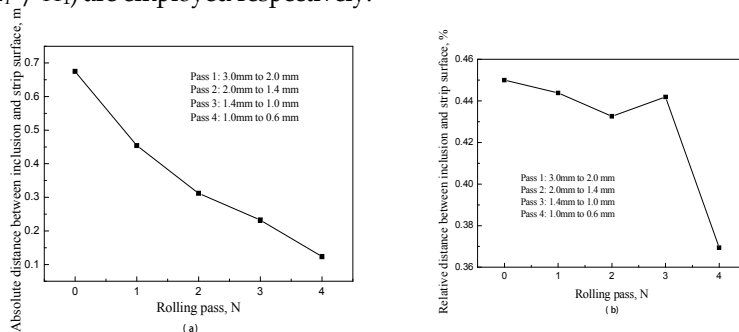


Fig. 57. Movement of inclusion during multi-pass rolling

Fig. 57 shows the change curve of the distance between the inclusion and the strip surface to the strip thickness. In Fig.57 (a), the absolute distance between the inclusion and the strip surface decreases as the rolling pass increases. When the strip thickness is 3.0 mm, the absolute distance between the inclusion and the strip surface is about 0.67 mm; and when the strip thickness is 0.6 mm, the absolute distance is only 0.1 mm. Fig.57 (b) shows the

relative distance between the inclusion and the strip surface. In the figure, the relative distance between the inclusion and the strip surface gradually decreases as the rolling pass increases. Meanwhile, the relative distance clearly decreases when the strip thickness is thin. From the results, it might be one kind of the justifications that the thinner the strip thickness is, the larger the effect of inclusions on the strip surface quality is.

Fig. 58 shows the distance between the inclusion and the strip surface under a variety of inclusion sizes, (a) for the absolute distance, (b) for the relative distance. As shown in the figure, the absolute distance decreases with decrease of the inclusion sizes. When the inclusion diameter is 10 μm , the relative distance changes slightly during multi-pass rolling, but when the inclusion diameter is 100 μm , the relative distance clearly decreases as the rolling pass increases. So, the inclusion will more easily approach the strip surface as the inclusion size increases, and the relative distance scarcely changes when the inclusion size is less than 10 μm .

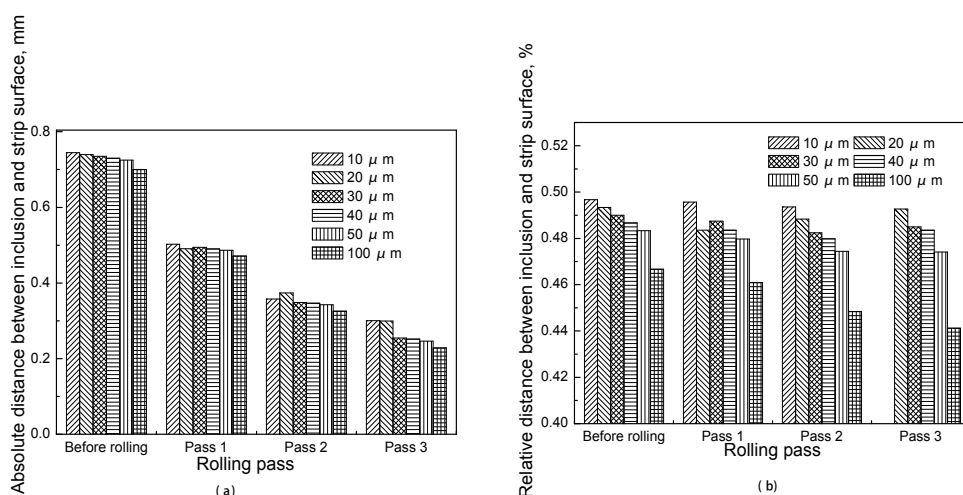


Fig. 58. Movement of inclusion for various inclusion sizes

Fig. 59 shows the distance between the inclusion and the strip surface under various inclusion deformation resistances, where (a) shows the absolute distance and (b) shows the movement distance. The distance between the inclusion and the strip surface for hard inclusions is less than that for soft inclusions when their sizes are the same.

Fig. 60 shows the distance between the inclusion and the strip surface under a variety of inclusion position, where (a) for the absolute distance and (b) for the relative distance, (c) and (d) are the scaled figure of inclusion in the position of 1/8 strip thickness shown in the Figs. (a) and (b). During rolling, the absolute distance decreases with increase of the rolling passes, and the difference between the absolute distances increases with increase of the initial absolute distance. But the inclusion easily approaches to the strip surface when the initial absolute distance decreases. As shown in Fig. (d), the relative distance decreases clearly as the rolling pass increases when the inclusion is in the position of 1/8 of strip thickness.

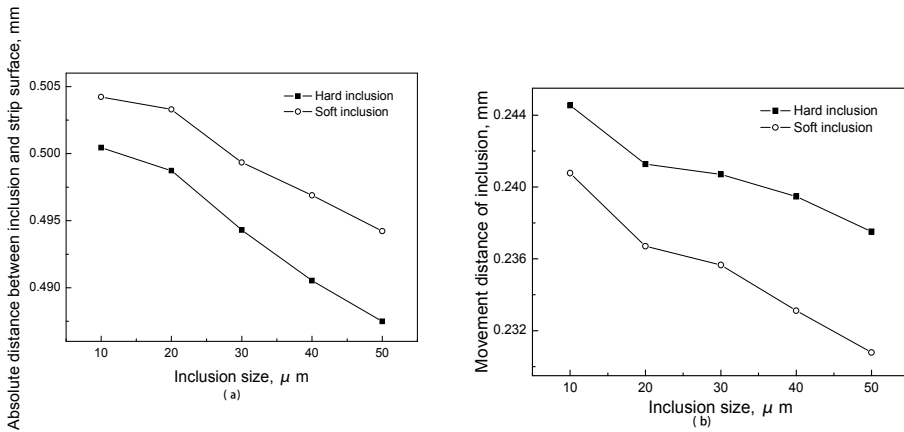


Fig. 59. Movement of inclusion for various deformation resistances of inclusions

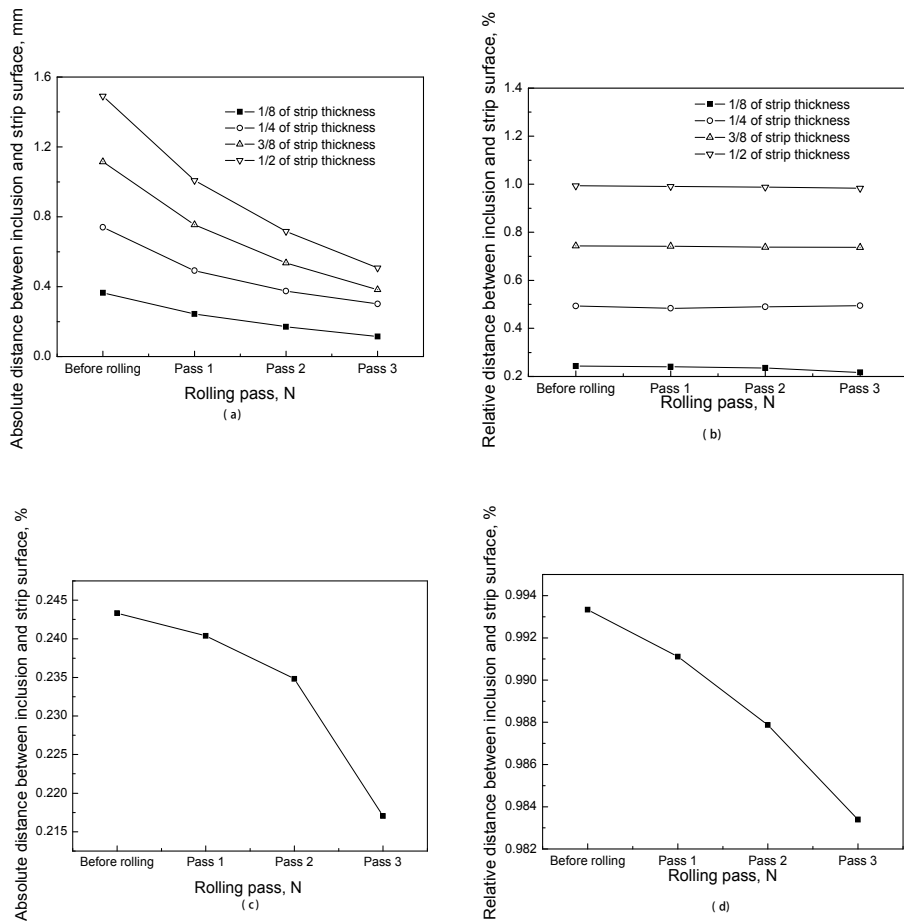


Fig. 60. Movement of inclusion for various inclusion positions (20 μm)

Fig. 61 shows the distance between the inclusion and the strip surface under a variety of work roll diameter, where (a) for the absolute distance and (b) for the relative distance. In the figure, the influence of work roll diameters on the movement of inclusion is small.

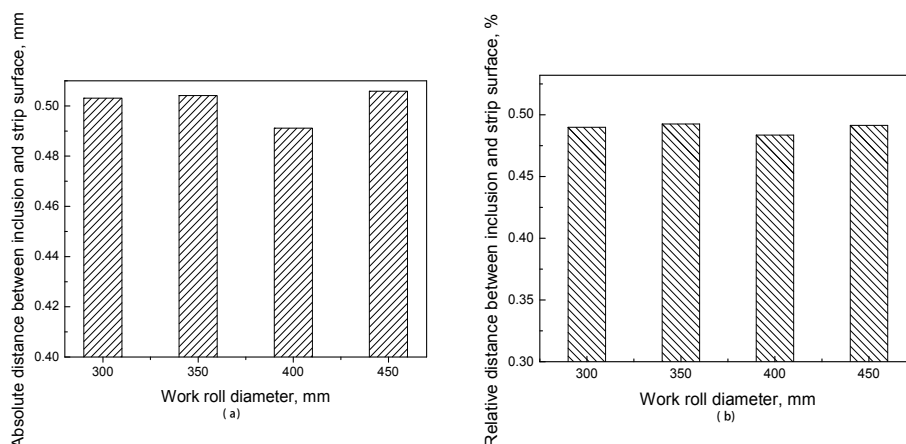


Fig. 61. Movement of inclusion for various work roll diameters (20 μm)

4.1.4 Discussion

During cold rolling, the cracks might appear between the inclusion and the strip matrix when the non-uniform deformation occurs between the inclusion and the strip matrix and whose bonding strength is low, as shown in **Fig. 62**. When the temperature of strip is high, the cracks might self-heal [62, 63], however, the iron atoms in strip generally do not migrate when the temperature of strip is less than 500 $^{\circ}\text{C}$, so the cracks can not automatically heal during cold rolling. For that it is significant for decrease of the defects in strip to analyze the relationships between the reduction of strip and the deformation of inclusions. [64] proposed the methods to control faults forming in the hot deformation of workpieces, (1) the plastic deformation should be used to control the shape of inclusion and reduce the concentration value of stress and strain caused by inclusion; (2) plastic inclusions should be dispersed to prevent cracks originating from inclusions in expatriation to meet the quality standard of flaw detection.



Fig. 62. Cracks near inclusion in strip after rolling

So the influence of the inclusion sizes on the crack generation is obvious, and the harm of inclusions during cold rolling will be more obvious if the large size inclusions are not dispersed into small ones during hot rolling.

From the results above, the deformation of inclusions decreases with reduction of the inclusion size. When the inclusion diameter is less than $10\text{ }\mu\text{m}$, the deformation of inclusion is very small. By contrary, when the inclusion diameter is larger than $20\text{ }\mu\text{m}$, the inclusion will deform with the strip matrix and elongate along the rolling direction. Meanwhile, it could be found that the deformation in front of inclusion is larger than that in rear of inclusion. Because the fatigue damage in bulk ductile metals usually originates from the strain localized region, where extrusions/intrusions form and the cracks might appear [65], the crack (void) size in front of inclusion is larger than that in rear of inclusion [56]. From the analysis above, the strip with a certain-size inclusion could but bear a certain reduction ratio; otherwise the cracks will appear in the strip. Meanwhile, from the results above, the closer the inclusion approaches the strip surface, the greater the deformation of inclusion occurs [66]. Besides the inclusion size and the inclusion position, some other influencing factors might affect the inclusion deformation, such as: the rolling reduction ratio and the rolling speed. Ervasti finds the void size increase with an increase of the rolling reduction ratio and decrease of the rolling speed.

The equivalent strain distribution of the strip matrix and the inclusion in the deformation zone is shown in Fig. 63. In the figure, the strain value of inclusion is much less than that of strip matrix, which is about 0.07 in the center of inclusion that is about 1/10 of the strain value of the strip matrix. Meanwhile, a zone of large strain value exists in the front and the rear of inclusion, for some kinds of strip matrix, where the cracks are easy to appear. Table 11 shows the maximum and the minimum strain and the strain gradient between the inclusions and the strip matrix for a variety of the inclusion sizes. When the inclusion diameter is $50\text{ }\mu\text{m}$, the difference of strain between the inclusion and the strip matrix reaches 1.025. So, the strip quality will deteriorate with increases of the inclusion sizes.

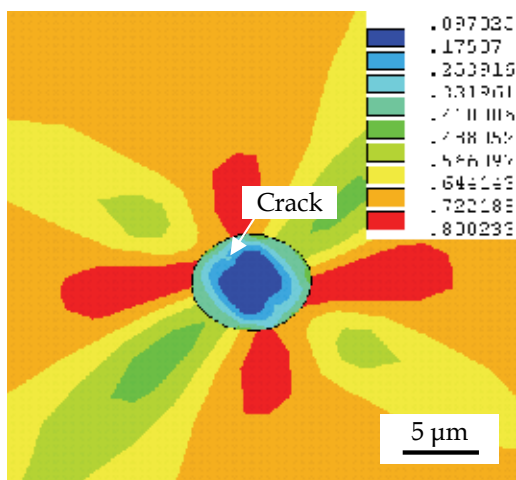


Fig. 63. Contour of the strain distribution of strip with a 20-micron inclusion after first pass in XY profile [67]

Inclusion size, μm	10	20	30	40	50
Strip matrix	0.803	0.800	0.945	1.026	1.246
Inclusion	0.044	0.098	0.168	0.243	0.221
Strain difference	0.759	0.702	0.777	0.783	1.025

Table 11. Strain gradient between the inclusion and the strip matrix for a variety of inclusion sizes [67]

4.2 Evolution of inclusions separate with matrix during rolling [68]

During continuous casting, the inclusions might exist in stainless steel. The behavior of inclusions will strongly affect the strip quality during rolling, especially the macroscopic inclusions. During rolling, the inclusions might be disjoined with the strip matrix, and then cracks generate; sometimes, the inclusions might move to the strip surface, which cause the scratch marks.

This part presents a new method to simulate the behavior of inclusions and the crack generation during rolling. The bond strength between the inclusions and the matrix can be adjusted through changing the material attribution of the transition layer which is established between the inclusions and the matrix. Meanwhile, the material failure is taken into account in the elements of the transition layer during deformation, with which the crack generation can be analyzed. The behavior of inclusions in type 304 stainless steel strips has been simulated by 3D FEM, and the influence of the bond strength between the inclusions and the matrix on the crack generation is analyzed.

4.2.1) Experimental

The SEM photograph of the inclusion in type 304 stainless steel strip was taken, as shown in Fig. 64. After corrosion, there is clear space between the inclusion and the strip matrix which is corroded. So there might be a transition layer between the inclusion and the matrix. According to the phenomenon, it can be assumed that there exists a transition layer between the inclusion and the strip matrix when simulating the deformation of inclusion during rolling, and that the bond strength between the inclusion and the matrix can be changed through adjustment of the material attribution of the transition layer.

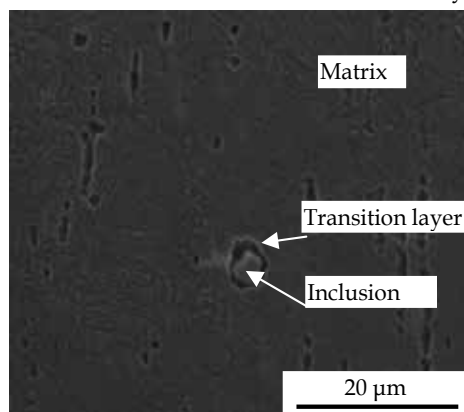


Fig. 64. SEM photograph of inclusion in 304 stainless steel strip with corrosion

4.2.2 FE analysis

(1) Basic parameters

In the simulation, the work roll diameter is 400 mm which is assumed to be rigid. The strip thickness is 3 mm before rolling and 2 mm after rolling. The diameter of inclusion is 48 μm which is assumed to be spherical and in the position of 1/4 of strip thickness (The inclusion size changes from a few to tens of microns, and a large value is employed for saving computer time.). And the thickness of the transition layer is 2 μm . During rolling, the inclusion is Al_2O_3 ; and the metrical parameters of transition layer are assumed to be the steel and the inclusion, each of which is occupying half and the material failure is considered. The friction coefficient is 0.15 between the strip and the roll. The main material parameters are shown in **Table 12**.

Parameters	Roll	Strip	Inclusion	Transition
Density, kg/m^3	7850	7830	3800	5815
Young's modulus, GPa	210	193	352	272.5
Poisson's ratio	0.3	0.36	0.24	0.3
Deformation resistance, MPa	-	205	263.9	238.8
Failure strain (F_s)	-	-	-	various

Table 12. Main material parameters in rolling process

(2) FE model

Owing to the symmetry of strip and rolls, a quarter of the strip and the rolls are considered in the geometrical model. The rolling model of the strip with inclusion is built with the parameters above. The model is meshed with 8-noded hexahedral elements. It just refines the elements in the inclusion, the transition layer and the matrix near the inclusion. And there are 6912 elements in the inclusion, 1728 elements in the transition layer and 80640 elements in the strip matrix. In rolling process, the roll rotates with a stable angular velocity, and the strip enters the roll with an initial velocity and exits under the friction force. The nodes on the middle cross section of strip thickness are constrained, $U_Y=0$; and the nodes in the middle cross section of strip width are constrained, $U_Z=0$ (X is rolling direction, Y is the strip thickness direction, Z is the strip width direction). The geometry and meshing of strip with the inclusion and the transition layer before rolling is shown in **Fig. 65**.

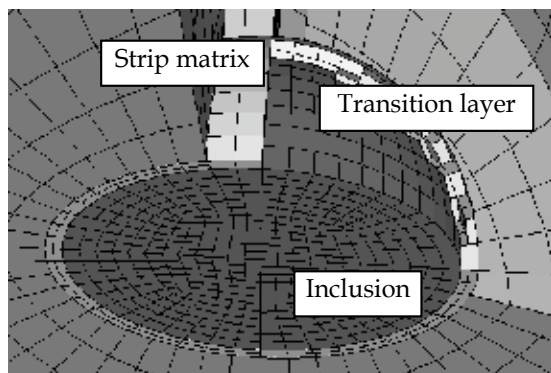


Fig. 65. FE meshing of strip with inclusion

4.2.3 Results

When the deformation of transition layer reaches the failure condition, a crack will appear between the inclusion and the strip matrix. The crack sizes can be obtained under the different failure conditions of transition layer. **Fig. 66** shows the transition layer shape after single-pass cold rolling process when the failure strain equals 0.15 (a), 0.1(b), 0.05(c) and 0.025(d). In the figure, the crack size in the transition layer between the inclusion and the matrix increases with the decrease of the bond strength between the inclusion and the strip matrix. When the failure strain is larger than 0.15, the cracks do not appear on the transition layer which deforms with the strip. When the failure strain is less than 0.1, the cracks appear on the transition layer, and the crack direction is about 45 degree with the rolling direction which is affected by the position of inclusion and the shape of deformation zone in the rolls gap. And with the decrease of the failure strain of transition layer, the cracks propagate. When the failure strain is less than 0.025, the inclusion almost sloughs off the strip matrix.

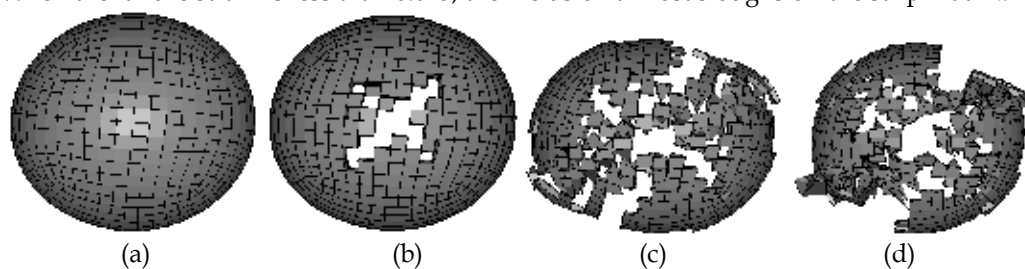


Fig. 66. Failure conditions of transition layer after rolling

Fig. 67 shows the inclusion shape after rolling. In the figure, the inclusion elongates along the rolling direction which is changed from a sphere to an ellipsoid; and the inclusion spreads slightly along the strip width direction compared with that along rolling direction.

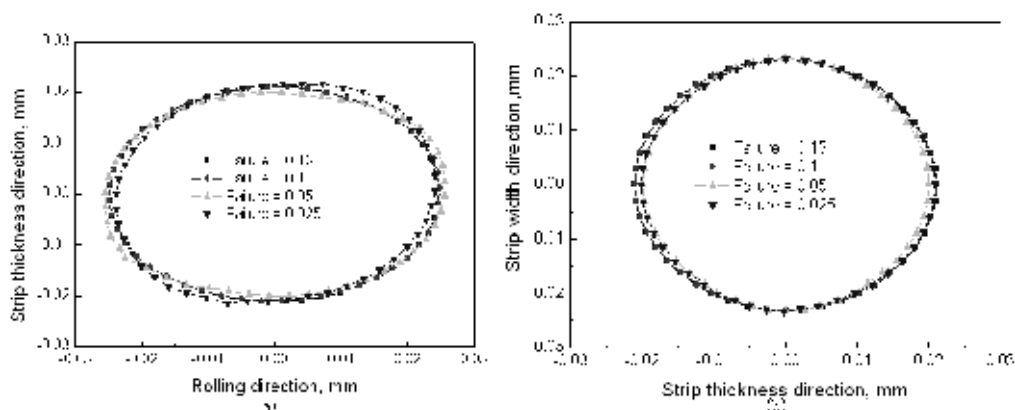


Fig. 67. Inclusion shape after rolling along rolling direction (a) and along strip width direction (b)

4.3 Evolution of inclusion with weak adhesion to strip matrix

4.3.1 Basic parameters and rolling conditions

Fig. 68 shows the schematic drawing of strip rolling with a circle (square, triangle) inclusion, where X is the rolling direction, Y is the strip thickness direction. **Fig. 69** shows three typical

kinds of shapes of inclusions in the type 304 stainless steel determined by SEM.

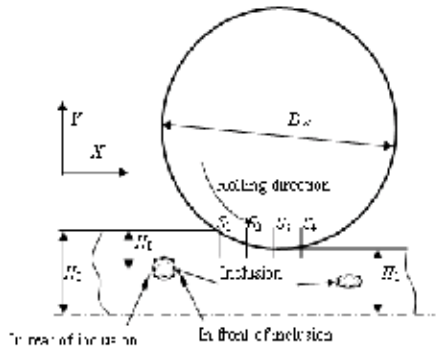


Fig. 68. Schematic drawing of strip rolling with an inclusion



Fig. 69. Shape of inclusions in a stainless steel

In the simulation, the work roll diameter (D_w) is 400 mm. The strip thickness is 3.0 mm before rolling and 2.0 mm after rolling. The circle, square, and triangle inclusions [57, 69] are employed respectively whose sizes are 10, 20, 30 and 40 μm separately and in the position of $1/2$, $3/8$, $1/4$ and $1/8$ of strip thickness (H_1/H_0). The Coulomb friction model is used in the FE models and the friction coefficient between the strip and the roll is assumed to be 0.15. The friction coefficient during cold rolling is about 0.04 ~ 0.231[70].

During rolling, the strip matrix is the type 304 stainless steel. And the hard and soft inclusions are employed respectively. The main material parameters of the roll, the strip matrix and the inclusions used in FE models are shown in **Table 13**.

Parameter	Roll	Strip matrix	Inclusion	
			Hard	Soft
Density, kg/m^3	7850	7830	3800	5000
Young's Modulus, GPa	210	193	352	120
Poisson's ratio	0.3	0.36	0.24	0.36
Deformation resistance, MPa	-	205	263.9	100

Table 13. Material parameters of roll, strip matrix and inclusion

4.3.2 FE model

Due to symmetry of strip and rolls, a half of the strip and the rolls are considered in the model. The rigid material model is employed for the roll and the bilinear isotropic material model for the strip matrix and the inclusions. The 2D geometrical models of the strip with an inclusion before rolling are built with the parameters above and are meshed with quadrilateral elements. The inclusions and the strip matrix are assumed to be weak adhesion [56], and the constant friction coefficient exists at the interface between inclusions and strip matrix. The FE meshing of strips with circle, square, and triangle inclusions before rolling and the whole rolling model are shown in Fig. 70. The element sizes around inclusions are much finer than the others for saving the computing time. In rolling process, the roll rotates with a constant angular velocity and the strip enters the roll with an initial velocity and exits under the action of friction force between the roll and strip. During rolling, cracks form for the contact between matrix surfaces and do not propagate.

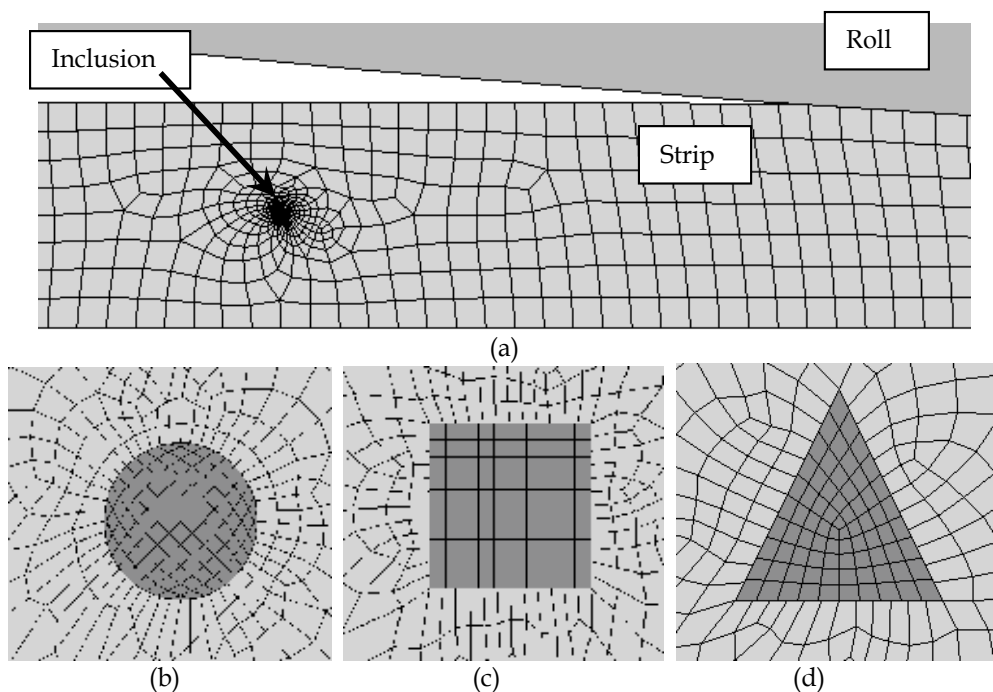


Fig. 70. FE meshing of the whole rolling model (a), and slab with circle (b), square (c), triangle (d) inclusions

4.3.3 Results and Discussion [71]

Fig. 71 shows the relationship between the hard circle inclusion and strip matrix at different positions in the deformation zone during rolling. In the position S_2 , there are two voids in front and rear of the inclusion, and the tip of voids gradually close, which change into cracks after rolling, as shown in Fig.71 (d). In the following, the relationship between the inclusion and the strip matrix in the position of S_4 during rolling is employed.

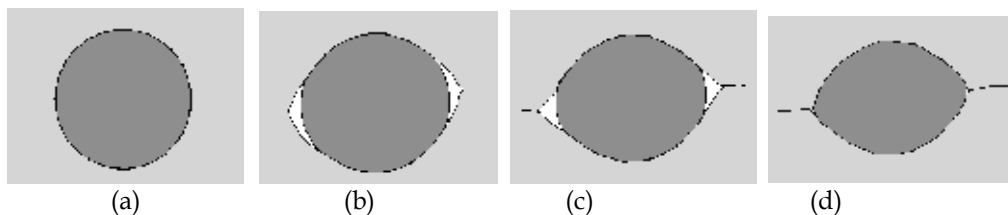


Fig. 71. Relationship between the inclusion and the strip matrix in S1(a), S2(b), S3(c), S4(d) position during rolling

Fig. 72 shows the relationship between the hard inclusion and the strip matrix after rolling when the inclusion size is $10\ \mu\text{m}$. If the inclusion shape is circle, there are two cracks in front and rear of the inclusion respectively. If the inclusion shape is triangle, there is one crack in rear of the inclusion. **Fig.73** shows the inclusion shape in type 304 stainless steel after cold rolling through SEM. Compared the relationship between the inclusion and the strip matrix in the **Fig. 72** and that in the **Fig. 73**, the calculated results are in good agreement with the one in the experiment.

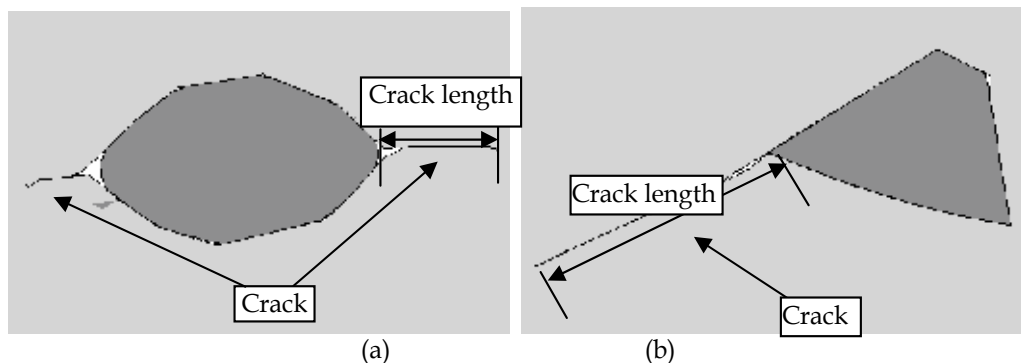


Fig. 72. Relationship between the inclusion and the strip matrix after rolling for 10-micron circle (a) and triangle (b) hard inclusions in the position of $1/4$ of strip thickness

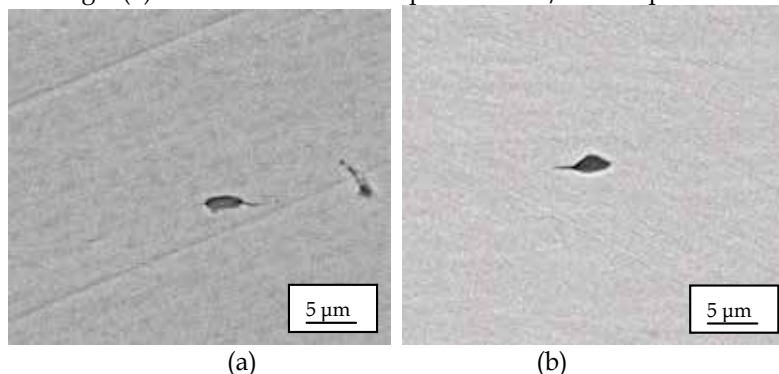


Fig. 73. SEM graph of cracks around circle (a) and triangle (b) hard inclusions in strip matrix after rolling

Fig. 74 shows the relationship between the soft inclusion and the strip matrix after rolling process when the inclusion size is $10\ \mu\text{m}$. If the inclusions are soft, there are no cracks

around the inclusions. Because the inclusion shapes are obtained in single pass, it is obvious that the inclusions will spread into a narrow strip after multi-pass rolling process, which will be similar to that of the experimental results through SEM as shown in Fig. 75.

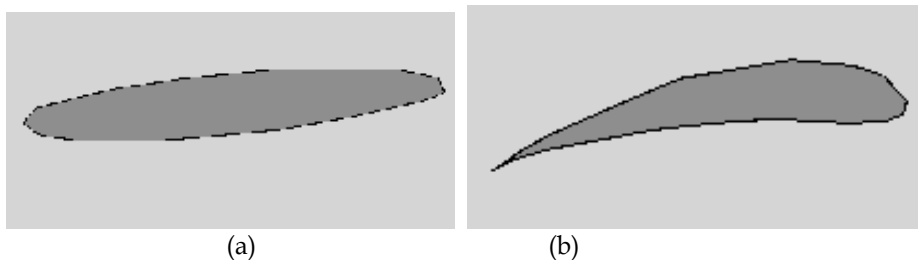


Fig. 74. Relationship between the inclusion and the strip matrix after rolling for 10-micron circle (a) and triangle (b) soft inclusions in the position of $1/4$ of strip thickness

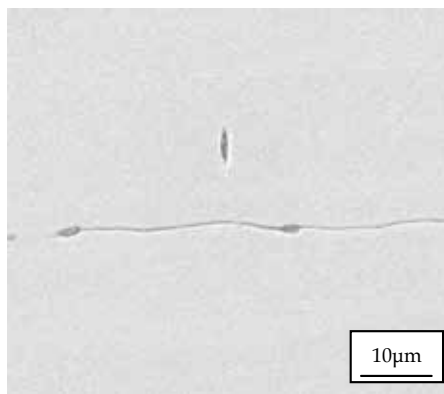


Fig. 75. SEM graph of the shape of soft inclusion in strip matrix after rolling

Table 14 shows the relationship between the inclusions and the strip matrix for various inclusion sizes both hard and soft inclusions when the inclusions are positioned at $1/4$ of the strip thickness. For hard inclusions, when the inclusions are circle, the cracks appear in front and rear of the inclusions and the crack direction is along the rolling direction. If the inclusions are square, two cracks appear in front and rear of inclusions separately and the crack direction is 45 degree to the rolling direction at the rolling conditions, meanwhile, the square inclusions deform to the ellipse inclusions similar to that of circle inclusions; when the inclusions are triangle, there just is one crack appear in rear of inclusions in the position of the strip thickness. When the inclusions are soft, there are no cracks around inclusions under various inclusion sizes for all the circle, square and triangle inclusions.

Table 15 shows the relationship between the inclusions of $20\ \mu\text{m}$ and the strip matrix under various inclusion positions for both hard and soft inclusions. For hard inclusions, when the inclusions are circle, the cracks appears in front and rear of the inclusions, and the crack size decreases with increase of the distance between the inclusions and the strip surface; when the shape of inclusions is square, two cracks appear in front and rear of inclusions which are in the position of $3/8$, $1/4$ and $1/8$ of the strip thickness, and four cracks appear in every corner of the square inclusions when the inclusions are in the position of $1/2$ of the strip

thickness; when the inclusions are triangle, one crack appears in rear of the inclusions that are in the position of $1/4$ and $1/8$ of the strip thickness and two cracks appears in front and rear of the inclusions that are in the position of $1/2$ and $3/8$ of the strip thickness. As to the soft inclusions, there are no cracks around inclusions after rolling process under various the positions of the inclusions along the strip thickness direction.



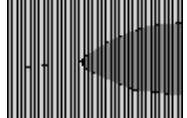

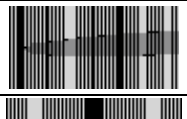








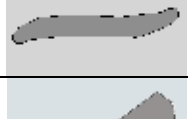






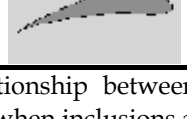
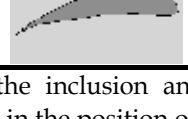
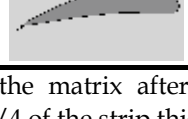
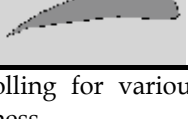
Inclusion size, μm		10	20	30	40
Circle inclusions	Hard				
	Soft				
Square inclusions	Hard				
	Soft				
Triangle inclusions	Hard				
	Soft				

Table 14. Relationship between the inclusion and the matrix after rolling for various inclusion sizes when inclusions are in the position of $1/4$ of the strip thickness

In tables 14 and 15, it also can be found that the inclusions rotate a little angle for both hard and soft inclusions. In table 14, the inclusions seemingly rotate with the same angle after rolling, so the inclusion sizes affect the rotation of inclusion slightly. And in table 15, when the inclusion is at the position of $1/2$ of strip thickness, the inclusions do not rotate after rolling process, and with the inclusions approaching to the strip surface, the rotation angle of inclusions increases. So the rotation of inclusion is clearly affected by the position of inclusion in strip.

Fig. 76 shows the crack length around the circle inclusions after rolling for various inclusion sizes and positions. **Fig. 76 (a)** shows the crack length in front of the inclusion, the crack length is about $3 \mu\text{m}$ when the size of inclusion is $10 \mu\text{m}$ which is in the position of $1/8$ of the strip thickness, and the crack length is about $17 \mu\text{m}$ when the inclusion size is $40 \mu\text{m}$ and the inclusion is in the position of $1/2$ of the strip thickness. The crack length increases as the inclusion size increases when the inclusions are in the same position of the strip, and which also increases with increase of the distance between the inclusion and the strip surface. **Fig. 76 (b)** shows the crack length in rear of the inclusions. The crack length in rear of inclusion

increases with increase of the inclusion sizes and the distance between the inclusion and the strip surface similar to that in front of inclusion.















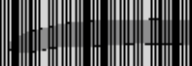



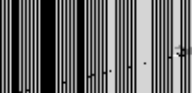
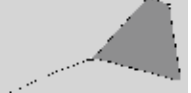



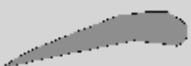
Inclusion position, $H_i/2H_0$		1/2	3/8	1/4	1/8
Circle inclusions	Hard				
	Soft				
Square inclusions	Hard				
	Soft				
Triangle inclusions	Hard				
	Soft				

Table 15. Relationship between inclusion and matrix after rolling for various inclusion positions in strip when inclusion size is 20 μm

Fig. 77 shows the crack length in front and rear of the square inclusions after rolling under various inclusion sizes and positions. As shown in the figure, the crack length in front of inclusion increase with increase of the inclusion sizes and the distances between the inclusion and the strip center. Compared **Fig. 77 (a) and (b)**, the crack length are similar between that in front of inclusion and that in rear of inclusion.

Fig. 78 shows the crack length in front and rear of the triangle inclusions after rolling under various inclusion sizes and positions. There are no cracks in front of inclusions when the inclusions are in the position of 1/8 and 1/4 of the strip thickness, and the crack length increases with increase of the inclusion sizes when the inclusions are in the position of 3/8 and 1/2 of the strip thickness. The crack length in rear of inclusions increases with increase of the inclusion sizes and with decrease of the distance between the inclusion and the strip surface. Compared the **Fig. 78 (a) and (b)**, the crack length in rear of the inclusions is larger than that in front of the inclusions when the inclusions are not in the position of 1/2 of strip thickness.

As shown in Figs 76 ~ 78, the crack size around inclusion after rolling process decreases with decreasing the inclusion size. If the inclusion size further decreases, the crack size also will decrease, and there will no cracks if the inclusion size is less than a certain value. Meanwhile, it is obvious that there are no cracks around soft inclusions after rolling, but cracks appear for hard inclusions. So with increasing the hardness of inclusion, the cracks will gradually increase which will reach the maximum value when the inclusions are rigid.

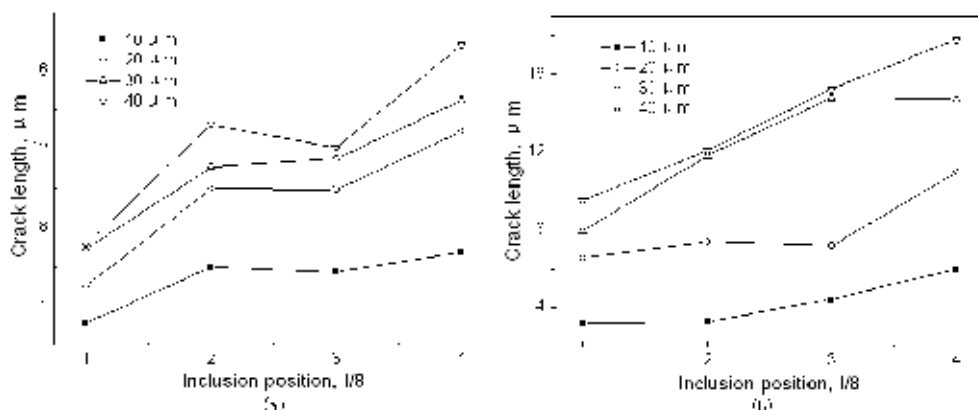


Fig. 76. Crack size in front (a) and rear (b) of circle inclusions

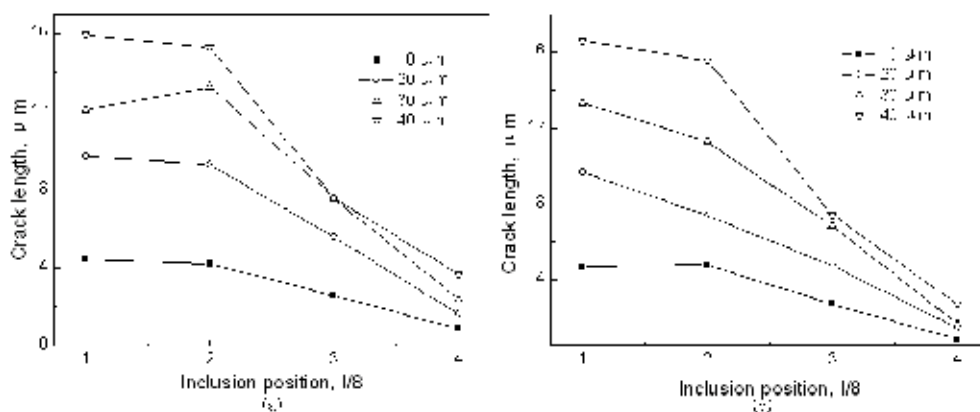


Fig. 77. Crack size in front (a) and rear (b) of square inclusions

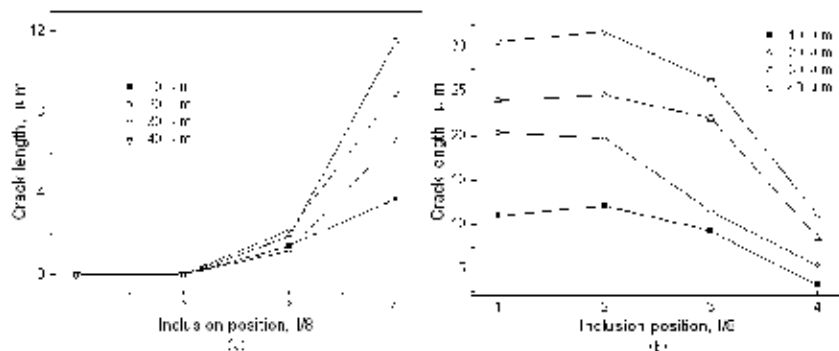


Fig. 78. Crack size in front (a) and rear (b) of triangle inclusions

Acknowledgements

The author gratefully acknowledges the financial supports by the Research Fund for the Doctoral Program of Higher Education through Grant No. 20090042120005. And the author also needs to thank the Prof. LIU Xianghua for his encouragement and guiding and helps, who is the previous director of State Key Laboratory of Rolling and Automation. His current

research fields include modelling and simulation of metal forming processes, advanced structural materials development and AI application in steel rolling process. He has been awarded three National Prizes of Science and Technology of China since 2000.

5. References

- [1] Yu H.L., Liu X.H. Coupling analysis to multipass V-H rolling process by FEM. *Hot Working Technology* 2007, 36(1):85-88. (in Chinese)
- [2] Li C.S., Yu H.L., Deng G.Y., Liu X.H. Numerical simulation to temperature field and thermal stress field of work roll in hot strip rolling. *Journal of Iron and Steel Research International*, 2007, 14(5): 18-23.
- [3] Yu H.L., Liu X.H. Friction and slab deformation during rough rolling. *International Journal of Surface Science and Engineering*, 2009, 3(5/6): 423-434.
- [4] Grass H., Krempaszky C., Reip T., Werner E. 3-D Simulation of hot forming and microstructure evolution. *Computational Materials Science* 2003, 28:469-477.
- [5] Yu H.L., Zhao X.M., Liu X.H. Tri-dimensional elastic-plastic FEM of rolling force during strip finishing rolling process. *Research on Iron and Steel*, 2005, 33(1): 14-16. (in Chinese)
- [6] Yu H.L., Liu X.H., Zhao X.M. Numerical analysis to rolling pressure distribution of plane view pattern controlling pass of plate rolling. *Journal of Plasticity Engineering*, 2005, 12(4): 51-53, 75. (in Chinese)
- [7] Yu H.L., Liu X.H., Li C.S. Influence of groove fillet radii of edger roll on slab deformation. *Iron and Steel*, 2006, 6: 47-51. (in Chinese)
- [8] Yu H.L., Liu X.H., Jiao Z.J., Zhao X.M. Application of geometry model updating method in FEM analysis of plane pattern of plate. *Journal of Iron and Steel Research*, 2006, 18(2): 26-29. (in Chinese)
- [9] Yu H.L., Liu X.H., Lee G.T. Analysis to rolls deflection of Sendzimir mill by 3D FEM. *Transactions of Nonferrous Metals Society of China*, 2007, 17(3): 600-605.
- [10] Knapik M. The numerical analysis of roll deflection during plate rolling. *Journal of Material Processing Technology*, 2006, 175: 257-265.
- [11] Yu H.L., Liu X.H., Yu Q.B. Tracing of macrosegregation in slab during multipass rolling. Submitted.
- [12] Li C.S., Song S.N., Mei R.B., Zhang G.L., Liu X.H. Online Arithmetic by fast finite element in strip rolling. *Journal of Mechanical Engineering*, 2009, 45(6): 193-198.
- [13] Liu X.H. 2.7 dimensional RPFEM analysis and its application in rolling. *Journal of Northeastern University of Technology*, 1990, 11(5): 468-473.
- [14] Wang L. Application of modified Cholesky factorization forcing matrix to be positive definite in rolling. Shenyang: Northeastern University, 2008.
- [15] Si LY, Lu C, Huynh NN, et al. Simulation of rolling behaviour of cubic oriented al single crystal with crystal plasticity FEM. *Journal of Materials Processing Technology*, , 2008, 201(1-3): 79-84.
- [16] Lue C, Si LY, Zhu HT, et al. Theoretical analysis of texture evolution in cold rolling process of single crystal aluminum. *Transactions of Nonferrous Metals Society of China*, 2007, 17, Sp.1: S254-S257.

- [17] Yu H.L., Liu X.H., Wang G.D. Analysis of crack tip stress of transversal crack on slab corner during V-H rolling process by FEM. *Journal of Iron and Steel Research International*, 2008, 15(3): 19-26.
- [18] Yu H.L., Liu X.H., Chen L.Q., Li C.S., Zhi Y., Li X.W. Influence of edge rolling reduction on plate-edge stress during finish rolling process, *Journal of Iron and Steel Research International*, 2009, 16(1), 22-26.
- [19] Yu H.L., Liu X.H. Behavior of subsurface blisters in slab during hot rolling using FEM. *The Second International Conference on Modelling and Simulation*, Manchester, 2009 May, 37-42.
- [20] Yu H.L., Liu X.H., Li C.S., Chen L.Q. Research on behavior of slab surface defects in forward slip zone during V-H rolling process. *Materials Science Forum*, 2008, 575-578: 243-248.
- [21] Yu H.L., Liu X.H., Bi H.Y., Chen L.Q. Research on behavior of inclusion in stainless steel during rolling by FEM. *Steel Research International*, 2008, 79, Special edition, Vol.1: 389-396.
- [22] Qunpanich D., Mutoh Y., Yoshii K. Fatigue behavior of hot-rolled steel plate at various stages of fabrication. *Materials Science and Engineering A* 2008, 477(1-2): 9-14.
- [23] Yamamoto K. Study on surface defects of coated sheets caused by the surface condition of a hot roughing mill roll. *SEASIS Quarterly (South East Asia Iron and Steel Institute)* 2007, 36 (3):54-58.
- [24] Riedel H., Andrieux F., Walde T., Karhausen K.-F. The formation of edge cracks during rolling of metal sheet. *Steel Research International* 2007, 78 (10-11): 818-824.
- [25] Ko D.C., Lee S.H., Kim D.H., Byon S.M., Park H.D., Kim B.M. Design of sizing press anvil for decrease of defect in hot strip. *Journal of Materials Processing Technology* 2007, 187-188: 738-742.
- [26] Burton D. S., Gao X., Brinson L.C. Finite element simulation of a self-healing shape memory alloy composite. *Mechanics of Materials* 2006, 38: 525-537.
- [27] Awais M., Lee H.W., Im Y.T., Kwon H.C., Byon S.M., Park H.D. Plastic work approach for surface defect prediction in the hot bar rolling process. *Journal of Materials Processing Technology* 2008, 201(1-3): 73-78.
- [28] Son I.H., Lee H.D., Choi S., Lee D.L., Im Y.T. Deformation behavior of the surface defects of low carbon steel in wire rod rolling. *Journal of Materials Processing Technology* 2008, 201(1-3): 91-96.
- [29] Ervasti E., Ståhlberg U. Behavior of longitudinal surface cracks in the hot rolling of steel slabs. *Journal of Material Processing Technology* 1999, 94(2): 141-150.
- [30] Ervasti E., Ståhlberg U. Transversal cracks and their behavior in hot rolling of steel slabs. *Journal of Material Processing Technology* 2000, 101(1) : 312-321.
- [31] Yukawa N., Ishikawa T., Yoshida Y., Koyachi A. Influence of rolling condition on deformation of surface micro-defect in plate rolling. *Tetsu-to-Hagané* 2005, 91(12): 861-867.
- [32] Yukawa N., Yoshida Y., Ishikawa T. Deformation analysis of longitudinal surface micro-defects in flat rolling. *Tetsu-to-Hagané*, 2006, 92(11): 19-24.
- [33] Liu X.H.. Rigid finite element method and its application in rolling process. *Metallurgical Industry Press*, Beijing, 1994. (in Chinese)
- [34] Sun Y.K.. Numerical models in hot flat continuous rolling. *Metallurgical Industry Press*, Beijing, 1979.(in Chinese)

- [35] Wang M., Yang H., Sun Z.C., Guo L.G., Qu X.Z. Dynamic explicit FE modeling of hot ring rolling process. Transactions of Nonferrous Metals Society of China 2006, 16: 1274-1280.
- [36] Yu H.L., Liu X.H. Thermal-mechanical FE analysis of evolution behavior of slab surface cracks during hot rolling, Materials and Manufacturing Processes, 2009, 24(5): 570-578.
- [37] Wei D., Han J., Jiang Z.Y., Tieu A.K. A study on crack healing in 1045 steel. Journal of Materials Processing Technology, 2006, 177(1-3): 233-237.
- [38] Zhou Y., Guo J., Gao M., He G. Crack healing in a steel by using electropulsing technique. Materials Letters 2004, 58(11): 1732-1736.
- [39] Xiong, S.W., Rodrigues, J.M.C. and Martins, P.A.F. Three-dimensional modeling of the vertical-horizontal rolling process. Finite Elements in Analysis and Design, 2003, 39: 1023-1037.
- [40] Jiang, Z.Y., Tieu, A.K. and Lu, C. Mechanical modeling of friction variation in slab edging with a finite element method approach, Tribology International, 2004, 37: 733-742.
- [41] Yu H.L., Liu X.H., Li C.S. FEM analysis to the slab corner metal flow condition during multi-pass V-H rolling process, 2005, 26(10): 982-985. (in Chinese)
- [42] Xie H., Xiao H., Zhang G., et al. Analysis on rolling pressure distribution of strip width different reductions by explicit dynamic FEM, Journal of Iron Steel Research, 2002, 14(6), 33-35.(in Chinese)
- [43] Ding G.: Algorithms and methods of ANSYS/LS-DYNA, Beijing Institute of Technology, Beijing, 1999, pp.104.
- [44] Yu H.L., Liu X.H., Zhao X.M., Kusaba Y. FEM Analysis for V-H rolling process by updating geometric method. Journal of Materials Processes Technology, 2006, 180(1-3): 323-327.
- [45] Yu H.L., Liu X.H., Li C.S., Lan F.F., Wang G.D. Research on the behavior of transversal crack in slab V-H rolling process by FEM. Journal of Materials Processes Technology, 2009, 209(6): 2876-2886.
- [46] Yu H.L., Liu X.H., Ren X.J. Behavior of longitudinal cracks in slab V-H rolling processes. Steel Research International, 2008, 79(7): 537-543.
- [47] Cai K.K., Han C.J., Sun Y.H. Crack control during continuous casting. Conference on the Continuous Casting Slab Technology, Nanjing, 2006, 1-15. (In Chinese)
- [48] Yu H.L., Liu X.H., Jiang Z.Y. FE analysis of behavior of internal-longitudinal cracks in slab during vertical-horizontal rolling. Proceeding of The 8th Asia Pacific Conference on Materials Processing, 2008, 6, Guilin-Guangzhou, China, p.799-804.
- [49] Liu X.H., Yu H.L. Analysis of crack healing in hot rolled deformed piece. Journal of Iron and Steel Research, 2009, 21(6): 38-41. (in Chinese)
- [50] Yu H.L., Liu X.H., Li C.S., Zhao X.M., Kusaba Y. Influences on the plastic strain distribution of slab of vertical roll shape during multi-pass V-H rolling process. Acta Metallurgical Sinica.(English Letters), 2006, 19(1): 51-56.
- [51] Matsumiya T., 2006. Recent topics of research and development in continuous casting. ISIJ Int. 46(12), 1800-1804.
- [52] Todoroki H., Mizuno K., 2004. Effect of silica in slag on inclusion compositions in 304 stainless steel deoxidized with aluminum. ISIJ Int. 44(8),1350-1357.

- [53] Luo C.H., Ståhlberg U., 2001. Deformation of inclusions during hot rolling of steels. *Journal of Material Processing Technology* 114, 87-97.
- [54] Luo C.H., Ståhlberg U., 2002. An alternative way for evaluating the deformation of MnS inclusion in hot rolling of steel. *Scand. J. Metall.* 31(1), 184-190.
- [55] Ervasti E., Ståhlberg U., 2005. Void initiation close to a macro-inclusion during single pass reductions in the hot rolling of steel slabs: A numerical study. *Journal of Material Processing Technology* 170, 142-150.
- [56] Hwang Y.M., Chen D.C., 2003. Analysis of the deformation mechanism of void generation and development around inclusions inside the sheet during sheet rolling process. *Proc. Instn Mech. Engrs Part B: J. Eng. Manuf.* 217(10), 1373-1381.
- [57] Melander A., 1997. A finite element study of short cracks with different inclusion types under rolling contact fatigue load. *Int. J. Fatigue.* 19(1), 13-24.
- [58] Zhang S.T., 2001. *Steel Materials Handbook- Stainless Steel*, Standards Press of China, Beijing.
- [59] Li R.J., 1995. *Ceramic-Metal Complex Materials*, Metallurgical Industry Press, Beijing.
- [60] Yu H.L., Liu X.H., Bi H.Y., Chen L.Q. Deformation behavior of inclusions in stainless steel strips during multi-pass cold rolling. *Journal of Materials Processing Technology*, 2009, 209(1): 455-461.
- [61] Yu H.L., Bi H.Y., Liu X.H., Chen L.Q. Numerical analysis of movement of inclusion in stainless steel during cold rolling. *Proceedings of First International Conference of Modelling and Simulation, Vol.V, Modelling and Simulation in Mechanics and Manufacture*, Nanjing, China, 2008, 8: 486-491.
- [62] Wei D.B., Han J. T., Tieu K., Jiang Z.Y.. Simulation of crack healing in BCC Fe. *Scripta Mater*, 2004, 51(6) , 583-587.
- [63] Yu H.L., Liu X.H., Li X.W. Crack healing in low carbon steels under hot plastic deformation. *Materials and Manufacturing Processes*, 2010, 25(4), in publish.
- [64] Ma Q.X., Wang Z.C., Zhong Y.X. The mechanism of faults originating from inclusions in the plastic deformation process of heavy forging. *Journal of Material Processing Technology*, 2002, 123: 61-66.
- [65] Zhang G.P., Takashima K., Higo Y.. Fatigue strength of small-scale type 304 stainless steel thin films. *Mater. Sci. Eng. A*, 2006, 426: 95-100.
- [66] Wang A., Thomson P.F., Hodgson P.D. A study of pore closure and welding in hot rolling process. *Journal of Material Processing Technology* , 1996, 60: 95-102.
- [67] Yu H.L., Bi H.Y., Liu X.H., Tu Y.F. Strain distribution of strips with a spherical inclusion during cold rolling. *Transactions of Nonferrous Metals Society of China*, 2008, 18(4): 919-924.
- [68] Yu H.L., Liu X.H., Li X.W. Analysis of inclusion deformation and crack generation during rolling with a super-thin layer by FEM. *Materials Letters*, 2008, 62(10-11): 1595-1598.
- [69] Xue Z.L., Weng Y.Q., Li, Z. B. Zero inclusion steel and its oxide inclusion behavior. *World Iron & Steel* , 2008, 8(4): 17-20. (in Chinese)
- [70] Zhao Z.Y., 1999. *Mechanics in Metal Forming*, Metallurgical Industry Press, Beijing.
- [71] Yu H.L., Bi H.Y., Liu X.H., Chen L.Q., Dong N.N. Behavior of inclusions with weak adhesion to strip matrix during rolling using FEM, *Journal of Materials Processing Technology*, 2009, 209(9): 4274-4280.

Finite element analysis of strip and rolling mills

Jian-guo CAO, Jie ZHANG, Ning KONG and Kai-fu MI
School of Mechanical Engineering, University of Science and Technology Beijing
P.R.China

1. Introduction to rolling process of strip

In metalworking, rolling is a metal forming process in which metal stock is passed through a pair of rolls. Compared with forging, extrusion, drawing and other processes, rolling has been more widely applied by high productivity, variety, continuity of the production process, easy mechanization and automation and information advantages. Steel is the largest consumption metal materials. Currently, about 90% of the steel is rolled into strip, section, tubes and wires and so on. As the backbone of steel products, strip that is so-called "universal steel" has widely been used in national economic departments of major products. Rolling of strip is classified according to the temperature of the metal rolled. If the rolling temperature of the metal is above its recrystallization temperature, then the process is termed as hot rolling. If the temperature of the steel is below its recrystallization temperature, the process is termed as cold rolling. In terms of usage, rolling processes of strip play an important roll in the manufacturing industry.

1.1 Hot rolling process

The whole hot rolling process is shown in Fig.1. This casting slab is room temperature in common and when it is taken into the hot rolling process, it should be heated in the heating furnace. In this stage, the slab must be heated to the temperature between 1050 and 1280°C. The temperature must be monitored to make sure it remains up to the temperature required, and then it will be taken out of heating furnace by slab extractor and moved to the next stage called descaling by high pressure water. When the slab is moved to the descaling box by roll table, the high pressure water flushes the slab so as to remove the iron oxide skins and avoid scratching the rolls and strip. The following stage is rough rolling process. It often contains one or two roughing mills in which the slab is hot rolled reversibly. When the slab arrives, it will be rolled 5 or 7 times repeatedly to reach the thickness requirement. What is also worth mentioning is that the roughing mill contains edger rolls which are used to roll the edge of slab and center it. After rolled by the roughing mill, the slab is called transfer bar

in common and it goes to which can uniform the temperature and decrease temperature drop of the whole transfer bar. Then the transfer bar will be cut the front and end by flying shear. The later stage is finishing rolling process which is the most important and complex process in hot rolling. This part contains seven 4-high rolling mill stands from F1 to F7, which contains kinds of effective control methods such as shifting and bending of work rolls. Transfer slab is rolled by the seven mill stands so as to reach the strip control requirements and then goes to the laminar cooling section. In this section, the strip will be cooled to the required temperature according to mechanical properties needed. The last stage is coiler, and then the strip will be sent to the product room. Also, some strips should be rolled by temper mill if needed and after that the whole hot rolling process is finished.

Hot rolling is used mainly to produce hot rolling strip steel, which provides to the raw material of cold rolling and to the process of equipment such as container.

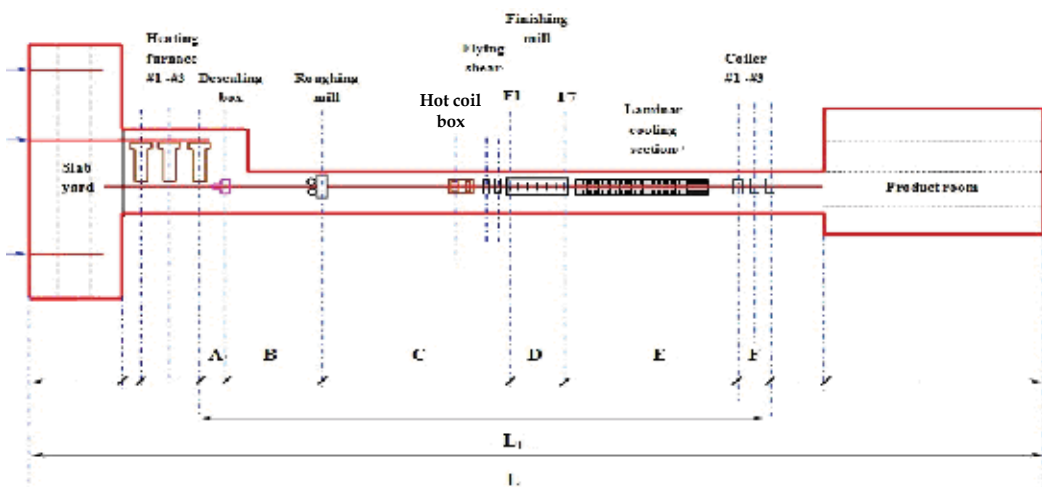


Fig. 1. Layout diagram of main equipment in hot rolling process

1.2 Cold rolling process

Cold rolling occurs with the steel below its recrystallization temperature (usually at room temperature), which increases the strength via strain. It also improves the surface finish and holds tighter tolerances. Commonly cold-rolled strips are usually thinner than the same products that are hot rolled. Because of their smaller size and greater strength, as compared to hot-rolled strips, 4-high, 6-high or reversing mills are used. But cold rolling cannot reduce the thickness of strips as much as hot rolling process. The whole cold rolling process reveals as follows.

Raw material of cold rolling is Hot-rolled strips. Steel strips are welded one after another by welder so that they are linked together. This is so called continuous rolling. These continuous strips are then sent to the pickling section for removing the iron oxide skins by sulfuric acid or hydrochloric acid in common. After that, the strips will be cleaned, dried, cut edge and sub-volumed. After acid washing machine, in order to roll the strips to final thickness and required strip profile and flatness, strips should be rolled by 4-high or 6-high tandem cold rolling mill of five stands named from S1 to S5 while generally without any intermediate annealing. In this stage, it takes into account of strip quality, rolling force,

allocated reduction and other factors for this tandem rolling process. The following stage is annealing used to eliminate cold hardening and soften the recrystallization strip steel so as to acquire the good ability of plasticity. The next stage is rolling the strip by temper mill. The last stage is galvanizing, tinning or colour coating of the cold-rolled strips, according to the requirements. Then these strips will be cut and packaged and the whole tandem cold rolling process is finished. Cold rolling is used mainly to produce cold-rolled strip, which provides to the production of auto sheet, home appliances sheet, electrical steel pieces and so on.

The main deformation processes of wide strip are rough and finishing rolling process of hot strip mills, tandem cold rolling process. New generation high-tech mills for profile and flatness control in wide strip rolling, such as 4-high or 6-high CVC (continuously variable crown) mill, SmartCrown mill, K-WRS (Kawasaki steel work-roll shifting) mill, ASR (asymmetry self-compensating rolling) mill, UCM (universal crown mill), UCMW mill includes an work roll shifting system in addition to the strip profile and flatness functions that are provided in the UCM mill, ECC (edge drop & crown compact) mill, T-WRS&C (taper work roll shift and cross) mill (Ginzburg V. B., 1993; Chen X.L. 1997; Cao J.G., 2006), have been developed and applied to the production of strip.

2. Finite element analysis of spalling on the backup roll of a roughing mill for stainless steel

Rolls are the important consuming parts that affect the efficiency of large industrial mills and the quality of mill products. The accidents of rolls, such as spalling, cracks and roll breakage, may happen during the service period, in which spalling is the primary form of damage (Li, H.C., 2007). There are a lot of factors that could lead to roll spalling including thermal shock loads, unreasonable distribution of roll surface pressure, rolling accidents, and inadequate cooling and so on (Ray, A.K., 2000; Chen, S.G., 2006).

2250 HSM (hot strip mill) of TISCO (Taiyuan Iron & Steel (Group) Co. LTD.) in China put into operation on June 29, 2006. The annual capacity of 2250 HSM is 4 million tons including 2 million tons of stainless steel, which is the largest stainless steel production equipment, and is the most advanced technology and most complete hot rolling production line in the world. This strip mill just set up a roughing mill for adapting to the stringent requirements on the temperature when rolling stainless steel slab, which leads to harsh conditions of rolling process, rolling difficulty, the complex mechanical behavior during service periods and less stable rolling process control performance. This roughing mill Occurred continuously 3 spalling accidents in a month, and each accident not only causes direct economic losses but also incident handling time is up to 2h even 5 ~ 6h, which directly threat to a long-term stable and normal operation of the 2250 HSM production line. So the study of spalling on the backup roll of a roughing mill for stainless steel sheet has important theoretical significance and engineering applications.

2.1 Backup roll spalling of 2250 roughing mill

3 spalling accidents occurred continuously in this roughing mill within a month of 2250 HSM on March 8, April 4 and April 8, 2008, e.g., the roughing rolls diameter of which is 1552mm occurred serious edge spalling in the April 4, 2008, the circumferential length of spalling is 4023mm (reached more than 82% of perimeter), the axial width is 1250mm (reached half of the length of roll body) and the maximum radius depth is

185mm(far more than the use range of the layer radius thickness 80mm of the new backup roll). The actual spalling fracture appearances are shown in Fig.2.



Fig. 2. Spalling of backup rolls in roughing mill of 2250HSM

A lot of tracking tests on the work roll and backup roll wear contours of the roughing mill have been finished. In general,work roll wear contours are the U-typed wear and the partial peak exists. Backup roll wear is generally non-uniform along the entire length of the roll barrel and there are overall and local backup roll wear contours, i.e. “cat ear” near strip edges.

Through the above analysis shows:(1)The alternating shear stress of A certain depth from the roll surface caused by roll contact pressure is the key reason of the backup roll spalling;(2)Spalling usually occurs in the maximum contact pressure area. The U-typed wear of conventional rolls makes contact stress between the edge of backup and work rolls increase significantly, contact stress between rolls in the middle of the roll had little change and spalling in the edge of the roll occurs sometimes. So the ideal solution is obtains uniform axial contact pressure distribution by improved backup roll and work roll contours(Cao, J. G.,1999).

2.2 Finite element analysis of backup roll spalling

A 3D FEM model of the rolls system should be modeled for the full analysis to 2250 wide strip roughing mill. The finite element model based on ANSYS softpackage(specific unit allocation of finite element model is shown in Table 1) is shown in Fig.3 (Yang, G.H., 2008). Axial distribution of contact stress between work roll and backup roll could be influenced by strip width, unit width rolling force and wear in different service period, and the above factors are analyzed individually by this model.

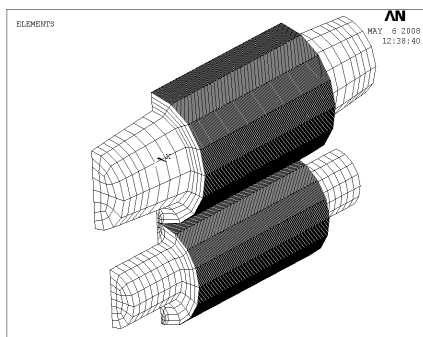


Fig. 3. Three-dimensional finite element roll system model of a 4-high mill

Work Roll	Number of solid elements	20184
	Number of contact elements	900
Backup Roll	Number of solid elements	15684
	Number of contact elements	900
The total number of units		35868
The total number of nodes		68814

Table 1. Unit allocation of finite element model

Fig.4 shows the axial distribution of contact stress between rolls of different strip width when unit width rolling force is 25kN/mm, and adopts the backup roll contour provided by SMS group. It could be known by comparative analysis: the peak of contact stress between rolls gradually increases with the increase of strip width. When the strip width B changes from 1010mm to 2050mm, the peak of contact stress between rolls increase 106.3%. Distribution trends of contact stress between rolls are the same for different width strips, and the non-uniformity coefficient of contact stress distribution between rolls are respectively 1.36, 1.38 and 1.41. It could be seen by analysis of contact pressure between rolls that the peak position of contact stress between rolls is essentially the same which is about 200mm from the edge side of backup roll barrel.

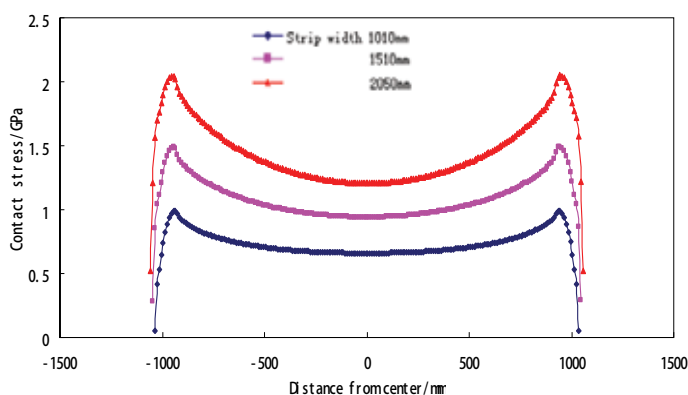
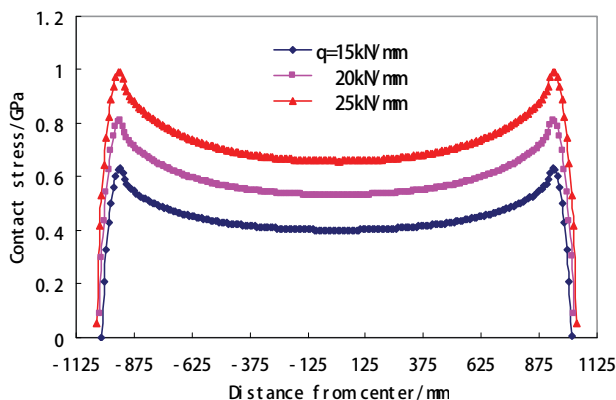
Fig. 4. Contact stress distribution of different strip width ($q=25 \text{ kN/mm}$)

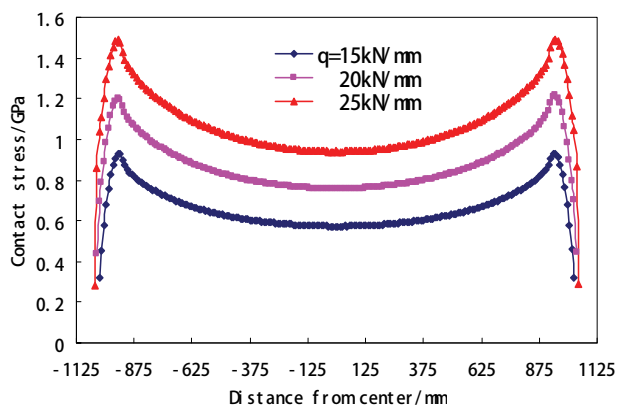
Fig. 5 shows the axial distribution of contact stress between rolls for the strip width 1010mm, 1510mm, 2050mm, when rolling force per strip width is 15kN/mm, 20kN/mm and 25kN/mm, and adopts the backup roll contour provided by SMS group. It could be known by comparative analysis: when the rolling force per strip width changes from 15kN/mm to 25kN/mm, the peak of contact stress between rolls increase respectively by 56.8%, 60.7%, 59.3% for the strip width 1010mm, 1510mm and 2050mm, the non-uniformity coefficient of contact stress distribution between rolls decreases with the increase of the unit width rolling force, the non-uniformity coefficient decreases respectively from 1.42 to 1.36, from 1.41 to

1.38, from 1.45 to 1.41 for the strip width 1010mm, 1510mm and 2050mm. Distribution trends of contact stress between rolls are the same for different rolling force per strip width, and the peak position of contact stress between rolls is essentially the same which is about 200mm from the edge side of backup roll barrel.

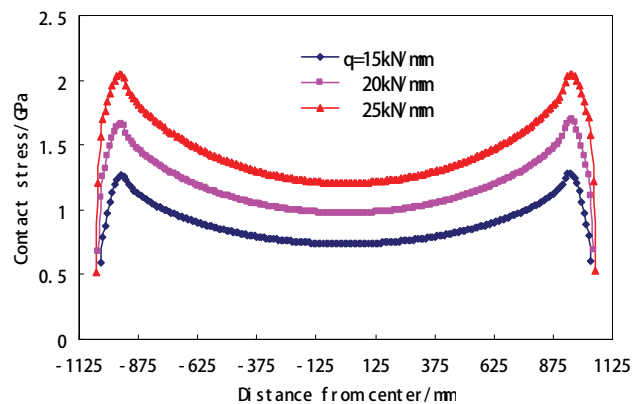
Fig. 6 shows the axial distribution of contact stress between rolls when the rolling force per strip width is 25kN/mm, work roll and backup roll are at different wear periods and adopts the backup roll contour provided by SMS group. WR signifies work roll, BR signifies backup roll, G signifies the roll contour in the early service period, W signifies the roll contour in the late service period. The three conditions in this figure is respectively: work roll adopts the contour in the early service period and backup roll adopts the contour in the late service period; work roll and backup roll both adopt the contour in the early service period; work roll and backup roll both adopt the contour in the late service period. It could be known by comparative analysis: when work roll is in the early service period, the peak of contact stress between rolls is slightly increases but not significantly as backup roll is in the late service period than in the early service period for the strip width 1010mm, 1510mm and 2050mm, and the asymmetry degree of contact pressure distribution between rolls basically



(a) Strip width 1010mm

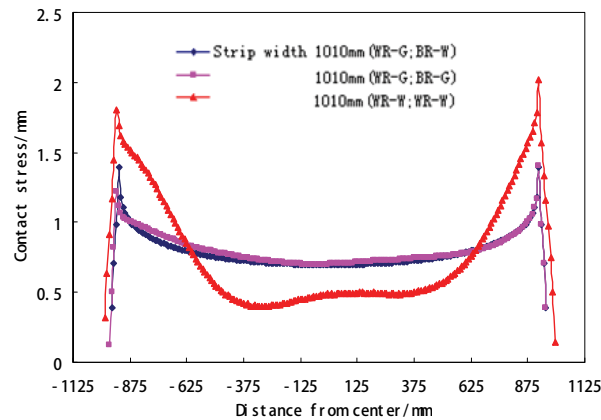


(b) Strip width 1510mm

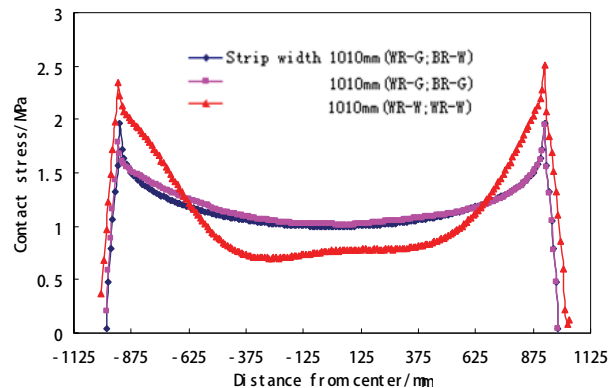


(c) Strip width 2050mm

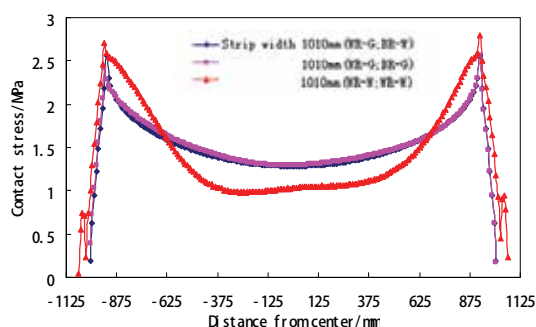
Fig. 5. Contact stress distribution of different rolling force



(a) Strip width1010mm



(b) Strip width 1510mm



(c) Strip width 2050mm

Fig. 6. Contact stress distribution of different service period

unchanged. When work roll and backup roll both are in the early service period, the peak of contact stress between rolls increase very significantly, the range of which is respectively 43.7%, 28.5% and 9.2%. The asymmetry coefficient of contact stress distribution between rolls increases respectively from 1.76 to 2.61, from 1.68 to 2.26, from 1.66 to 1.95. In comparison, there is greater impact by the wear when rolling narrow strip.

2.3 Simulation Conclusions

The simulation results show: the peak of the roll contact stress increase sharply when strip width and rolling force per width changing. And the peak and the dissymmetry of the roll contact stress increase when the work rolls and backup rolls are at the middle and late stage of the rolling campaign. Further more, there exists peak stress at the edge of backup roll at a distance of 200mm, which closes to actual experiences. Based on this study, according to the principle of varying contact rolling (Chen, X.L., 1994, 2000), The SCR (smart contact backup rolls) and matching work roll contours for different stage of rolling campaigns technology have been developed and applied to the 2250 hot strip mill at TISCO of China and gains a good result without spalled backup rolls during the continuous production of 6.082 million tons and more from July 4, 2008 until now.

3. Finite element analysis of roll contour configuration on the flatness control performance of non-oriented electrical steel sheets in hot strip rolling mills

Cold-rolled non-oriented electrical steel sheet, widely used as an iron core in electric devices, automobile motors and reactors, is given priority to development by China. National key production base of cold-rolled non-oriented electrical steel sheet, i.e. WISCO (Wuhan Iron & Steel (Group) Corp.), Baosteel (Shanghai Baosteel Group Corp.), TISCO (Taiyuan Iron & Steel (Group) Co. LTD.), and ANSTEEL (Angang Steel Co. LTD.), are expending production capacity of cold-rolled electrical steel sheet and other backbone enterprises which have the potential productivity are to plan and prepare production lines of cold-rolled electrical steel sheet. Recently with the automation level and energy requirements improvement of industrial users, the precision of profile and flatness of cold-rolled non-oriented electrical steel sheet is becoming an important factor in determining strip quality and mill

productivity. According to industrial tracking test results and theory study, we found that the finishing rolling process of HSM is the critical procedure of profile and flatness control performance of non-oriented electrical steel sheets (Cao, J. G. & Ou, Y.Y., 2005; Cao, J. G., 2006). In the chapter, we investigated the 1700mm hot rolling mill in WISCO that have produced the largest amount and the most comprehensive available cold-rolled electrical steel sheets in China. The finite element models were developed to calculate the profile and flatness control performance of different typical roll contours configurations in 1700 HSM.

3.1 Roll contour configuration and variety on electrical steel sheets in hot rolling

The 1700 mm hot strip mill at WISCO was designed and built by Ishikawajima-Harima Heavy Industries Co., Ltd. (IHI) and was commissioned in 1978. From 1992 to 1994, the mill completed a series of upgrades, i.e. the profile and flatness control system and mathematics model of high precision, on the finishing mill supplied by SIEMENS VAI CLECIM. The mill is one of the largest hot strip mills in China with a capacity of 3.45 Mt/y. The control system has the following characteristics after a series of upgrades: F4 through F7 apply K-WRS with a long stroke work roll shifting system ($\pm 150\text{mm}$), work roll heavy-duty bending system (200t/chock), conventional backup and work rolls. From 1997 to 2000, VCR (varying contact backup rolls) technology developed by USTB and WISCO has applied to F1 through F7 of 1700mm HSM (Chen, X.L., 2000). The rolling period of a newly ground backup roll set in hot strip mills exceeds two weeks, the varying contact backup rolls, being geared to different steel kinds and sizes, only go halfway towards solving the problem of significant transverse thickness profile difference by wear contour patterns of conventional work roll contours. From 2003 to 2008, the developed ASR (asymmetry self-compensating work rolls) technology has applied to the production of 2.3 mm \times 1050 mm conventional width electrical steel sheets rolling campaign successfully since November 2004 (Cao, J. G., 2006) and 2.3 mm \times 1200-1300 mm wide electrical steel sheets rolling campaign since October 2005 (Cao, J. G., 2008). The 4-high ASR mill with work roll shifting devices is functionally superior to other mills, such as CVC mill, SmartCrown mill and K-WRS mill, when they have ability of both wear control and overall profile control covering the strip crown, edge drop and high spot in downstream stand and suitable to the schedule free rolling of hot strip mills. The ASR work roll contours for stand F5 of 1700 hot strip mill with a 72-coil electrical steel rolling campaign is shown in Fig. 7.

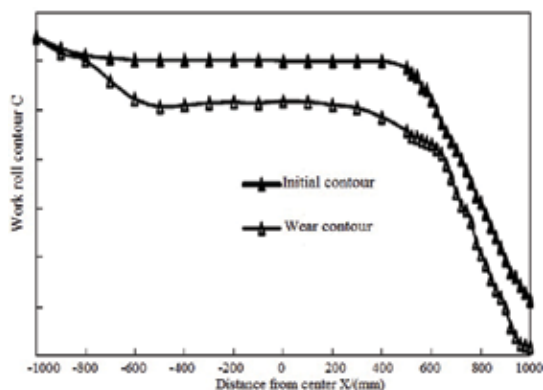


Fig. 7. ASR work roll contours for stand F5 of 1700 hot strip mill with a 72-coil electrical steel rolling campaign (2.55 mm \times 1280 mm)

During rolls servicing period, the change of work roll and backup roll contours which are resulted by wear and thermal expansion are important conditions that determine the mill type and control ability. The CVC and SmartCrown mill can provide a wide crown control range that is incapable of the wear control. The severe wear of work roll in downstream results in the failure and distortion of crown control ability for CVC, SmartCrown and CVC plus mills. The K-WRS mill is an effective way to alleviate the severe work roll wear contours by application of long stroke work roll shifting system with conventional work roll contours that is incapable of the crown control. In this chapter, we investigated three typical roll contour configurations of hot rolling mills in downstream stands for non-oriented electrical steel sheets in 1700 mm HSM, i.e. conventional backup/work rolls of K-WRS mill, VCR/conventional work roll contours, and VCR/ASR contours of ASR mill.

Fig. 8 shows the work roll contours measured in the conventional work roll (initial and wear contours) on the F5 stand in the rolling of the non-oriented electrical steel sheets and the match temperature of the work rolls. Conventional backup roll contours of 145 rolls, VCR of 94 rolls, conventional work roll contours of 37 rolls and ASR of 14 rolls that were used in WISCO 1700mm hot finishing mill have been measured. According to the large amount of comparison of work roll and backup roll (initial and wear contours), the following conclusions can be concluded: (1)The roll wear contour did not come out the “cat ear” roll contour from F1 through F3, but the roll wear contours, i.e. “cat ear” near strip edges (about -500mm to 500mm near strip edges), were significant from F4 to F7; (2) A maximum value of the work roll wear appears in stand F4, whereas the rolls in stands F1 and F3 show a very small increase in roll wear and the rolls in stands F5, F6 and F7 wear rapidly but smaller than those in stand F4; (3) Comparing to conventional strip, the value of work roll wear contour of non-oriented electrical steel sheets was larger, even reached to two times or three times larger; (4) If the conventional backup roll was applied, it would cause uneven wear resulting in bad self-maintenance of the roll contour from the top to the end stands; (5) If VCR was applied, the work roll wear is generally uniform along the entire length of the roll barrel.

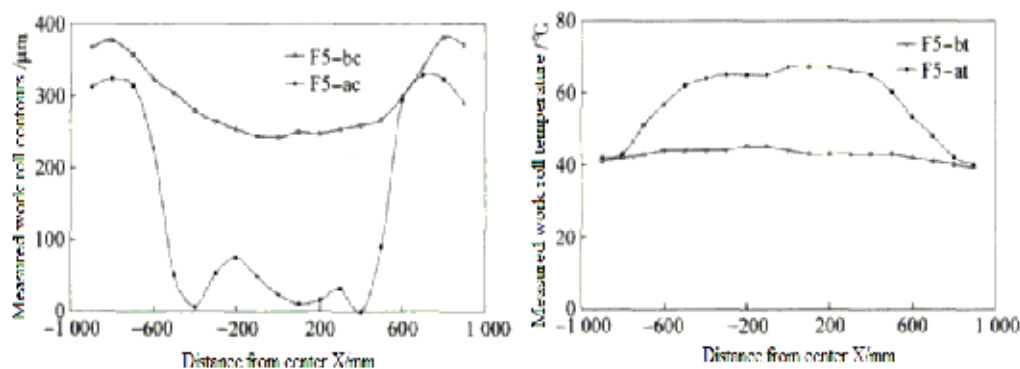


Fig. 8. Measured work roll contours and surface temperatures of non-oriented electrical steel sheet rolling campaign

3.2 Finite element analysis model of the roll stacks

In order to analyze the effect of roll contour configuration on the profile and flatness control performance of non-oriented electrical steel sheets in hot rolling, the two-dimensional finite

element analysis model with varying thickness of the 4-high hot finishing mill system, as shown in Fig. 9, was developed for effect of typical roll contour configuration on the

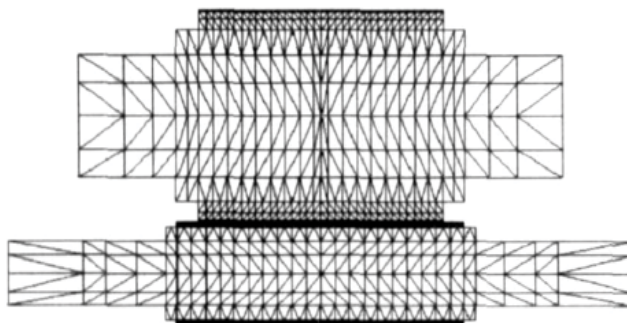


Fig. 9. Two-dimensional finite element model with varying thickness of roll stacks in 1 700 mm hot strip mill

characteristics of roll gap profile. In the model, adaptive contact elements are set at the contact interface between rolls, equivalent thickness on theory are set according to the equal principle of anti-flattening, and the thickness of physical units in each layer of the rolls is set by the equivalent principle of anti-bending modulus. The model has characteristics of high accuracy, fast calculating speed and a wide range engineering applications.

3.3 Profile and flatness control performance of electrical steel in hot rolling

As the disturbance variables, such as the change in roll contours, strip width and the roll force fluctuation, exist during the rolling, especially in downstream stands, leading to the variation of the roll gap crown, they are expected that the crown control range by the roll bending force can increase and the roll gap can maintain its crown relatively stable. The roll gap stiffness can be defined as the ratio of the roll force fluctuation to the correspondingly change of the roll gap crown. The effect of roll contours configuration of the ASR mill on characteristic of roll gap profile applied in 1 700 mm hot strip mill (F5) is shown in Fig 10. In Fig.10, CN represents conventional backup and work rolls of the K-WRS mill, 0% of roll bending force; CX represents conventional backup and work rolls of the K-WRS mill, 100% of roll bending force; VN represents varying contact backup rolls and conventional work rolls, 0% of roll bending force; VX represents varying contact backup rolls and conventional work rolls, 100% of roll bending force; AN represents varying contact backup rolls and asymmetry self-compensating work rolls of the ASR mill, 0% of roll bending force; AX represents varying contact backup rolls and asymmetry self-compensating work rolls of the ASR mill, 100% of roll bending force. It can be seen from Fig.10 that the crown control range

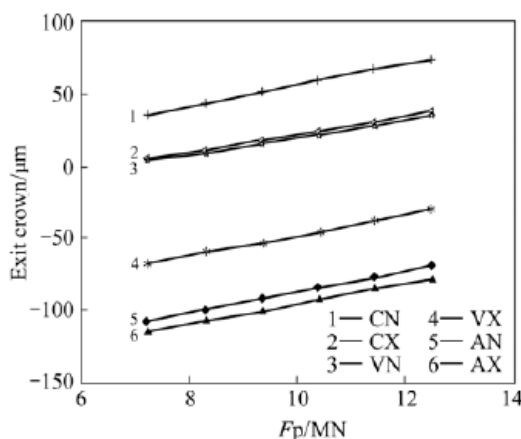


Fig. 10. Effect of typical roll contour configuration of the rolling mill on the characteristic of roll gap profile applied in a 1700mm hot strip mill

by the roll bending force on stand F5 with varying contact backup rolls and conventional work rolls, i.e. the area between the lines VN and VX, is enhanced from $103.5\mu\text{m}$ before application of varying contact backup rolls to $112.0\mu\text{m}$, increased by 8.23% compared with that of the lines CN and CX with conventional backup and work rolls of the K-WRS mill. The crown control range by the roll bending force in the ASR mill, i.e. the area between the lines AN and AX, is $116.7\mu\text{m}$, increasing 12.79% greater than that with conventional backup and work rolls of the K-WRS mill. And at the same time the roll gap stiffness shows an increase from $0.1350\text{ MN}/\mu\text{m}$ with conventional backup and work rolls to $0.1578\text{ MN}/\mu\text{m}$ with VCR and conventional work rolls, and to $0.1691\text{ MN}/\mu\text{m}$ by increasing 25.26% in the ASR mill (Cao, J. G., 2008).

The "as-ground" work roll contour is designed by the operators and provides effects on the strip crown magnitude during the initial phase of rolling campaign. The roll thermal contour is a result of thermal expansion during hot rolling that consists of basic and periodic components. Since the periodic component is localized on the roll surface, the roll thermal contour is usually taken into account the basic component only. Based on the statistical analysis of the data obtained during rolling at the 1700 mm hot strip mill, the roll thermal crown increases rapidly during the first 10 coils in a rolling campaign and amounts to the maximum steady value exceeding $300\mu\text{m}$ after rolling 30 coils (about 60 min) during the steady rolling pace that comprises of the rolling time and the idle or gap time between rolling of coils. If there are some changes during rolling, such as rolling pace, variations in strip width and cooling water temperatures, small but rapidly varying changes in the roll thermal contour occur. Work roll wear effect is another major reason for the variation of the roll contour during rolling of a hot strip mill cycle. As can be seen from the collected data, the work roll wear is generally non-uniform along the entire length of the roll barrel. There are different work roll wear contours for different stands. Generally speaking, work roll wear of electrical steel rolling campaign increases greater than that of other rolling campaign. Fig. 11 shows flatness control efficiency curves of typical roll contours configuration applied in F5 in a 1700 mm hot strip mill, in which the rolling force is 10.4 MN and the bending force is 800 kN per chock. In Fig. 11, BW-G represents conventional backup

and work roll initial ground contours and BW-W represents conventional backup and work roll comprehensive contours including wear contour and thermal expansion after a rolling campaign. VW-G represents VCR and conventional work roll initial ground contours and VW-W represents VCR and work roll comprehensive contours including wear contours and thermal expansion after a rolling campaign. VA-G represents VCR and ASR initial ground contours and VA-W represents VCR and ASR comprehensive contours including wear contours and thermal expansion after a rolling campaign.

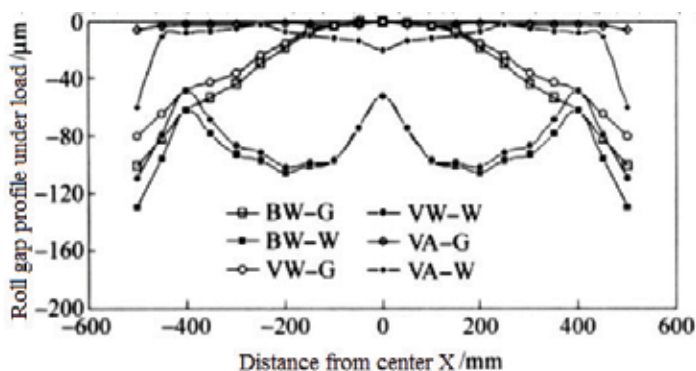


Fig. 11. Flatness control efficiency curves of typical roll contour configuration applied in a 1700mm hot strip mill

It can be seen from Fig. 11 that significant change of roll gap profile with conventional work and backup roll contours in downstream stands of hot rolling lead to non-oriented electrical steel transverse thickness profile difference, i.e. body crown and especially edge drop (Fig.12). It is hard to meet the improvement requirement of industrial users to the quality of

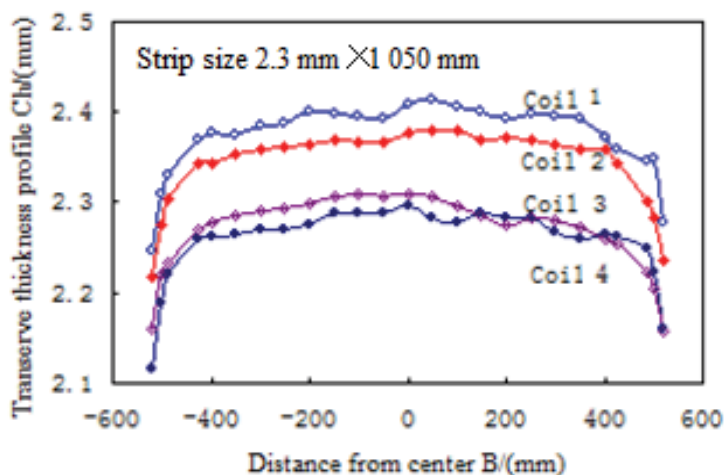


Fig. 12. Measured transverse thickness profile of non-oriented electrical steel strip rolled with long stroke work roll shifting system of conventional work rolls on 1700 mm hot strip mill (2.3 mm×1050 mm)

strip profile and flatness. The rolling period of a newly ground backup roll set in hot strip mills exceeds two weeks, the developed varying contact backup rolls, being geared to different steel kinds and sizes, only go halfway towards solving the problem of significant transverse thickness profile difference by wear contour patterns of conventional work roll contours. A rather smooth local work roll contour near strip edges and an increase in length rolled can be obtained by application of long stroke work roll shifting system with conventional work roll contours that is incapable of the crown control. Having ability of both wear control and overall profile control covering the strip crown, edge drop and high spot in downstream stand and suitable to the schedule free rolling of hot strip mills, the ASR mill with the asymmetry self-compensating work rolls, varying contact backup rolls, corresponding work roll axial shifting strategy and bending force mathematic models have been developed and applied to 1700 mm hot strip mill at WISGCO. According to the rule of work roll wear contour patterns in rolling process, the wear contour of work rolls with one-side tapered contours in the ASR mill can be change from U-type to L-type by shifting the work rolls to the axial direction with a special shifting strategy. As can be seen from Fig. 7, the work roll wear contours of ASR mill have only one side and open another side of U-type wear, and the hot-rolled strips within an entire rolling campaign can be always in the relatively flat area of work roll contours and remove 'from wide to narrow' constraint of coffin rolling regulations by severe U-type work roll wear. This is the on-line automatic compensation process of work roll wear contours in the ASR mills. It can be seen from Fig. 11 that roll gap profile under load is more flat and controllable by applying VCR/ASR technology in the ASR mill, the strip profile, especially edge drop, and flatness control ability of electrical steel can be enhanced notably, the larger transverse thickness profile difference problem of electrical steel sheets caused by severe U-type wear can be solved efficiently and the proportion of lower crown even "dead flat" electrical steel will raised significantly.

The non-uniformity of distribution of roll contact pressure is the specific value of the maximum value and mean value of the pressure along contact length between rolls. The usual rolling force is 10.4MN and the bending force is 800kN for conventional backup and work rolls, VCR/conventional work rolls, and VCR/ASR on 1700 mm hot strip mill. In this condition, The non-uniformities of distribution of roll contact pressure are 1.97, 1.15 and 1.25 in the early stage of work roll. The non-uniformities of distribution of roll contact pressure are 2.35, 1.40 and 1.48, due to roll contour changing caused by wear and thermal expansion. The distribution of pressure between rolls become smooth by application of VCR/conventional work roll and VCR/ASR configuration in the rolling mill. The peaks of pressure between rolls of the above three roll contours configuration are 14.64, 8.28 and 8.75 kN/mm in the early service stage and the peaks of pressure between rolls of the above three roll configuration are 17.68, 10.22 and 10.36kN/mm in the end service stage. Compared with conventional roll configuration, the peaks of contact pressure between rolls the service stage of VCR/ASR from early to late stage decrease 10.23% and 41.40%. The new ground roll contour in the ASR mill has a well self-maintenance and expresses the ability of strip profile and flatness control steadily within the entire rolling campaign. It also has the ability of avoiding roll spalling and enhancing the service performance.

3.4 Industrial application and its effect

The VCR has been tested in stand F3 of 1700 mm HSM at WISCO since May 1997 and then improved in F6, F4, F5 and applied in all 7 stands with 14 backup rolls since September 1999. The application of VCR enhances the strip control ability and stability of finishing stands in hot rolling and improves the quality of strip profile and flatness distinctly: Through strip profile and flatness production data statistics of 95.2 thousand coils before (50.8 thousand coils) and after (44.4 thousand coils) application of the varying contact backup rolls, the ratio of strip over-limit crown decreased from 33.90% to 13.84%, the ratio of strip flatness overall length over 7 I-units decreased from 19.28% to 6.91%. It can be concluded from the industrial tracking test data that the backup roll wear contours of VCR is well-distributed along roll barrel which can enhance the self-maintenance of backup roll profile efficiently, ameliorate the change feature of roll profile remarkably. It can not only maintain the stability of strip control effect, but also avoid backup roll spalling, extend life of service from 2 weeks to 3 weeks or even more 5 weeks for test in upstream stands of 1700 mm HSM, increase rolling yield.

In order to verify the performance of ASR mill type and strip control technology and to improve the quality and quantity of electrical steel strip, the ASR has been applied to 1700 mm hot mill at WISCO for test in 2003, 8. According to applied to stands F6, F4, F5 and so on for test and considering edge crack of electrical steel sheets in hot rolling production practice, the ASR applied to downstream stand in rolling service of electrical steel sheets can control edge drop efficiently, but a more serious edge crack maybe caused by stress concentration. Considering the factors as mentioned above, ultimately, the ASR mill has been industrial tested successfully and applied to downstream stand F5 of 1700mm hot strip mill on 1700 hot strip mill in July 2004 and applied to the industrial production of 2.3 mm×1050mm conventional width electrical steel rolling campaign since November 2004 and applied to wide electrical steel rolling campaign since October 2005. The ASR mill can achieve a decrease in the crown of the roll gap and an increase in the roll gap stiffness for the stability of the roll gap profile controllable during rolling, which can better the affection of work roll severe wear to strip profile and flatness control ability of electrical steel sheets edge drop and make a profound contribution to low crown decreasing non-oriented electrical steel sheets production. Measured transverse thickness profile of non-oriented electrical steel strip rolled in the ASR mill is shown in Fig.13, compared with that of long stroke work roll shifting system of conventional work rolls in Fig.12 on 1700 mm hot

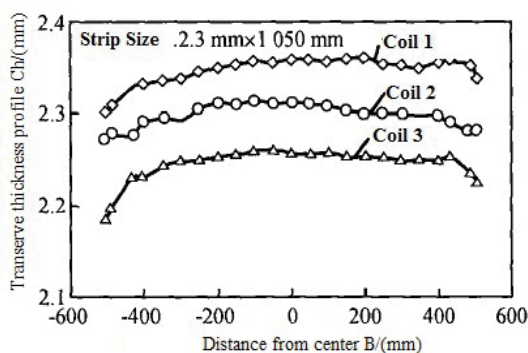


Fig. 13. Measured transverse thickness profile of electrical steel rolled in ASR mill (2.3 mm × 1050 mm)

strip mill(2.3 mm×1050 mm).The industrial results were that the rate of measured crown of the same strip-width 2.3 mm×1 050 mm conventional width electrical steel strips for test less than 52μm was increased from 13.70% to 81.25%, the rate of measured crown less than 45μm is increased to 93.15% by the continuous application, the life of rolls expanded from 60 coils to 80 coils and more, whereas the rate of measured crown of the same width 2.55mm×1200~1300mm wide electrical steel strips for test less than 45μm was increased from 41.8% to 98.2%, the rate was 92.7% for application, the rate of measured wedge less than 22μm was increased from 68.54% to 95.48%, the life of rolls expanded from 40 coils to 70 coils and more.

4. Finite element analysis of shape buckling load for cold-rolled strips

Strip flatness performance obtained during cold reduction is considered to be the main factor in achieving optimum productivity of downstream facilities and final product quality. The distribution of the strip extension along the width direction is usually uneven in tandem cold rolling process. Due to this, the strip after rolled always has internal residual stress distributed unevenly along the direction of strip width, some part of the strip is tensile, and some others is compressive. When the stress of the compression part goes over a certain threshold, strip deformation occurs, kinds of buckle turns up, such as edge wave, center buckle and their compositions. The root causes of strip buckling are that lengths of the longitudinal fibers are different because the strip is unevenly rolled along the width direction and reach the certain threshold.

According to energy principle and variation calculus, using thin elastic small displacement theory, thresholds of strip buckling under every kind of internal stress can be solved without any expression of buckling conditions, numerical method is the effective way to get the value of critical stress under every rolling conditions (Nappep, C., 1997; Fischer, F. D., 2003; Liu,Y.L., 2007). So it's important to ensure the threshold of strip buckling correctly, and this will be the fundamental of establishing a shape control model in cold rolling.

The 1700mm tandem cold rolling mill at the Wuhan Iron and Steel Company (WISCO) was designed and built by SMS and was commissioned in 1978. From 2003 to 2004, the mill completed a series of upgrades on the PL-TCM (pickling line combined with tandem cold rolling mill) etc. supplied by VAI Clecim and SIEMENS. The mill is one of the largest tandem cold rolling mills in China with a capacity of 1.85Mt/y including 500 thousand tons of electrical steel now. The mill still use the previous flatness target model of the automatic shape control system after technical upgradion, which can't reflect actual production in industrial process. With the ANSYS finite element method, a calculated models to analyze and calculate strip buckling was established. With corresponding loads and boundary conditions, the threshold of strip buckling is brought out, which is valuable for the construction of flatness target model in electrical steel rolling.

4.1 Finite element modeling of strip buckling

As the object of study here is wide and thin strip in cold rolling, the thickness of the strip, which is only 0.5 mm, is much less than its width of 1050 mm and more, and length of 1000 km and more, this can be seen as a thin strip buckling problem. So modeling with a two-dimensional element shell63 provided by ANSYS, the value of the strip thickness is imported as a real parameter. The width of the finite element model is the same as that of

the strip, and the length of the model is equal to half wave length L of the buckling strip. The areas of the divided elements is less than $20 \text{ mm} \times 20 \text{ mm}$. The detailed parameters of the model: Width $b=1300 \text{ mm}$. The calculated parameters: Young's modulus $E=2.1 \times 10^{11}$, Poisson's ratio $\mu=0.28$.

Class 1 of Chebyshev polynomial system is adopted to represent the external loads which will be used in the simulation, specific expression are as follows:

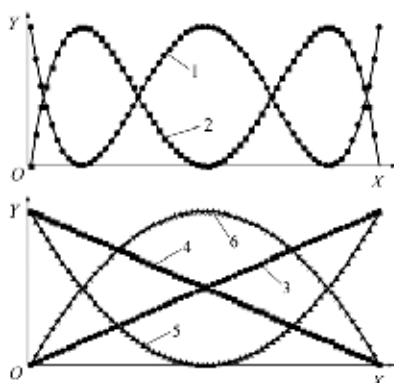
$T_1(x)=x$, $T_2(x)=2x^2-1$, $T_4(x)=8x^4-8x^2-1$. Shape of cold-rolled strip stress is shown in Fig.14.

The external loads is assumed to remain constant in buckling process of strip, i.e. the instability phenomena during stress relaxation is ignored. Different external loads is replaced by different constraint conditions. In addition, try various constraint conditions to each external load. Carrying out a number of calculations to ensure which constraint condition is closest to the actual situation, and use it as the simulation condition. Loads of long center buckle is shown in Fig.15.

4.2. Calculation and analysis

Since the special structure of Cold-rolled wide and thin strip, the buckling has already

happened when $\sigma_{cr} \ll \sigma_s$, so it is unnecessary to take the nonlinear characteristics into account. The main parameters for calculation are strip width, thickness and half wave length L . According to the practical experience and measured data, the half wave length L varies usually between 0.25 and 1.25 times of the width, so the minimum buckling threshold and the length of half wave L can be got by a number of calculations with changes in the parameter range. The method of construction buckling analysis provided by ANSYS is accomplished by increasing the nonlinear load step by step, with which the threshold of the buckling load is brought up. In the calculation, when the initial shape of the strip is flat with a plate load, there would be no deformation or buckling. A reasonable static displacement is needed in order to get predeformation in the strip before calculation, so the model is buckling. Using nonlinear calculation module BUCKLING of ANSYS package, the threshold of buckling will be calculated. Buckling of long center is shown in Fig. 16.



1-1/4 wave, 2- center buckle with edge wave, 3- left side long edge wave, 4-right side long edge wave, 5-center wave, 6-two sides edge wave

Fig. 14. Shape pattern of cold-rolled strip stress

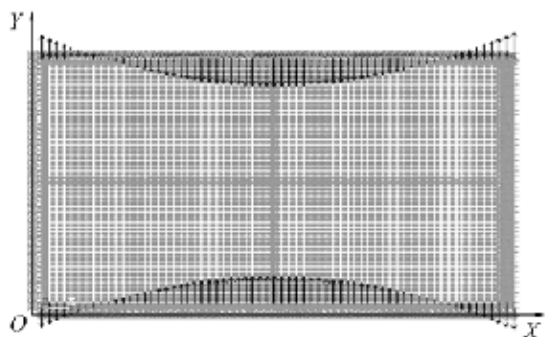


Fig. 15. Loads of long center buckle

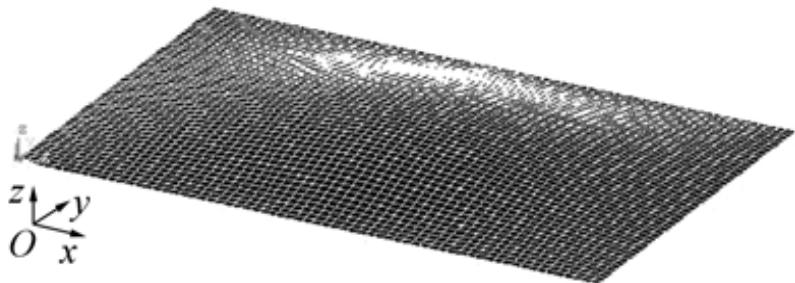


Fig. 16. Buckling of long center

The results of strip critical threshold for typical different waves with strip size 0.5 mm×1 300 mm are shown in Table 2 and Fig.17.

Loads of the long edge wave		Loads of the long center wave		Loads of two sides symmetrical wave		Loads of center buckle with edge wave		Loads of 1/4 symmetrical wave	
L/mm	Threshold	L/mm	Threshold	L/mm	Threshold	L/mm	Threshold	L/mm	Threshold
0.80	0.646 95	0.80	0.881 26	0.60	0.592 74	0.30	2.052 30	0.40	1.512 60
0.90	0.634 86	0.90	0.839 76	0.64	0.590 97	0.40	1.818 00	0.46	1.500 50
1.00	0.626 04	1.00	0.814 43	0.70	0.589 86	0.46	1.761 80	0.50	1.497 40
1.10	0.621 20	1.10	0.801 08	0.72	0.589 83	0.50	1.729 70	0.52	1.497 00
1.16	0.620 10	1.16	0.798 90	0.76	0.590 19	0.52	1.733 80	0.56	1.498 10
1.20	0.620 00	1.20	0.798 51	0.80	0.590 99	0.56	1.746 90	0.60	1.500 70
1.26	0.620 66	1.26	0.800 52	0.90	0.594 29	0.60	1.802 10	0.70	1.508 90
1.30	0.621 52	1.30	0.804 08	1.00	0.598 38	0.70	1.899 10	0.80	1.514 90
						0.80	1.975 90	0.90	1.521 00

Table 2. Buckling load results of different strip waves

As can be seen from Fig. 17, different loads lead to different critical thresholds. The biggest buckling load appears in the center buckle with edge wave, the smallest appears in the side long edge, and the long center ranks the middle. Under the same internal residual stress, the edge wave is most easily generated, while it is the most difficult to produce the center buckle with edge wave.

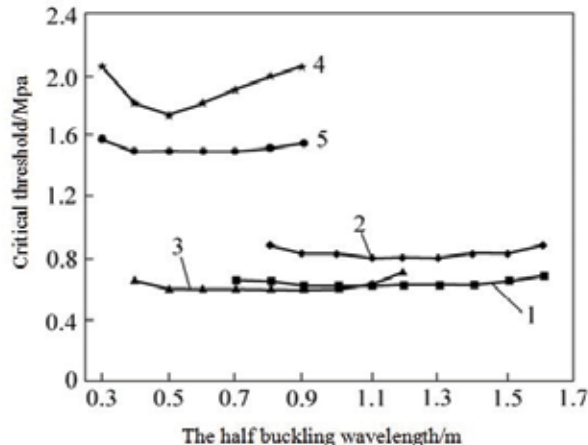


Fig. 17. Buckling load of five typical waves

When the length of edge wave and center wave becomes larger, the half buckling of the center buckle with edge wave is 50%~60% of width, the long center and the single-side long edge is about 80%, the symmetrical dual-side long edge is about 70%. These meet the facts that the center buckle with edge wave is short wave and the long center and the single-side long edge is long wave.

Taking the long edge wave and the center buckle as examples, relationship between buckling load and tensile strain of thin and wide strip is shown in Fig.18. When a tensile strain is applied to the strip, the critical thresholds obviously enlarge with a linear pattern of strain. So a proper tensile stress can prevent the strip from buckling in a cold rolling process.

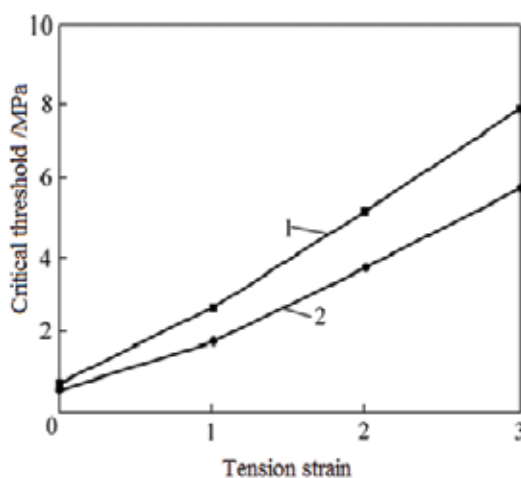
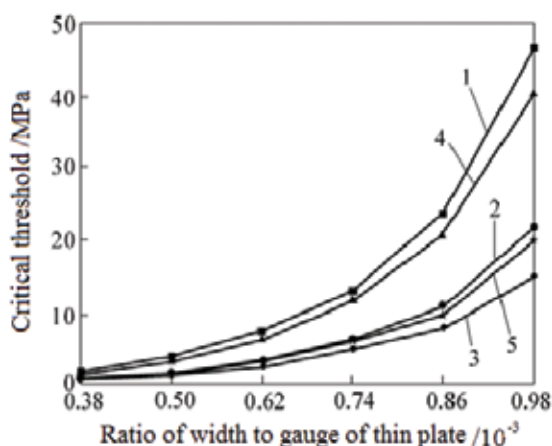


Fig. 18. Relationship between buckling load and tensile strain

The results of critical buckling threshold for 5 typical wave or buckle with different ratio of width to gauge of thin plate are shown in Fig. 19.



1-center buckle with edge wave, 2-center wave, 3-two sides edge wave, 4-1/4 wave, 5-one side long edge wave

Fig.19. Relationship between buckling load and ratio of width to gauge of thin plate

As can be seen from Fig.19, critical buckling thresholds increase significantly with the increase of ratio of width to gauge of thin plate. And with the same increased ration, the critical buckling thresholds of composite waves increase more quickly than that of simple waves such as center buckle and edge wave, i.e. the slope of the curve is greater than the other two wavy's slope. Therefore, the greater the thickness of the plate, the more difficult buckling. The ratio of width to gauge of thin plate is an important factor to the critical buckling threshold of thin and wide strip. In general, the analysis of strip buckling is valuable for the construction of flatness target model in tandem cold rolling mills.

5. Finite element analysis of comprehensive edge drop control technology of strip in 4-high ECC(edge drop & crown compact) tandem cold rolling mills

Edge drop control is a new research field in the strip profile and flatness control because of the increasing requirements for strip profile and flatness (Campas, J.J.,1995; Wang, J.S.,2001; Chi, W.M.,2003). Nowadays the main edge drop control technology are applied by the new built 6-high mill type, such as 6-high UCMW developed by Hitachi with tapered work rolls(K-WRS) (Hiruta, T.,1997; Yarita, I.,1998; Lu, H.T. , 2006; Zhou, X.M., 2007)and 6-high CVC with EDC system developed by SMS Schloemann Siemag AG(Hartung, H.G.,1998; Lackinger, C.V.,2002). The 6-high rolling mill in which the edge drop control effect is improved by shifting the intermediate rolls and the work rolls with an effective edge drop control device along the axial direction of rolls and using the smaller diameter of work rolls. For a tandem cold rolling mill composed of conventional 4-high mills is not provided with an effective edge drop control device, the edge drop of cold-rolled strip is mainly determined by the hot-rolled strip profile as shown in Fig. 20. Even taking various measures such as no trimming before rolling and great cutting-edge after rolling, increasing

the positive bending for center buckle rolling at the expense of flatness control for effects of edge drop, the average edge drop of strip is 15~20 μm . The collected measured data of strip which have the different sizes such as 0.5mm \times 1195mm, 0.5mm \times 1210mm, 0.5mm \times 1240mm and etc. show that the average edge drop is 19 μm , the proportion of the edge drop less than 10 μm is only 29.2%. The 4-high rolling mill in which the edge drop control effect is improved much by shifting the tapered work rolls (K-WRS) can't meet the required quality nowadays, such as the 1700 tandem cold rolling mill upgraded by VAI CLECIM and SIEMENS on March 2004 (Lu, H.T., 2006). Generally, the main edge control technology has applied to the new built 6-high tandem cold rolling mills, such as 6-high UCMW mill at yawata works of Nippon Steel in Japan, No.2 silicon steel works of WISCO and No.3 tandem cold rolling mill of Baosteel in China, Gwangyang plant of POSCO in Korea, 6-high CVC mill with EDC work roll and cooling system at Beeckerwerth cold rolling mill of TKS (Thyssenkrupp steel) in Germany, No.2 tandem cold rolling mill of WISCO and Hangang tandem cold rolling mill of HBIS (Hebei Iron & Steel Group Co. Ltd) in China. Based on ANSYS software and industrial rolling of edge drop control, a 3-dimensional finite element model of roll stacks was built for 1700 mm 4-high tandem cold rolling mill. The developed EDW (edge drop control work rolls) and matched VCR (varying contact backup rolls) by integration design thought of roll contours, mathematical model of roll shifting and bending for edge drop control were integrated to the comprehensive edge drop control technology of 4-high ECC mill type developed by the project team of USTB & WISCO (University of Science and Technology Beijing & Wuhan Iron and Steel Company). The comprehensive edge drop control technology has applied to the production of non-oriented electrical steel strip in 1700mm 4-high tandem cold rolling mill at WIS(G)CO of China since August 2006.

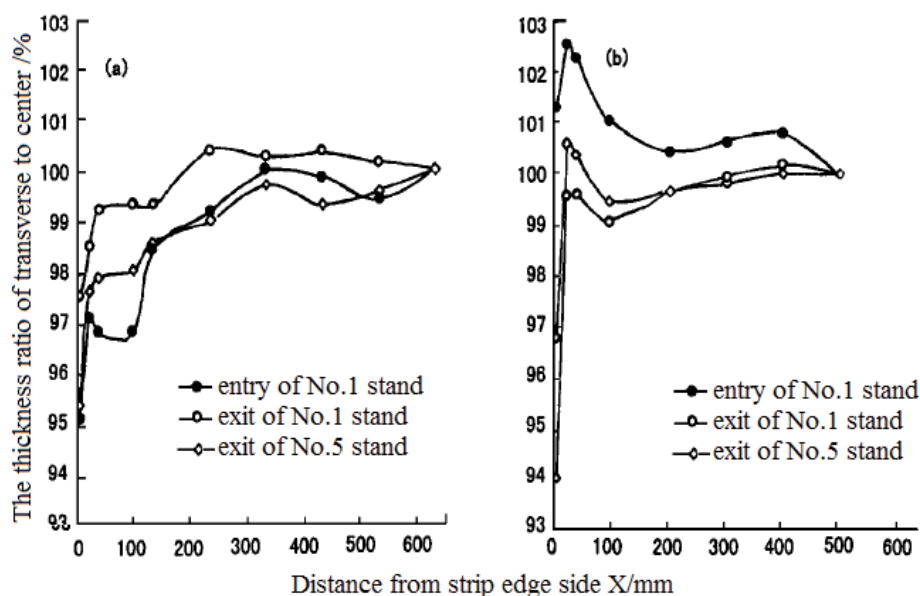


Fig. 20. Edge drop of conventional 4-high tandem cold rolling mills

5.1 FEM for roll stacks and strips

In order to find out the crucial factors which affect the edge drop, an explicit dynamic finite element method was adopted to build a combination simulation model for roll stacks and strip, as shown in Fig. 21. Modeling parameters are shown in Table 3.

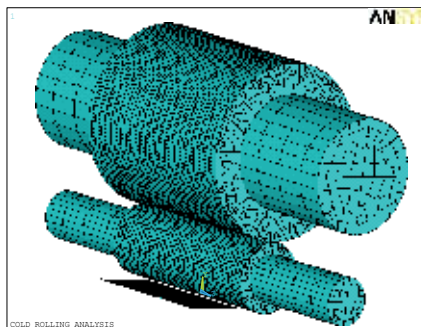


Fig. 21. Finite element model of roll stacks and strip in 1700 tandem cold rolling mill

5.2 The analysis of influencing factor of the edge drop

The main rolling parameters which have a great effect on the strip edge drop are strip width, strip thickness, deformation resistance, friction and roll diameter, etc. Aims at some important parameters such as strip thickness and deformation resistance integrating with the model features, a simulation is finished.

The effect of strip width on the strip edge drop is analyzed by designing three conditions that strip width B is 800mm, 1000mm and 1200mm. The simulation results show that there is a little effect of strip width on the strip edge drop during the process of the cold-rolled strip, the strip edge drop reduces with the increase of strip width, and the rule that the lateral thick difference changes with strip width is invariability after strip thickness reduces, but both of the edge drop are reduced.

Parameter		Value
Measure of the BR body	$D_B \times L_B / mm$	1500×1700
Measure of the BR neck	$D_E \times L_E / mm$	980×520
Measure of the WR body	$D_W \times L_W / mm$	600×1700
Measure of the WR neck	$D_N \times L_N / mm$	406×520
Length of the strip	L / mm	500
Width of the strip	B / mm	800/1000/1200
Thickness of the strip at maunting side	H / mm	1.7/2.3/3.0
The type of the FE		3D Solid164
The quantity of the strip FE		13158

Table 3. Modeling parameters

The analysis results of strip rolling process thought three conditions that entry thickness was 1.7mm, 2.3mm and 3.0mm separately. The result shows that the strip edge drop is reduced with the reduction of strip thickness. The actual production needs to control the

edge drop ,the width range which is 1000mm~1200mm and the edge drop which changes with $8\mu\text{m}$ through simulation calculation. The strip edge drop changes the $36\mu\text{m}$ when strip thickness reduces from 3.0mm to 1.7mm and the strip edge drop changes the $21\mu\text{m}$ when strip thickness reduces from 3.0mm to 2.3mm. Therefore, there is a very little effect of strip width on the strip edge drop and the effect of strip width can be ignored during the control of edge drop.

The effect of deformation resistance on the strip edge drop was analyzed by designing three conditions that deformation resistance was 350MPa, 270MPa and 195MPa ,and the effect of deformation resistance with different width on the strip edge drop was analyzed. The simulation results show that there is a great effect of deformation resistance on the strip edge drop and with the increase of deformation resistance, the strip edge drop increases rapidly. The reason is that the increase of deformation resistance leads to the degree of rolling force heterogeneity more serious and makes the roll flattening more heterogeneous, so the strip edge drop increases rapidly.

5.3 Comprehensive edge drop control technology of strip in 4-high ECC mill

Based on the FEM analysis and industrial experiment, the EDW(Edge drop control work rolls) and matching VCR(Varying contact backup rolls) were designed and proposed. In order to control the strip edge drop contiguously, EDW provide a continuously variable taper to compensate for uneven flatten that can also effectively avoid small edge wave and other flatness problems on strip edge. According to the developed shifting strategy and mathematical models,the EDW improve edge drop effect and are suitable for different widths of strip by shifting top and bottom work rolls of multi-stands to the axial opposite direction in the 4-high ECC(Edge drop & crown compact) mill(Fig.22). Transverse thickness and profile of the sampling strip is shown in Fig.23 before and after applying EDW technology to No.1 stand of 1700 mm tandem cold rolling mill at WISCO since August 2006 .The rate of edge drop less than $10\mu\text{m}$ defined as the difference of thickness of 115 mm and 15 mm from the strip edge side improved from 62.5% of the 4-high tapered work rolls (K-WRS) mills to 91.3% of the 4-high ECC mill with EDW rolls. According to the thoughts mentioned above and the electrical steel production program of PL-TCM(pickling line combined with tandem cold rolling mill), VCR which is compatible with EDW is designed to eliminate harmful contact area between rolls to improve edge drop control effect.

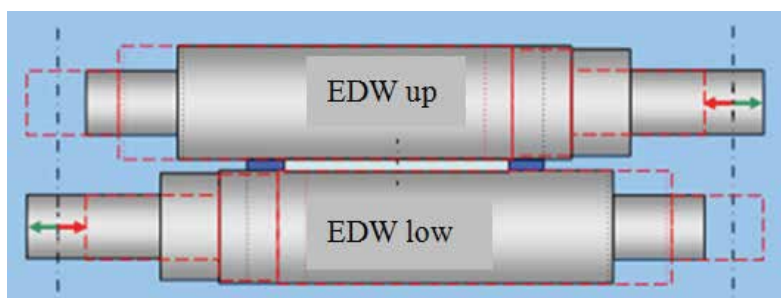


Fig. 22. Principle of EDW(edge drop work rolls) in 1700 tandem cold rolling mill

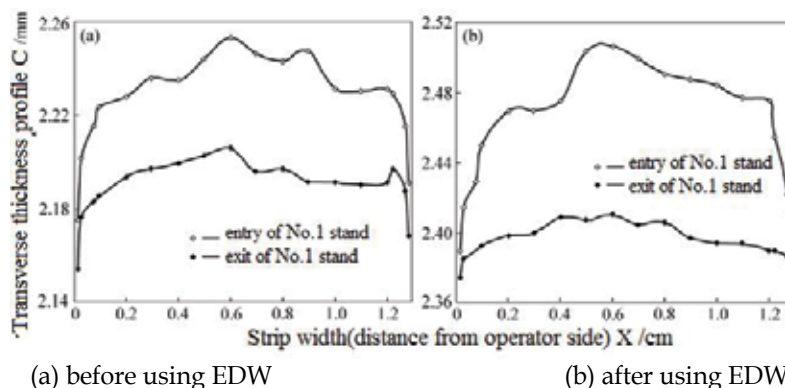
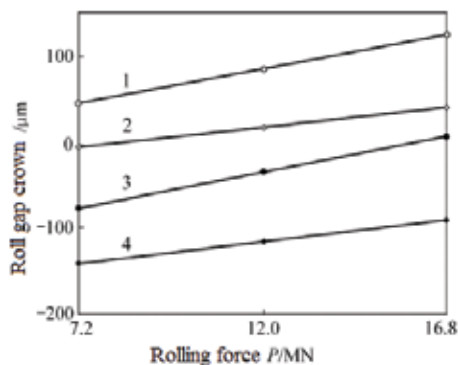


Fig. 23. Comparison of strip profile before and after using of EDW(edge drop work rolls) in 1700 tandem cold rolling mill

A finite element model was established to analyze the quantitative relationship between the length of the contact line between the work and backup rolls and the strip width of PL-TCM under different strip width, rolling force and roll bending conditions. According to the properties of roll system deformation, the contact length between rolls can match the strip width automatically under the action of the rolling force, so the harmful contact area is reduced or eliminated, also the bending deformation of roll gap becomes small, reducing the crown of the roll gap, increasing the roll stiffness, raising the control range and efficiency of bending force and improving the flexibility of the roll gap. Under the same roll shifting stroke, roll gap stiffness of the 4-high ECC mill with EDW compatible with conventional backup rolls or VCR was obtained from simulation (Fig.24). The results show that the roll gap stiffness of the configuration roll system of EDW and VCR is $148\text{kN}/\mu\text{m}$, 19.1% more than conventional back-up rolls $124\text{kN}/\mu\text{m}$, obviously increasing stability of roll gap control. The roll gap crown adjustable zone of EDW and VCR is $133.1\mu\text{m}$, compared with that of conventional back-up rolls $120.2\mu\text{m}$, increases 10.7%. This is not only mean an improvement of crown control, but also an increasing of edge control ability. The contact pressure distribution between rolls of EDW with VCR and conventional backup rolls was obtained from simulation and the pressure peak was 21.6% smaller than the former.



1-bending force is 0 with conventional BR(backup roll); 2- bending force is 0 with VCR; 3-bending force is 0.5 MN with conventional BR; 4- bending force is 0.5 MN with VCR
Fig. 24. Roll gap stiffness of EDW compatible with VCR and conventional back-up rolls

An explicit dynamic finite element method is adopted to build a combination simulation model for roll stacks and strip. The area method of edge drop equivalent is put forward to build the shifting mathematical model of edge drop control based on the mechanism of edge drop control and characteristics of the 4-high ECC mill with EDW and VCR. The parameters of model are determined by simulation and industrial experiment (Cao, J. G., 2008).

5.4 Industrial experiment and application

WISCO of China recently established a 4-high tandem cold mill rolling technology that assure excellent profile and flatness by installing ECC(edge drop & crown compact) mill including the comprehensive edge drop technology mentioned above since August 2006. In July 2007, the data which come from the continuous rolling of the total weight of 23266 tons of electric steel with different sizes such as 0.5mm×1150mm, 0.5mm×1195mm, 0.5mm×1240mm show that the proportion of the edge drop less than 10 μm is 100%. According to the user's growing severe demands, the proportion of the edge drop less than 7 μm is 98.22% and the proportion of the transverse thickness deviation along the strip length of the same strip less than 10 μm is 97.25%, this significant performance reaching to the same high precise control level of new-built 6-high 1450UCMW tandem cold rolling mill during the same rolling period. The proportion of the transverse thickness deviation along the strip length of the same strip less than 10 μm is 92.06% by applying this technology on 1700 mm tandem cold rolling mill for the production of 114 800 t non-oriented electrical steel on first half of 2008.

6. Conclusion

Finite element method is an effective way to the behavior analysis of mechanical, thermal, deformation and other characters of strip and rolling mills. With the help of theory research, simulation and industrial experiment, the developed hot strip mill and tandem cold rolling mill which provide both wear control and profile control ability, comprehensive edge drop control ability are applied to the industrial production and gain better strip profile and flatness quality, process improvements and productivity increase.

7. References

- Ginzburg, V. B. (1993). *High-quality Steel Rolling: Theory and Practice*. Marcel Dekker Inc, ISSN 0-8247-8967-9, New York
- Chen, X.L. (1997). Flatness control in new generation high-tech mills for wide strip rolling, *Journal of University of Science and Technology Beijing*, Vol.19, No. S1, 1-5, ISSN 1001-0521
- Cao, J. G., Zhang, J., Song, P., et al. (2006). ASR technology for controlling profile and flatness of non-oriented electrical steel strip in hot rolling. *Iron and Steel*, Vol. 41, No. 6: 43-46, ISSN 0449-749X
- Li, H.C., Jiang, Z.Y., Tieu, A.K., et al. (2007). Analysis of premature failure of work rolls in a cold strip plant. *Wear*, No. 263: 1442-1446, ISSN 0043-1648

- Ray, A.K., Mishra, K.K., Das, G., et al. (2000). Life of rolls in a cold rolling mill in a steel plant-operation versus manufacture. *Engineering Failure Analysis*, No. 7: 55-67. ISSN 1350-6307
- Chen, S.G., Wang J.Z., Kang Y.L., et al. (2006). Barrel spalling analysis of 45Cr4NiMoV forged steel backup roll for hot rolling. *Shandong Metallurgy*, Vol. 28, No. 3: 51-53. ISSN 1004-4620
- Cao, J. G., Chen, X.L., Yang Q., et.al. (1999). Preventing method and its effect for roll spalling in hot wide strip mill. *Metallurgical Equipment*, No. 4: 6-9,45. ISSN 1001-1269
- Yang, G.H., Cao, J.G., Zhang J., et.al. (2008). Backup roll contour of a SmartCrown tandem cold rolling mill. *International Journal of Minerals, Metallurgy and Materials*, Vol.15, No.3: 357-361. ISSN 1005-8850
- Chen, X.L., Yang Q., Zhang Q. D., et.al.(1994/1995). Varying contact backup roll for improved strip flatness. *Steel Technology International*, (yearly): 174-178. ISSN: 0953-2412
- Chen, X.L., Zhang, J., Zhang, Q.D., et.al. (2000). Development in profile and flatness control system of hot strip mills. *Iron and Steel*, Vol.35, No.7: 28-33. ISSN 0449-749X
- Cao, J.G., Zhang J., Chen X.L., et al.(2002). Control of roll contour for strip profile and flatness in hot rolling, *Proceedings of the 44th Mechanical Working and Steel Processing Conference and the 8th International Rolling Conference*, pp.1001-1010, Florida, Sept. 2002, The Iron & Steel Society, Orlando
- Cao, J. G., Zhang, J., Chen, X.L., et al. (2005). Selection of strip mill configuration and shape control. *Iron and Steel*, Vol.40, No.6: 43-46. ISSN 0449-749X
- Ou, Y.Y., Cao, J. G., Zhang J., et al. (2005). Relation between profile and flatness control and rolling trait of the hot strip mill for non-oriented silicon steel. *Metallurgical Equipment*, No.1: 6-9. ISSN 1001-1269.
- Cao, J. G., Zhang, J., Gan, J.B., et al. (2006). Work roll wear prediction model of non-oriented electrical steel sheets in hot strip mills. *Journal of University of Science and Technology Beijing*, Vol.28, No. 3: 286-290, ISSN 1001-0521.
- Cao, J. G., Wei, G.C., Zhang, J., et al. (2008). VCR and ASR technology for profile and flatness control in hot strip mills. *Journal of Central South University : Science & Technology of Mining and Metallurgy*, Vol.15, No. 3: 264-270, ISSN 1005-9784.
- Nappez, C., Boulot, S., McDermott, R.C.(1997). Control of strip flatness in cold rolling: a global approach. *Iron and Steel Engineer*, Vol. 74, No. 4: 42-45, ISSN 0021-1559
- Fischer, F. D., Rammerstorfer, F.G. & Friedl, N.(2003). Residual stress-induced center wave buckling of rolled strip metal. *Journal of Applied Mechanics*, Vol.70, No.1: 84-90, ISSN 0021-8936
- Liu, Y.L., Fan, J., Mike, L., et al. (2007). Three-dimensional simulation of local shape defects in continuous cold rolling. *Iron and Steel Technology*, Vol.4, No.8: 70-80, ISSN 1547-0423
- Campas, J.J., Terreaux, S., Roches, L.V D, et al.(1995). New on-line gage for edge drop measurement and effect of tapered work rolls. *Iron and Steel Engineer*, Vol.72, No.12:27-32, ISSN 0021-1559
- Wang, J.S., Zhao, Q.L., Jiao, Z.J., et al.(2001). The research on the principle and application for T-WRS&C cold mill. *Heavy Machinery*, No.6:8-11, ISSN 1001-196X.
- Chi, W.M., Gao, W., Wu, B., et al. (2003). Cold strip rolling of new technology-edge drop control. *Mining and Metallurgy*, Vol.12, No.2:91-94. ISSN 1005-7854.

- Hiruta, T., Akagi, I., Mizushima, N. (1997). Development of advanced transverse thickness profile control of thin hard steel strips at tandem cold rolling mill. *Kawasaki Steel Technical Report*, Vol.37, No.10:19-24, ISSN 0388-9475
- Yarita, I., Kitahama, M., Hiruta, T., et al. (1998). Transverse thickness profile control in hot and cold strip rolling by tapered-crown work roll shifting (K-WRS) mill. *SEAIQ Quarterly* (South East Asia Iron and Steel Institute), Vol.27, No.3:26-34, ISSN: 0129-5721
- Lu, H.T. , Cao, J.G. , Zhang, J. , et al. (2006).Edge drop control of a taper roll during continuous cold rolling. *Journal of University of Science and Technology Beijing*, Vol.28, No. 8: 774-777, ISSN 1001-0521.
- Zhou, X.M., Zhang, Q.D.,Wang,C.S., et al. (2007). Automative edge drop control system of UCMW cold mill and its optimization. *Iron and Steel*, Vol.42, No.9: 56-59. ISSN 0449-749X
- Hartung, H.G., Hollmann, F.W., Holz, R., et al. (1998). A new way to reduce the edge drop. *Steel research*, Vol.69, No.4-5: 143-147, ISSN: 0177-4832
- Lackinger, C.V.,Nettelbeck, H.J.,Oemkes, H.(2002). The new tandem mill at ThyssenKrupp Stahl's Beeckerwerth cold strip. *Stahl und Eisen*, Vol.122, No.2: 25-32, ISSN 0340-4803
- Cao, J.G. , Qi, J.B., Zhang, J., et al. (2008).Backup roll contour for edge drop control technology in tandem cold rolling mill. *Journal of Central South University(Science and Technology)*, Vol.39, No. 5: 1011-1016, ISSN 1672-7207.
- Cao, J.G., Mao, N., Zhang, J., et al. (2008). Study on shifting mathematical model of edge drop control in tandem cold rolling mill. *Iron and Steel*, Vol.43, No.8: 57-60. ISSN 0449-749X.

Strain variations on rolling condition in accumulative roll-bonding by finite element analysis

Tadanobu INOUE

*National Institute for Materials Science,
Japan*

1. Introduction

Bulk ultrafine-grained (UFG) materials with grain sizes of tens to hundreds of nanometers showing improved mechanical properties without the addition of alloying elements have attracted the attention of researchers in materials science (Inoue et al., 2010a; Kimura et al., 2008), and microstructural evolution and hardness variation in the UFG materials fabricated by a plastic deformation process such as equal-channel angular pressing (ECAP) (Horita et al., 2000; Segal, 1995), accumulative roll bonding (ARB) (Saito et al., 1998), caliber rolling (Inoue et al., 2007b; 2009c; Mukai et al., 2010), and high-pressure torsion (HPT) shown in Fig. 1, have been studied in detail (Inoue et al., 2009a; 2010b; Todaka et al., 2008). Since the microstructural evolution of plastically deformed materials is directly related to the magnitude of plastic deformation, the understanding of the phenomenon associated with the strain development is very important (Inoue et al., 2001; 2007a; 2008). For example, in ARB which is a severe plastic deformation process for realizing UFG microstructures in metals and alloys, the microstructure and texture in a sheet processed by one ARB cycle without lubricant dramatically change depending on the thickness location of the sheet (Kamikawa et al., 2007). In a rolling process, including ARB, it is reported that these changes are caused by the redundant shear strain imposed by large friction between rolls and sheet.

Rolling is an excellent plastic deformation process in the mass production of a metallic sheet, and many reports have been published regarding the rolling characteristics, shape control, and microstructure control through theory, numerical simulations, and many experiments. Figure 2 shows a typical illustration of a rolling process. The classical rolling theory (Underwood, 1952) has been used as a method to qualitatively understand variations of the rolling characteristics of a rolling force and torque against the processing parameters such as the roll diameter, reduction, rolling speed, and friction condition. On the other hand, deformation in sheet metals by rolling has been studied in detail through many experiments and finite element simulations. Flow of metals in rolling experiments makes the presence of the shear deformation clear and qualitative relation between equivalent strain including shear deformation and microstructure through sheet thickness has been reported in the literature (Lee et al., 2002; Sakai et al., 1988; Matsuoka et al., 1997). However, only a few

studies have been carried out on quantitative correlation between microstructure and strain through a combination of experiments and finite element simulations (Um et al., 2000; Mukhopadhyay et al., 2007). The embedded-pin method (see Fig. 3) is often employed to measure the shear strain through thickness experimentally (Cui & Ohori, 2000; Hashimoto et al., 1998), but magnitude of shear strain and equivalent strain obtained by this method do not exhibit the exact value (Inoue & Tsuji, 2009b). Therefore, for controlling the microstructures, it is essential to understand the deformation behavior in the sheet accurately and quantitatively through a combination of rolling experiments and finite element simulations.

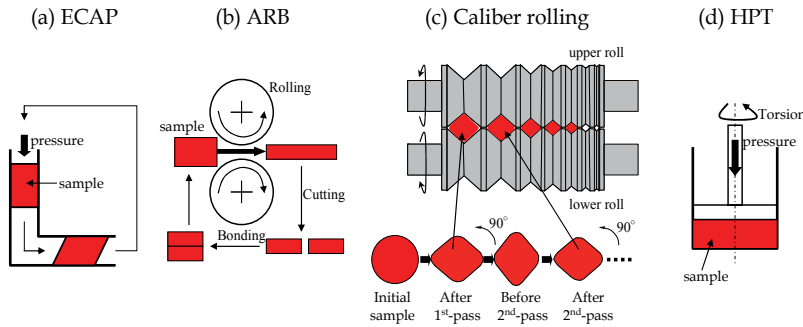


Fig. 1. Major severe plastic deformation processes.

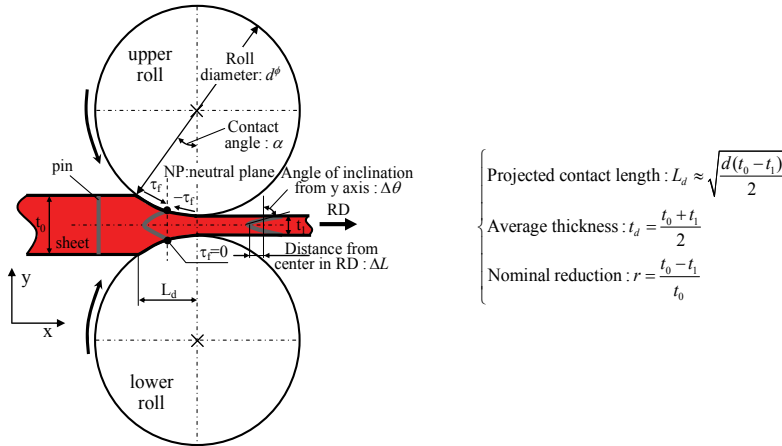


Fig. 2. Schematic illustration showing the geometry of the rolling process.

The shear strain is caused by not only the friction between the rolls and the material surface but also the roll bite geometry, L_d/t_d (Backofen, 1972; Dieter, 1988). Here, L_d denotes the projected length of the contact arc to the horizontal plane, and t_d is the average sheet thickness shown in Fig. 2. In interstitial free (IF) steel sheets rolled by 50% with and without lubrication, Um et al., 2000 investigated the variations of the strain distribution and the flection of an embedded pin against the roll bite geometry by changing the initial sheet thickness and roll diameter. However, the distribution of the shear strain is not shown and

the effect of the roll bite geometry on strain distribution in sheet rolling is not systematically studied. Furthermore, it is not verified that the roll bite geometry is a universal parameter on the strain distribution under unlubricated condition. Hence, it is important to systematically explore the effect of the L_d/t_d ratio on the magnitude and distribution of strains under various friction conditions using numerical simulations. Moreover, if the magnitude and distribution of strains through the thickness in a rolled sheet can be quantitatively estimated by using experimental data (Kamikawa et al., 2007; Sakai et al., 1988), ΔL and $\Delta\theta$, measured from the embedded-pin method, as shown in Fig. 3, a simulation result would provide useful guidelines for analyzing the evolution of microstructures in the ARB process as well as designing the microstructure of the sheet metal by conventional rolling.

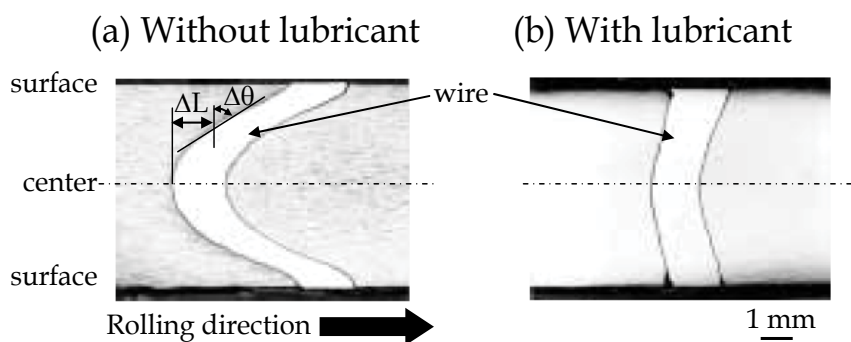


Fig. 3. Flection of stainless wire in sheet after rolling observed by the embedded pin method (Hashimoto et al., 1998). Here, interstitial free steel of 10 mm in initial thickness was rolled by a reduction of 40% per pass at 973 K using a two-high mill with a roll diameter of 300 mm at a rolling speed of 300 mm s⁻¹. The L_d/t_d indicates 3.1 under this rolling condition.

This study aims to exhibit a quantitative correlation between strains and L_d/t_d in metal sheet rolled under various friction coefficients using finite element analysis (FEA), which is a powerful tool for understanding deformation behaviors during a plastic deformation process. The each strain component and equivalent strain at various thickness locations in the sheet during and after rolling were studied in detail, including ΔL and $\Delta\theta$ measured from the embedded-pin method in rolling experiments. Furthermore, the problem associated with universality of the L_d/t_d parameter on the magnitude and distribution of strains is discussed.

2. Finite element model

The elastic-plastic FE simulation was carried out using the FE-code ABAQUS/Explicit ver.6.5-4. A 4-node linear element in a plane strain model, element type CPE4R, was used for sheets of 2 mm and 5.3 mm in initial thickness, t_0 , and the rolls were regarded as the rigid body. No remeshing was carried out in the analysis because the deformed mesh by rolling corresponds to the flection of the pin in the embedded-pin method. The Coulomb condition was used as the frictional condition between the rolls and the sheet, $\tau_f = \mu p$, where

τ_f denotes the shear stress, μ is the friction coefficient, and p is the contact pressure. Assuming the Coulomb law, a condition to pull a sheet into rolls bite is given by $\mu > \tan \alpha$, where α denotes the contact angle shown in Fig. 2.

The flection of the pin in the embedded-pin method makes the presence of the shear deformation clear. The pin flection in pure aluminum (Lee et al., 2002), interstitial free (IF) steel (Matsuoka et al., 1997; Sakai et al., 1988; Um et al., 2000; Kamikawa et al., 2007), and Type 304 stainless steel (Zhang et al., 1996; Sakai & Saito, 1999) sheets rolled at various temperatures has been observed using this method. In the present study, the condition of a commercial 1100 Al sheet rolled at ambient temperature without lubricant and without front and back tensions reported by Lee et al., 2002 was referred to as the main rolling condition: initial thickness, $t_0=2$ mm; nominal reduction per pass, $r=50\%$; roll diameter, $d^\phi=255$ mm; and rolling speed, 170 mm s^{-1} . The L_d/t_d indicates 7.5 under this rolling condition, and its value becomes smaller with decreasing d^ϕ . Nine rolling conditions, Numbers 1-9, employed in the present analysis are listed in Table 1. Numbers 1-6 denote the rolling conditions in which the 1100 Al of $t_0=2$ mm is rolled by $r=50\%$ using a rolling simulator with $d^\phi=310, 255, 201, 118, 40$, and 15 mm, resulting in $L_d/t_d=8.3, 7.5, 6.7, 5.1, 3.0$, and 1.8 , respectively. In order to study the universality of the L_d/t_d parameter on the strains through sheet thickness, Nos. 7-9 at $L_d/t_d = 5.1$ are also analyzed under various combinations of t_0 , r , and d^ϕ . The stress σ –strain ε relationships of the 1100 Al at 301 K employed in the analysis were described by $\sigma=28+105.67 \varepsilon^{0.32} \dot{\varepsilon}^{0.017}$ MPa depending on the strain rate, $\dot{\varepsilon}$ (Ataka, 2006), and a Young's modulus of 70 GPa and a Poisson's ratio of 0.35 were used as the elastic modulus.

No.	Initial sheet thickness t_0 (mm)	Exit thickness t_1 (mm)	Nominal reduction r (%)	Roll diameter d^ϕ (mm)	Projected length L_d (mm)	Average sheet thickness t_d (mm)	Roll bite geometry L_d/t_d	Rolling speed v (mm/s)	Minimum friction coefficient $\mu_{\min}=\tan\alpha$
1	2	1	50	310	12.4	1.500	8.3	170	0.08
2				255	11.3		7.5		0.09
3				201	10.0		6.7		0.10
4				118	7.7		5.1		0.13
5				40	4.4		3.0		0.23
6				15	2.7		1.8		0.38
7	5.3	1.41	29.5	255	8.7	1.705	5.1	170	0.07
8		1.76	12	768	9.6	1.880			0.03
9		2.65	50	310	20.2	3.975			0.13

Table 1. Rolling conditions used in the present study.

In the analysis, the classical metal plasticity models with Mises yield surface, *PLASTIC, HARDENING=ISOTROPIC as keyword in ABAQUS/Explicit, 2006, were employed. The equivalent strain, ε_{eq} , imposed by rolling is defined as follows:

$$\varepsilon_{eq} = \int_0^{t(\text{steady})} \frac{d\varepsilon_{eq}}{dt} dt \quad (1)$$

where $d\varepsilon_{eq}/dt$ denotes the incremental equivalent strain, and $t(\text{steady})$ is the rolling time. Since an incremental strain in the x direction, $d\varepsilon_{xx}/dt$, is equal to a minus incremental strain in the y direction, $-d\varepsilon_{yy}/dt$, under the plane strain condition, the incremental equivalent strain, i.e., the equivalent strain rate, $d\varepsilon_{eq}/dt$, is represented as below:

$$\frac{d\varepsilon_{eq}}{dt} = \frac{2}{\sqrt{3}} \sqrt{\left(\frac{d\varepsilon_{xx}}{dt}\right)^2 + \frac{1}{4}\left(\frac{d\gamma_{xy}}{dt}\right)^2} \quad (2)$$

In equation (2), $d\gamma_{xy}/dt$ is the incremental shear strain. Since the direction of the shear stress, τ_t , in a roll bite changes to opposite directions before and after a neutral plane NP, shown in Fig. 2, the total shear strain, γ , must be expressed as follows:

$$\gamma = \int_0^{t(NP)} \frac{d\gamma_{xy}}{dt} dt + \int_{t(NP)}^{t(steady)} \left| \frac{d\gamma_{xy}}{dt} \right| dt \quad (3)$$

Here, the first term in the above equation denotes a positive shear strain, $\gamma+$, induced by the shear stress, τ_t , before NP, and the second term is the negative shear strain, $\gamma-$, by the τ_t after NP. In other words, the γ represents the total magnitude of the shear strain γ_{xy} taking into account the deformation history during rolling. At the thickness center with no shear deformation, the $d\gamma_{xy}/dt$ is always zero, and the $d\varepsilon_{xx}/dt$ is constant throughout rolling. Hence, ε_{eq} at the center agrees with the value $2/\sqrt{3} \ln\{1/(1-r)\}$ calculated simply from a reduction in thickness independent of the deformation history, where $\varepsilon_{xx} = -\varepsilon_{yy} = \ln\{1/(1-r)\}$ and $\gamma = 0$.

3. Simulation results

3.1 Mesh dependence of strain in rolled sheet

At first, the appropriate mesh division in the FEA was examined because the magnitude of strain depends certainly on mesh size. Figure 4 shows the variation of the equivalent strain, ε_{eq} , at a surface against the initial element length in the thickness direction, t_{el} , for FEA using $\mu=0.3$ under the rolling condition No. 2 in Table 1, where the initial element length in the rolling direction, L_{el} , is 0.0231 mm. The figure inset describes the FE mesh near the sheet surface. It is found that ε_{eq} at the surface depends strongly on the t_{el} as expected. The ε_{eq} increases with decreasing t_{el} and tends to be almost constant at t_{el} below 0.03 mm. On the other hand, ε_{eq} at the thickness center exhibited about 0.80 regardless of t_{el} below 0.26 mm. This magnitude corresponds to a value, $2/\sqrt{3} \ln\{1/(1-r)\}$, of equivalent strain calculated theoretically under the plane strain condition, where the reduction r is 0.5 in the present condition. In the present analysis, the sheet thickness was divided into 66 elements, i.e., $t_{el}=0.03$ mm was used throughout the sheet. Moreover, the sheet length was determined through some simulations on a steady-state of deformation where the strains remains constant along the RD. As a result, the finite element mesh in the sheet with dimensions of 2 mm^t×15 mm^L included 20167 nodes and 19800 elements as illustrated in Fig. 5. Here, a minimum L_{el} is 0.0231 mm at the mid-length (center element), and L_{el} gradually increases toward the front and back from the center. The L_{el} at the front and back edges of the sheet is the same 0.0923 mm. Similar, for the rolling condition No. 9 of $t_0=5.3$ mm in Table 1, $t_{el}=0.03$ mm was used throughout the sheet. Thus, the finite element mesh in the sheet with dimensions of 5.3 mm^t×15 mm^L included 26488 nodes and 26100 elements.

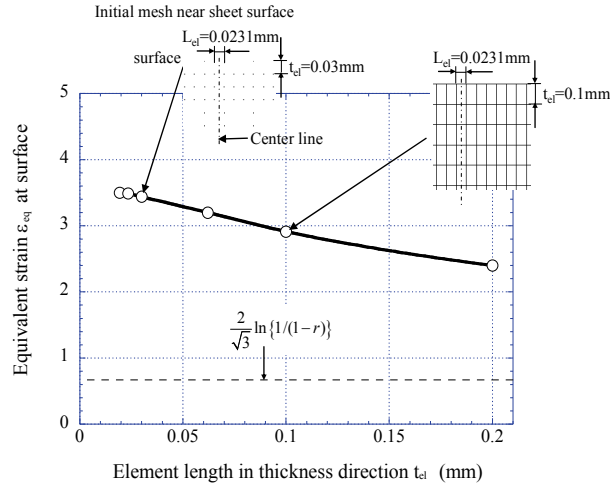


Fig. 4. Variation of ε_{eq} at surface in the rolled Al with $\mu=0.3$ against initial element length in thickness direction. Here, rolling condition No.2 in Table 2 was used.

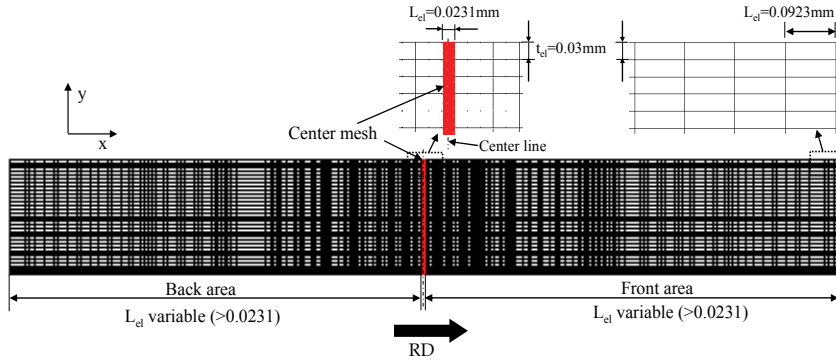


Fig. 5. Initial finite element mesh used in the analysis. Here, L_{el} denotes the mesh length in rolling direction (RD).

3.2 Variations of ΔL_s , $\Delta\theta_s$, and ε_{eq} at the sheet surface with roll bite geometry

Figure 6 shows the effect of the L_d/t_d ratio on the distance from center to surface in RD, ΔL_s , the angle of inclination from the y axis, $\Delta\theta_s$, and the equivalent stain, ε_{eq} , at the surface of a 50% rolled Al sheet in the friction coefficient range of $\tan \alpha \leq \mu \leq 0.4$. The slip areas in Fig. 6 denote a condition of $\mu < \tan \alpha$ on the basis of the Coulomb law, and the open symbol in Fig. 6(a) represents the observed result, $\Delta L_s = 1.1$, of the pin method reported by Lee et al., 2002. Although the ΔL_s have no large difference by L_d/t_d under a low friction of $\mu \leq 0.2$, this difference becomes larger with increasing L_d/t_d under high friction. Similarly, in Fig. 6(b), the $\Delta\theta_s$ becomes larger with increasing L_d/t_d and μ . However, under $L_d/t_d = 8.3$ of the largest roll bite geometry in the present study, the $\Delta\theta_s$ decreases from $\mu = 0.35$ to 0.4.

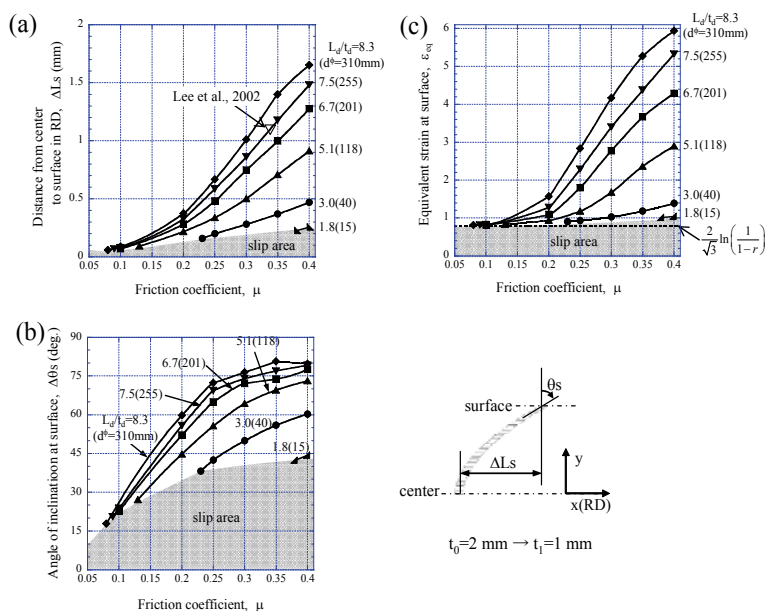


Fig. 6. (a) ΔL_s , (b) $\Delta\theta_s$, and (c) ϵ_{eq} at the surface of a 50% rolled Al sheet against μ under different roll bite geometries L_d/t_d (Nos.1-6 shown in Table 1).

Furthermore, the $\Delta\theta_s$ indicates a large value of 45° at $\mu = 0.4$ even if the roll bite geometry is very small, $L_d/t_d = 1.8$. The ϵ_{eq} increases with increasing L_d/t_d under the same μ . Although the tendency is slight under $\mu \leq 0.2$, it becomes remarkable under $\mu > 0.2$ at $L_d/t_d \geq 5.1$. At $\mu = 0.4$ of a high friction condition, the ϵ_{eq} exhibits a very large value of 6.0, which is seven times higher than that at the center, $2/\sqrt{3} \ln[1/(1-r)]$, under $L_d/t_d = 8.3$. On the other hand, under $L_d/t_d = 1.8$ the ϵ_{eq} is 1.0, which is slightly higher than that at the center. The variation of ϵ_{eq} at the sheet surface as functions of μ and L_d/t_d is very similar to that of ΔL_s . Using Fig. 6(a,c), it is considered that the friction coefficient was $\mu \sim 0.34$ in the rolling experiment under $L_d/t_d = 7.5$ reported by Lee et al., 2002 and the ϵ_{eq} of 4.23, which is five times higher than that at the center, was introduced to the sheet surface. Kamikawa et al., 2007 reported that the microstructural parameters, fraction of high-angle grain boundaries and average misorientation, at the surface in the IF steel processed by one ARB cycle at 773K without lubricant correspond to those at the center in the IF steel processed by five ARB cycles with lubricant. The FE result shown in Fig. 6 is consistent with their microstructural results, although a material is different. The results of Fig. 6(c) indicate that the ϵ_{eq} introduced to the sheet surface by rolling under the unlubricated condition is significantly different by the roll diameter even if the initial thickness, nominal reduction, and friction condition are the same. It is clear that ϵ_{eq} at the surface depends on not only the friction condition between the rolls and the sheet but also the roll bite geometry. It is evident from Fig. 6(b,c) that the ϵ_{eq} and $\Delta\theta_s$ at the surface after rolling with $\mu = 0.4$ under $L_d/t_d = 5.1$ agree with those with $\mu = 0.25$ under $L_d/t_d = 8.3$. However, in Fig. 6(a) the value of ΔL_s does not agree at these conditions. Furthermore, although the $\Delta\theta_s$ for $\mu = 0.4$ under $L_d/t_d \geq 6.7$ indicates almost the same value,

the ΔL s and ε_{eq} are different under these conditions. These differences are attributed to the complicated deformation history near the surface, as described in Subsections 3.3.3 below.

3.3 Deformation behavior at various thickness locations in a rolled sheet

3.3.1 Flection of pin after rolling

Figure 7 shows deformed meshes of a center part of the FE mesh (center mesh in Fig. 5) after rolling with different μ , corresponding to the pin described in Fig. 2. Here, the rolling condition No.2 in Table 1 was used. The flection through the thickness becomes larger with increasing μ . This feature is consistent with the experimental results in which the flection of the embedded pin is larger in a sheet rolled without lubricant than that with lubricant, as shown in Fig. 3. Although the shear strain and equivalent strain are always calculated from the inclination of the pin, based on equations proposed by Sakai et al., 1988, the embedded-pin method does not exhibit the exact value of strains, because these equations are derived under three assumptions; I) The ratio of incremental shear strain to incremental compressive strain is constant during rolling; II) The incremental compressive strain is uniform through the thickness; III) Plane strain conditions prevail in the deformation zone. And, the effect of reverse shear deformation after the neutral plane, NP, shown in Fig. 2, is not considered in these assumptions. As a result, it is clarified from strain histories during rolling shown in Fig. 9 below that two assumptions I) and II) are not proper.

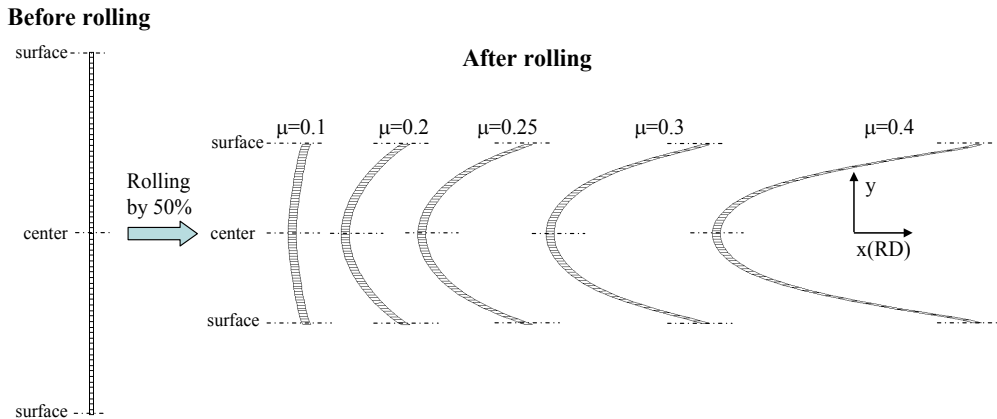


Fig. 7. Initial mesh and deformed mesh after rolling with various friction coefficients μ under rolling condition No.2 in Table 1.

3.3.2 Variations of ΔL and $\Delta\theta$ through sheet thickness after rolling

Figure 8 shows the distributions of the distance from the center in RD, ΔL , and the angle of inclination from the y axis, $\Delta\theta$, through the sheet thickness after rolling with various μ under $L_d/t_d = 7.5$ ($d^\phi = 255$ mm) and $L_d/t_d = 5.1$ ($d^\phi = 118$ mm), respectively. Here, these data were obtained from the values of the integration point for each element divided into 33 through the sheet thickness. In Fig. 8(a), the ΔL increases with increasing μ throughout thickness locations, becomes gradually larger toward the surface from the center, and takes a maximum at the surface. For the same μ , the ΔL under $L_d/t_d = 7.5$ is larger than that under $L_d/t_d = 5.1$ throughout thickness locations. In Fig. 8(b), the $\Delta\theta$ similarly increases with

increasing μ , but the distributions of $\Delta\theta$ through thickness are different by μ . Under $L_d/t_d = 7.5$, the $\Delta\theta$ for $\mu = 0.1$ slightly increases from $y=0$ (center) to 0.45 and there is a steep increase toward the surface thereafter. The $\Delta\theta$ for $\mu = 0.2$ monotonously increases toward the surface. The $\Delta\theta$ for $\mu \geq 0.25$ increases with the distance y , takes a maximum at a thickness location near the surface, and decreases toward the surface from the location. The feature can be seen clearly in the flexion of pin embedded in IF-steel sheets rolled at 973 K under the unlubricated condition, as shown in Fig. 3 (Matsuoka et al., 1997; Hashimoto et al., 1998; Kawabe et al., 1996). The location where the $\Delta\theta$ takes a maximum is slightly away from the surface with increasing μ . The distributions of $\Delta\theta$ for $\mu = 0.2$ and 0.4 under $L_d/t_d = 5.1$ are similar to those for $\mu = 0.1$ and $\mu \geq 0.25$ under $L_d/t_d = 7.5$. Furthermore, the magnitude of $\Delta\theta$ near the surface for $\mu = 0.4$ under $L_d/t_d = 5.1$ agrees with that for $\mu = 0.3$ under $L_d/t_d = 7.5$. At this time, the γ at the surface for these conditions was almost the same value of ~ 4.7 , but the ε_{eq} was different, where $\varepsilon_{eq} = 3.4$ for $\mu = 0.3$ under $L_d/t_d = 7.5$ and $\varepsilon_{eq} = 2.9$ for $\mu = 0.4$ under $L_d/t_d = 5.1$ (Fig. 6(c)).

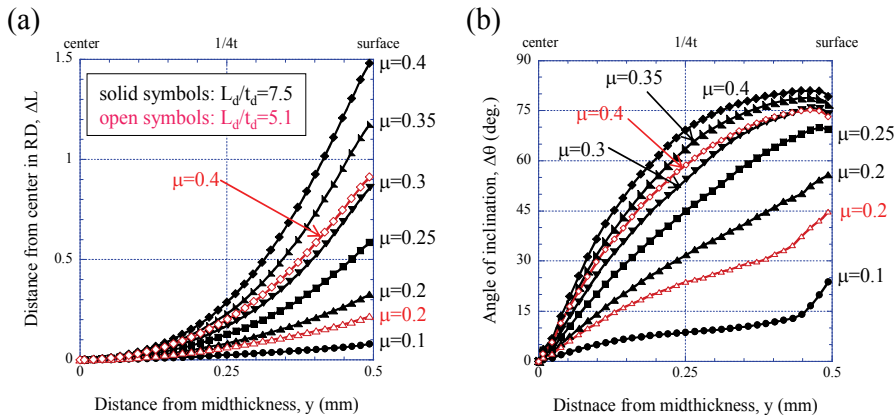


Fig. 8. Variations of (a) ΔL and (b) $\Delta\theta$ through sheet thickness after rolling with friction coefficient μ under $L_d/t_d = 7.5$ (No. 2 in Table 1) and $L_d/t_d = 5.1$ (No. 4 in Table 1).

3.3.3 Strain histories during rolling

The histories of the total strain in the x direction, ε_{xx} , the total shear strain, γ , and the equivalent strain, ε_{eq} , at five elements, i.e., e1, e4, e7, e17 and e33, through sheet thickness during rolling with $\mu = 0.4$ under $L_d/t_d = 7.5$ are shown in Fig. 9, including deformed meshes at times of 0.0875, 0.1, and 0.14 s illustrated in the inset. Here, the NP denotes the time at which the shear stress τ_f for each element equals zero and the stress in the y direction becomes a maximum (Underwood, 1952). The e1 corresponds to the element located in the sheet surface, e4, to the element for which the $\Delta\theta$ indicated a maximum value for $\mu = 0.4$ (Fig. 8(b)), e33, to the element located in the center, e17, to the element at 1/4 thickness located between e1 and e33, and e7, to the element between e4 and e17. From Fig. 9, all strains at e17 monotonously increase with increasing time. This means that each strain rate, $d\varepsilon_{xx}/dt$, $d\gamma/dt$, and $d\varepsilon_{eq}/dt$, at thickness locations from the center to the 1/4 thickness is almost constant during rolling. Furthermore, the strain rates at e17 increases as a result of

the shear deformation more than that at e33 without γ . Hence, the ϵ_{eq} at the 1/4 thickness (e17) becomes larger than that at the center (e33) which agrees with the value $2/\sqrt{3} \ln\{1/(1-r)\}$ calculated simply from reduction in thickness due to $\gamma=0$ constantly. On the other hand, all strain rates at e7, e4, and e1 near the surface vary during rolling by the γ effect, and the ϵ_{xx} and ϵ_{eq} increase as a thickness location approaches the surface (Fig. 9(a,c)). However, the γ at e4 corresponding to the thickness location for which the $\Delta\theta$ indicated a maximum value in Fig. 8(b) is smaller than that at e7 (Fig. 9(b)). The shear strain rate, $d\gamma/dt$, at e4 is initially faster than that at e7, but the γ indicates the same value, $\gamma^+ \approx 2.9$, before the NP, because the time of $d\gamma/dt \approx 0$ during rolling is longer in e4 than in e7. After the NP, the $d\gamma/dt$ at e7 inversely becomes faster than that at e4. Namely, although the positive shear strain, γ^+ , is the same at these two locations, the negative shear strain, γ^- , at e4 is smaller than that at e7.

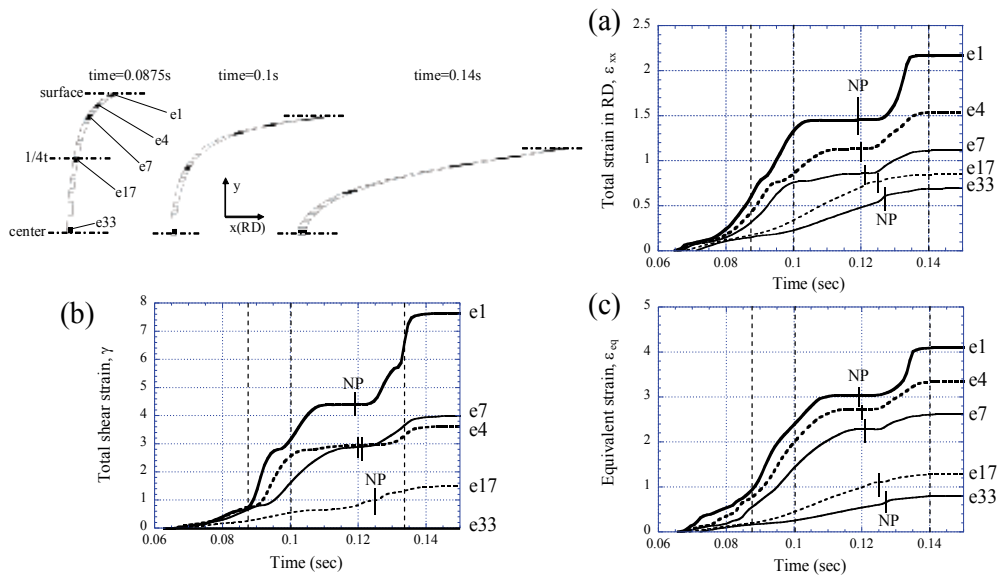


Fig. 9. Histories of (a) ϵ_{xx} , (b) γ , and (c) ϵ_{eq} at five thickness locations (as inset) during rolling with $\mu=0.4$ under $L_d/t_d=7.5$ (No. 2 in Table 1).

Such shear strain history was not seen in $0.13 \leq \mu \leq 0.4$ under $L_d/t_d = 5.1$ ($d^\phi=118$ mm). This feature is further exhibited in Fig. 10, where the variations of γ , γ^+ , and γ^- through thickness are shown. Here, the thickness locations corresponding to five elements illustrated in the inset of Fig. 9 are displayed. In Fig. 10(a), the γ^+ initially increases with the distance y . Its incremental rate becomes larger from e17 ($y=0.25$) to e7 ($y=0.4$), becomes constant until e4 ($y=0.45$) thereafter, and shows a steep increase from e4 to e1 ($y=0.5$). On the other hand, the γ gradually increases from e33 ($y=0$) to e7 ($y=0.4$), decreases from e7 to e4 ($y=0.45$), and shows a sharp increase thereafter. The magnitude of γ^- is smaller than that of γ^+ throughout thickness locations. As a result, the γ which expresses these sums shows a distribution with a sudden dip at $y=0.45$ near the surface, where the $\Delta\theta$ indicated a

maximum value (Fig. 8(b)). Although a similar variation of $\Delta\theta$ with a maximum was seen in $\mu = 0.4$ under $L_d/t_d = 5.1$, the γ distribution with the sudden dip in Fig. 10(a) does not appear in Fig. 10(b), where the variations of γ , γ^+ , and γ^- through thickness after rolling with $\mu = 0.4$ under $L_d/t_d = 5.1$ are shown. In Fig. 10(b), the magnitude of γ^- is smaller than that of γ^+ throughout thickness locations except the surface, and the γ^- exhibits the same magnitude as the γ^+ at the surface. The feature which γ^+ equals γ^- at the surface was the same in the friction coefficient range of $0.13 \leq \mu \leq 0.4$ under $L_d/t_d = 5.1$. This feature was also observed in $0.1 \leq \mu \leq 0.3$ under $L_d/t_d = 7.5$, as shown in Fig. 6(b) of T.Inoue & N.Tsuiji, 2009. Consequently, the magnitude of the shear strain γ^+ before the NP is larger than that of the reverse shear strain γ^- after the NP through thickness, but these magnitudes exhibit the same value at the sheet surface. Provided that a sheet is rolled with a high friction condition under a large roll bite geometry, the magnitude of γ^+ becomes larger than that of γ^- throughout thickness due to the sudden dip of γ^- near the surface.

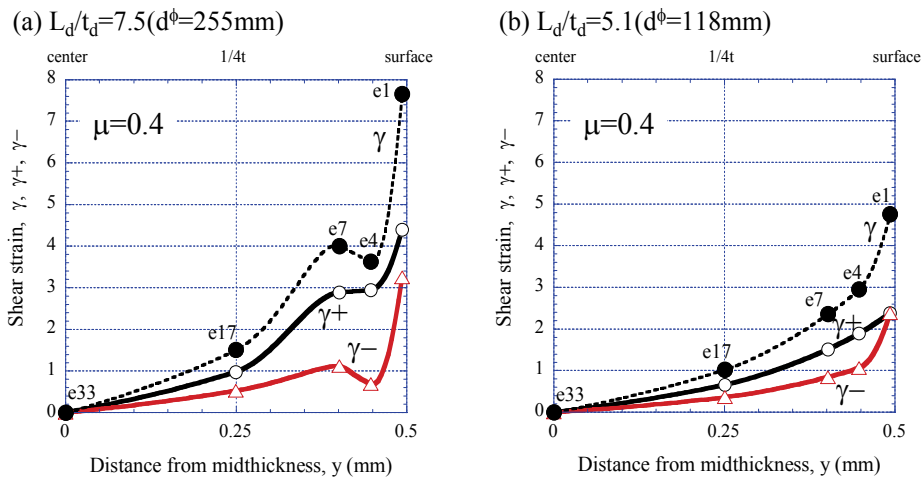


Fig. 10. Variations of the total shear strain γ , positive shear strain γ^+ , and reverse shear strain γ^- through sheet thickness after rolling with $\mu=0.4$. Here, e1, e4, e7, e17 and e33 are displayed in the inset of Fig. 9.

3.4 Strain distributions through sheet thickness after rolling

According to the embedded-pin method (Sakai et al., 1988; Matsuoka et al., 1997), the “apparent” shear strain, $\gamma_{(pin)}$, shows a maximum at the thickness location of 0.1~0.2 from surface, and, hence, the “apparent” equivalent strain, $\varepsilon_{eq(pin)}$, also takes the maximum there. Here, in Sakai et al., 1988, a large roll bite geometry of 9.1 had been employed. This is clear from the variation of $\Delta\theta$ through sheet thickness in Fig. 8(b). The $\gamma_{(pin)}$ is calculated from the inclination of the pin after rolling. Provided the direction of shear deformation remains unchanged during rolling, the $\gamma_{(pin)}$ and $\varepsilon_{eq(pin)}$ must take their maximum at the surface. However, since the direction of shear deformation changes to the opposite direction at the NP, the inclination of mesh at the surface becomes smaller in the surface layer under high friction and large roll bite geometry conditions, as shown in Fig. 8(b). Therefore, the $\varepsilon_{eq(pin)}$

measured by the embedded-pin method would be underestimated in the surface layer compared with the “substantial” equivalent strain obtained in the present study.

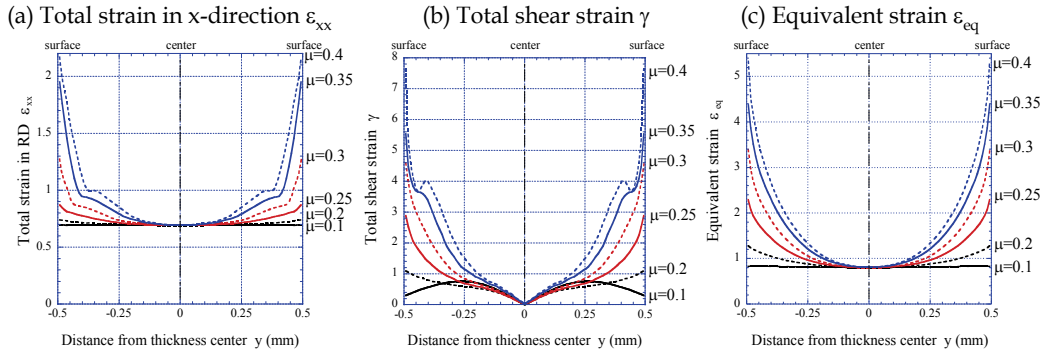


Fig. 11. Distributions of ε_{xx} , γ , and ε_{eq} through sheet thickness for various friction coefficients μ under $L_d/t_d = 7.5$ (No. 2 in Table 1).

Figure 11 represents the distributions of the “substantial” γ and ε_{eq} through the sheet thickness after rolling with various μ under the rolling condition No. 2 in Table 1. Similarly, the distribution of ε_{xx} ($= -\varepsilon_{yy}$) is shown in Fig. 11 because all the strains are associated as shown in Fig. 9. The γ at the thickness center is always zero in Fig. 11(b). The γ for $\mu=0.1$ takes a maximum at $y=\pm 0.25$ mm ($1/4t$) and decreases toward the surface thereafter. It is likely that this distribution results from the roll bite geometry L_d/t_d , because the corresponding strain ε_{xx} is constant through the sheet thickness as shown in Fig. 11(a). The ε_{xx} for friction coefficient other than $\mu=0.1$ gradually increases toward the surface from the center, but for $\mu=0.35$ and 0.4 there is a sharp raise at the thickness location of $0.05\sim 0.15$ from surface. In Fig. 11(b), γ for $\mu=0.2$ to 0.3 increases toward the surface from the center. However, the distributions of γ for $\mu \geq 0.35$ are different from those for $\mu=0.2$ to 0.3 . Especially for $\mu=0.4$ there is a sudden dip at the thickness location near surface. It is considered that this behavior is attributed to the decrease in the reverse shear strain γ_- as shown in Fig. 10(a). If the γ_+ and γ_- are the same magnitude regardless of the change in μ , the total shear strain γ might gradually increase toward the surface from the center such as γ for $\mu=0.2$ to 0.3 , and, furthermore, larger total shear strain might occur at the surface. On the other hand, in Fig. 11(c), ε_{eq} at the thickness center is constant regardless of μ , and its magnitude indicates 0.80 because no shear strain is imposed at the center. Although ε_{eq} for $\mu=0.1$ looks almost constant throughout the thickness, the ε_{eq} showed a distribution in the strain range of $0.8 \leq \varepsilon_{eq} \leq 0.83$; it increases toward the surface from zero ($y=0$) at the center, takes a maximum at $y \approx \pm 0.45$ mm, and decreases thereafter. It is considered that this distribution is the effect of L_d/t_d . The ε_{eq} for all other μ gradually increases toward the surface from the center and shows a distribution with the maximum at the surface. The maximum ε_{eq} increases with increasing μ , and ε_{eq} for $\mu=0.4$ reaches 5.33 , which is six times higher than that at the center. The $\varepsilon_{eq}=5.33$ corresponds to a 99% reduction in plane strain compression. In the distribution of ε_{eq} , a sharp raise or a sudden dip, as seen near the surface of ε_{xx} and γ are not observed. This means that the equivalent strain ε_{eq} that denotes

scalar amount varies continuously throughout sheet thickness regardless of the roll bite geometry and the friction between the rolls and the material surface. On the other hand, strains in each component that denote vector amount vary complicatedly in the sheet thickness. Hence, it is found that a much larger ε_{eq} can be introduced to the sheet surface through a complicated deformation by the shear strain effect.

3.5 Strain distributions in ARB

There are numerous reports on microstructure evolution in various sheet materials by ARB process (Hidalgo et al., 2010; Kolahi et al., 2009; Xing et al., 2002). In most papers, a rolling by a 50% reduction in thickness is conducted without lubricant, and equivalent strain calculated simply from reduction in thickness has been used regardless of a high friction condition. However, it is found from previous results that equivalent strain introduced in a rolled sheet is different by the effects of not only friction condition but also roll diameter used.

From Figs. 6(a,c) and 11(b,c), γ and ε_{eq} at the surface imposed by the first cycle of ARB at ambient temperature reported by Lee et al., 2002 show the values of 5.60 and 4.23, respectively. That is, a much larger equivalent strain, which is five times higher than that at the center, is introduced to the surface. Here, it is considered that the friction coefficient is $\mu \approx 0.34$. Moreover, assuming that the same deformation is repeated in the following cycles, γ and ε_{eq} at the surface in the multi-cycle ARB-processed 1100 Al can be estimated by $\gamma = 5.60 N$ and $\varepsilon_{eq} = 4.23 N$, respectively, where N denotes the number of cycles. On the other hand, the γ and ε_{eq} at the center are calculated by $\gamma = 5.60 (N-1)$ and $\varepsilon_{eq} = 0.8 + 4.23 (N-1)$, respectively. Utsunomiya et al., 1999 showed the flection of the pin in the 1100 Al processed by one ARB cycle at 473 K with and without lubricant under the same roll bite geometry as Lee et al., 2002. Here, ΔL s indicated ~ 0.15 with lubricant and ~ 1.5 without it. Since the effect of temperature in the rolled Al or IF steel sheet on the relation among ε_{eq} , γ , and ΔL s is small in comparison with the friction effect (Um et al., 2000; Inoue & Tsuji, 2009), it is likely from Figs. 6(a,c) and 11(b,c) that the strain at the surface is $\gamma \approx 0.59$ and $\varepsilon_{eq} \approx 0.97$ for rolling with lubricant and $\gamma \approx 7.66$ and $\varepsilon_{eq} \approx 5.34$ for rolling without it. Here, it is considered that the friction coefficient is about ~ 0.12 for rolling with lubricant and ~ 0.40 for rolling without it. That is, equivalent strain and shear strain at the surface in a sheet rolled without lubricant are five times and twelve times higher than the strains with it. Namely, ε_{eq} at the surface in the 1100 Al processed by one ARB cycle without lubricant would correspond to ε_{eq} in that processed by five ARB cycles with lubricant. Consequently, since ARB process shown in Fig. 1(b) is usually conducted by a 50% reduction of thickness without lubricant to aid bonding, a rolling condition with large roll bite geometry is desirable for fabricating ultrafine-grained materials efficiently.

3.6 Variations of the total shear strain γ and equivalent strain ε_{eq} against ΔL or $\Delta \theta$

The embedded-pin method has a limitation on the quantification of strains imposed by rolling, as mentioned previously. In particular, the strains near the surface are determined by very complicated histories during rolling, as shown in Fig. 9. However, if the magnitude and distribution of strains in a rolled sheet can be quantitatively estimated by using experimental measurements, such as ΔL or $\Delta \theta$ observed from the pin flection shown in Fig. 3, these quantitative strain analyses would be useful for designing the microstructure of the

sheet metal by rolling. Figure 12(a,b) shows the variations of γ and ϵ_{eq} with ΔL through sheet thickness for different values of μ under $L_d/t_d = 7.5$ and Fig. 12(c,d) shows the variations of γ and ϵ_{eq} with $\Delta\theta$. Here, all data were obtained from the values of the integration point for 33 elements in sheet thickness, as shown in Fig. 8. The $\Delta L = 0$ and $\Delta\theta = 0$ correspond to a location of the thickness center.

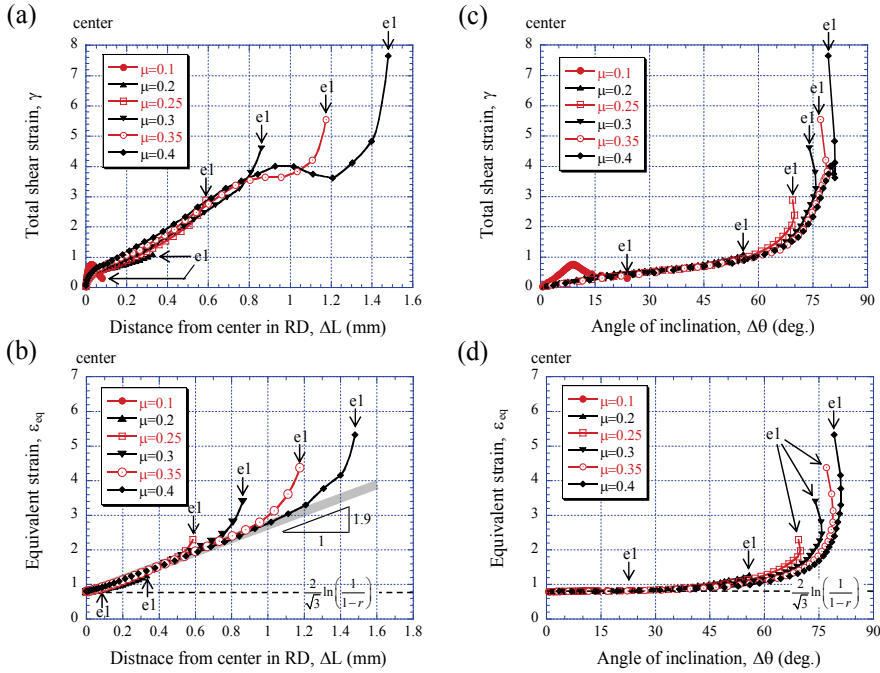


Fig. 12. Variations of (a) γ and (b) ϵ_{eq} with ΔL and (c) γ and (d) ϵ_{eq} with $\Delta\theta$ through thickness in 50% rolled Al for various friction coefficients μ under $L_d/t_d = 7.5$ (No. 2 in Table 1). Here, e1 denotes data at the sheet surface shown in the inset of Fig. 9.

In Fig. 12(a,c), the $\gamma - \Delta L$ and $\gamma - \Delta\theta$ relations for $\mu = 0.1$ are significantly different from the relations for other μ . The γ for the μ close to $\tan \alpha$, which is a condition to pull a sheet into the roll bite, took a maximum at $y \approx \pm 0.25$ mm (1/4 thickness) and decreased toward the surface thereafter, as shown in Fig. 11(b). This is because the distribution of γ is dominated by the roll bite geometry L_d/t_d due to the small effect of μ . The γ for $\mu \geq 0.2$ increases as the ΔL or $\Delta\theta$ becomes larger, i.e., the thickness location approaches a sheet surface. However, since there is a decrease of the negative shear strain γ_- near the surface, as shown in Fig. 10(a), the γ for $\mu \geq 0.35$ initially increases with increasing the ΔL or $\Delta\theta$, decreases near the surface, and shows a steep increase thereafter. The $\gamma - \Delta\theta$ relations, except $\mu = 0.1$, have a good correlation, compared to the $\gamma - \Delta L$ relation. In particular, the variation of γ with $\Delta\theta$ until 60° is almost the same despite μ . It is likely that the $\Delta\theta$ depends strongly on the shear deformation. It is evident from Fig. 12(b,d) that the ϵ_{eq} , including all strain components as defined in equation (2), has very good correlation with the ΔL and $\Delta\theta$. The ϵ_{eq} monotonically increases with increasing the ΔL , and the relation is expressed by $\epsilon_{eq} = 1.9 \Delta L$

+ 0.8, but the ε_{eq} sharply increases near the surface by the effect of the shear strain (Fig. 12(b)). On the other hand, in Fig. 12(d), the ε_{eq} remains 0.8 calculated by $2/\sqrt{3} \ln\{1/(1-r)\}$ in the range of $0 \leq \Delta\theta \leq 30^\circ$, slightly increases with $\Delta\theta$ thereafter, and sharply increases at $\Delta\theta \geq 70^\circ$ by the shear deformation effect. It is clear from Fig. 12 that the ε_{eq} has better correlation with the ΔL and $\Delta\theta$, than γ . This means that ΔL and $\Delta\theta$ are determined by a combination of the strains in the x and y directions and the shear strain and their histories during rolling. As shown in Figs. 6 and 8, the deformation behaviors in a rolled sheet depend strongly on not only the friction but also the roll bite geometry, especially when rolling with high friction. Figure 13 shows the variations of ε_{eq} with ΔL and $\Delta\theta$, respectively, in a 50% rolled Al with $\mu=0.4$ under various L_d/t_d of five rolling conditions, Nos. 1-5, in Table 1. These correlations agree with the $\varepsilon_{eq} - \Delta L$ and $\varepsilon_{eq} - \Delta\theta$ relations shown in Fig. 12(b,d). Consequently, the L_d/t_d ratio has a similar effect on μ regarding the relation among ε_{eq} , ΔL , and $\Delta\theta$ through sheet thickness after rolling.

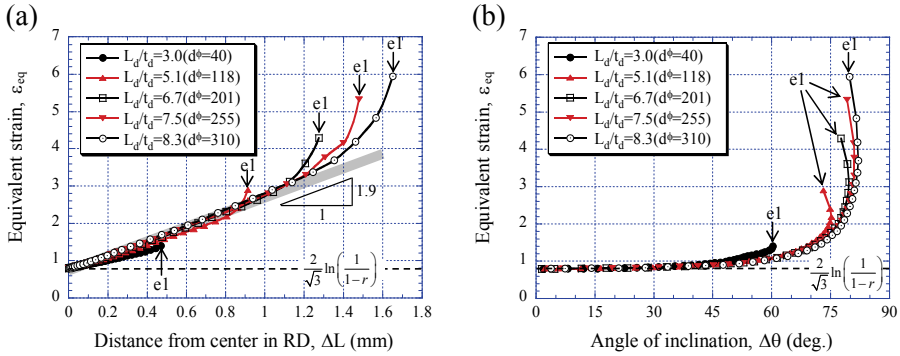


Fig. 13. Variations of ε_{eq} with (a) ΔL and (b) $\Delta\theta$ through thickness in 50% rolled Al with $\mu=0.4$ under various L_d/t_d (Nos. 1-5 in Table 1). Here, e1 denotes data at the surface shown in the inset of Fig. 9.

3.7 Universality of the roll bite geometry on strains through thickness in a rolled sheet

As a condition to impose a large equivalent strain by rolling, the roll bite geometry is as important a processing parameter as the friction condition. The roll bite geometry L_d/t_d is determined by the thickness before and after rolling, t_0 and t_1 , and the roll diameter, d^ϕ , as shown in Fig. 2. However, even if the L_d/t_d is the same value, the deformation behaviors through sheet thickness may be different by a combination of t_0 , t_1 , and d^ϕ because the deformation histories during rolling are very complicated under high friction conditions. Figure 14 shows the distributions of $\Delta\theta$ and equivalent strain added by the shear deformation, $\varepsilon_{eq(shear)}$, through thickness in the rolled Al with $\mu=0.3$ under the rolling conditions Nos. 4, 7-9 of $L_d/t_d = 5.1$ in Table 1. Here, the horizontal axis was normalized by the half of sheet thickness after rolling, i.e., $2y/t_1 = 0, \pm 0.5$, and ± 1 indicate the center, $1/4$ thickness, and surface, respectively, in the sheet. The $\varepsilon_{eq(shear)}$ represents the equivalent strain added by shear deformation as defined by $\varepsilon_{eq} - 2/\sqrt{3} \ln\{1/(1-r)\}$ and equals zero at the center of $\gamma=0$. It is found from Fig. 14(a) that, for all conditions, the $\Delta\theta$ through thickness

does not exhibit exactly the same distribution and magnitude. In particular, the magnitudes for Nos. 4 and 9 at the same r are significantly different from those for Nos. 7 and 8.

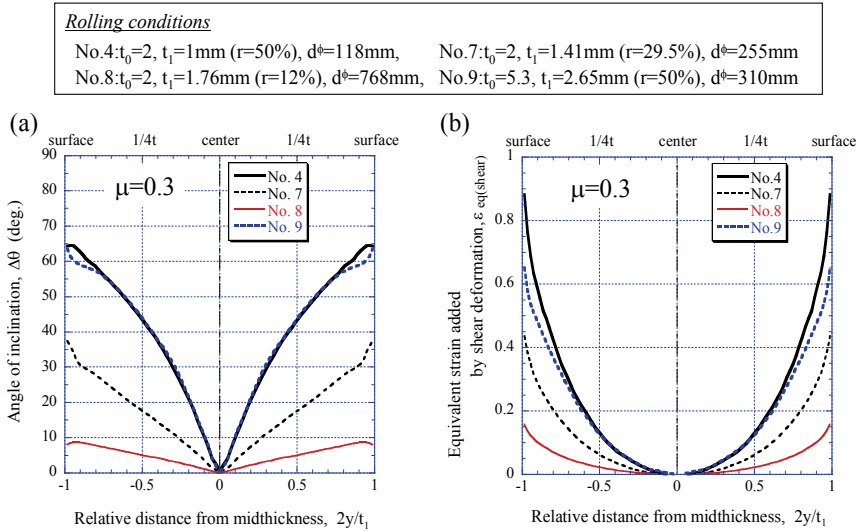


Fig. 14. Distributions of (a) $\Delta\theta$ and (b) equivalent strain added by shear deformation, $\varepsilon_{eq(shear)}$, through sheet thickness in rolled Al with $\mu=0.3$ under $L_d/t_d=5.1$. Here, full details of the rolling conditions are shown in Table 1.

Numbers 4 and 9 indicate the same magnitude at all thickness locations except near $y/t_1 = \pm 0.9$. The magnitude of $\Delta\theta$ decreases with decreasing r under the same t_0 (Nos. 4, 7, 8). The friction coefficient range of $0.13 \leq \mu \leq 0.4$ showed a similar feature. This result indicates that the $\Delta\theta$ measured from the embedded-pin method is different by r even if the L_d/t_d has the same value under the same t_0 . Therefore, we should use a reduction of the thickness, r , as a variable of the L_d/t_d . In Fig. 14(b), the $\varepsilon_{eq(shear)}$ also exhibits the same tendency. At the surface, the $\varepsilon_{eq(shear)}$ for No. 4 of $r=50\%$ is 0.88, which is twice as high as that for No. 7 of $r=29.5\%$ and five times higher than that for No. 8 of $r=12\%$. These values are similar to the relations of the equivalent strain calculated simply from the r , i.e., ε_{eq} at the center; $\varepsilon_{eq} = 0.8, 0.4, 0.15$ for $r=50\%, 29.5\%, 12\%$, respectively. Furthermore, although the $\Delta\theta$ at the surface in Nos. 4 and 9 indicated the same value in Fig. 14(a), the $\varepsilon_{eq(shear)}$ is larger in No. 4 than in No. 9. This feature was the same in the range of $\mu \geq 0.3$, and the difference of $\varepsilon_{eq(shear)}$ at the surface increased with increasing μ . However, the $\varepsilon_{eq(shear)}$ at the surface in the Nos. 4 and 9 indicated the same magnitude in the range of $\mu \leq 0.25$. On the other hand, the $\varepsilon_{eq(shear)}$ from the center to the $1/4$ thickness in No. 4 is in good agreement with that in No. 9 (Fig. 14(b)). A similar feature was seen in the friction coefficient range of $0.13 \leq \mu \leq 0.4$. As shown in Fig. 9, at thickness locations from the center (e33) to the $1/4$ thickness (e17), the strain rates are almost constant during rolling. In other words, the magnitude of $\varepsilon_{eq(shear)}$ at the thickness locations which take a constant strain rate during rolling agrees under the L_d/t_d ratio with the same r . However, the $\varepsilon_{eq(shear)}$ from the $1/4$ thickness to the surface is not the same magnitude in the range of $\mu \geq 0.3$, and the $\varepsilon_{eq(shear)}$ at the locations becomes larger with

decreasing t_0 for the same r . Figure 15 shows the distributions of $\Delta\theta$ ($-1 \leq 2y/t_1 \leq 0$) and $\varepsilon_{eq(shear)}$ ($0 \leq 2y/t_1 \leq 1$) for $\mu=0.25$. It can be seen that the $\varepsilon_{eq(shear)}$ and $\Delta\theta$ in No. 4 are almost the same magnitude and distribution as those in No. 9. Consequently, even if the L_d/t_d is the same value, the strain imposed by rolling does not exhibit exactly the same magnitude and distribution through sheet thickness because the deformation behaviors during rolling are not simple due to the effect of the shear deformation. Under the $L_d/t_d = 5.1$, the $\varepsilon_{eq(shear)}$ increases with increasing r for the same t_0 . However, provided that the L_d/t_d is employed under the same r , the ε_{eq} through sheet thickness agrees regardless of the combination of t_0 and d^ϕ in the friction coefficient range of $\mu \leq 0.25$. For high friction of $\mu \geq 0.3$, the ε_{eq} at locations from $1/4$ thickness to the surface becomes larger as t_0 decreases, and this tendency becomes more remarkable with increasing μ .

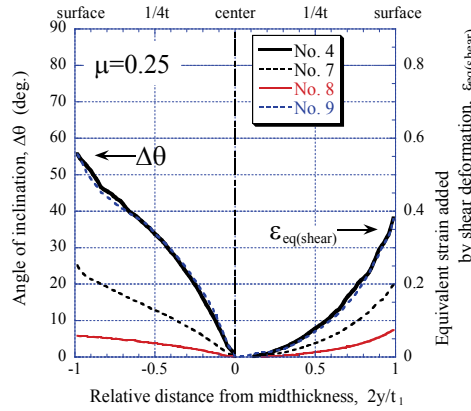


Fig. 15. Distributions of $\Delta\theta$ (lefthand side) and $\varepsilon_{eq(shear)}$ (righthand side) through sheet thickness in rolled Al with $\mu=0.25$ under $L_d/t_d=5.1$.

As a result, the L_d/t_d parameter shown in Fig. 1 should be expressed by the following equation including t_0 , d^ϕ and r , using the relation of $t_1=t_0/(1-r)$.

$$\frac{L_d}{t_d} = \frac{1}{2-r} \sqrt{\frac{2dr}{t_0}} \quad (4)$$

The relation among the roll bite geometry L_d/t_d , roll diameter d^ϕ , nominal reduction r , and initial thickness t_0 is plotted in Fig. 16 for various combinations of d^ϕ , r , and t_0 . Here, the open symbols indicate the rolling conditions employed in the literature related to the microstructure change through sheet thickness in a rolling process, including ARB. The L_d/t_d gradually increases with increasing d^ϕ under constant t_0 and r (Fig. 16(a)) and monotonically increases with increasing r under constant d^ϕ and t_0 (Fig. 16(b)). Under d^ϕ and r constant (Fig. 16(c)), the L_d/t_d gradually increases with decreasing t_0 , and there is a steep increase at $t_0 < 5$ mm. Compared to the increases of d^ϕ or r , the decrease of t_0 has a large influence on the increase of L_d/t_d . It is found from Fig. 16 that the L_d/t_d is different with regard to t_0 , d^ϕ , and r . Even if the friction condition is the same, this causes the

difference in the magnitude and distribution of strains, as shown in the present study, which results in the changes of the microstructure and texture depending on the thickness location of a rolled sheet.

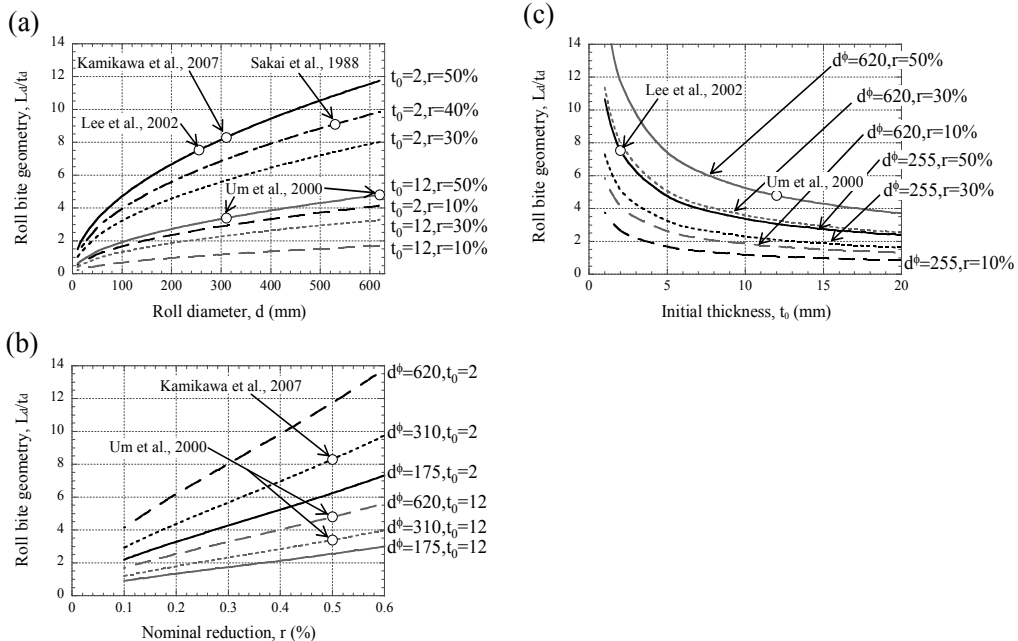


Fig. 16. Variations of the roll bite geometry L_d/t_d with (a) roll diameter d^ϕ , (b) nominal reduction r , and (c) initial thickness t_0 . Here, the open symbols indicate the rolling conditions employed in the literature.

4. Outlook

It is commonly accepted that the distribution of the equivalent strain through thickness in a rolled sheet depends on the presence of shear deformation due to friction between the rolls and the sheet and brings about a change of microstructure (grain size, texture) through sheet thickness. It is found from the present results that the roll bite geometry defined by equation (4), as well as the friction coefficient, is an important parameter for studying microstructural changes through thickness in metal sheet rolling including ARB process. The grain size decreases with an increase of the equivalent strain, and, hence, a rolling condition with high reduction using a mill with a large roll diameter without a lubricant is effective for the refinement of crystal grains. Such rolling conditions can be seen in hot rolling (high friction and large roller) rather than cold rolling (low friction condition and small roller) and in a commercial rolling mill of product level (large roller) rather than a rolling mill of laboratory level (small roller). Moreover, the microstructural evolution depends strongly on not only equivalent strain but also shear strain. Inoue et al., 2002 reported that in the ferrite grain transformed from deformed austenite, the grain size in the area with the shear strain is finer than that in the area without the shear strain under the same equivalent plastic strain. Cho et al., 2004; Kang et al., 2010 demonstrated significant advances in high angle grain

boundaries and the subdivision of grains under the deformation with shear strain than under without shear strain using Ni-30Fe alloy and low carbon steel. In other words, it is noted that the shear deformation plays a important role for refining crystal grains.

Also, it is known that a shear texture different from a conventional rolling texture develops at the locations where shear deformation is introduced by rolling. Control technology of texture associated with the shear deformation is positively used to improve the deep drawability (ND//<111>) and magnetic characteristic (Goss orientation: {110}<001>). The texture of ND//<111>, which contributes to improve the deep drawability, is produced by a large shear deformation in fcc metals, but its deformation is disadvantageous in bcc metals (Matsuo, 1989). Therefore, for a study of deep drawability, using a large roller is important for the case of aluminum alloys (fcc), and a small roller is recommended for the case of IF steels (bcc). In study of deep drawability in ferritic rolling of extra low carbon steels, Kawabe et al., 1996 reported that the r-value in the sheet improved more significantly when using a small than a large roller under the same friction condition. They confirmed, using the embedded-pin method, that the improvement in the r-value was the result of the decrease of the shear deformation through sheet thickness. Consequently, in order to study universal relation between microstructure (grain size, texture) and plastic deformation, it is essential to understand the quantity (magnitude) and quality (component) of the strain into materials introduced by a plastic deformation process through a combination of numerical simulations and experimental measurements and observations. And a quantitative relation among the strains, friction, and roll bite geometry obtained from numerical simulations would provide useful guidelines for studying the microstructure design in a rolled sheet and for understanding the quantitative correlation between microstructures and strain in ARB process.

5. Conclusions

The effect of the roll bite geometry, L_d/t_d , on the magnitude and distribution of strains imposed in a metal sheet by rolling under various friction conditions was studied using a finite element simulation that takes the deformation history into account. The relation among the strains, distance from center in RD, ΔL , and angle of inclination from the y axis, $\Delta\theta$, was shown through thickness in a rolled 1100 Al. The present results will provide useful guidelines for studying the correlation between microstructures and strain in accumulative roll-bonding (ARB) process as well as the microstructure design in a rolled sheet. The main results are as follows:

1. The deformation through sheet thickness during and after rolling depends strongly on not only the friction condition between the rolls and the sheet but also the roll bite geometry L_d/t_d . The L_d/t_d ratio has a similar effect on friction coefficient μ regarding the relation among ε_{eq} , ΔL , and $\Delta\theta$ through sheet thickness after rolling. In other words, the roll bite geometry is as important a processing parameter as the friction condition for studying microstructural changes through thickness in metal sheet rolling and ARB process.
2. The ε_{eq} in a rolled sheet gradually increases toward the surface from the center and shows a distribution with the maximum at the surface. The maximum ε_{eq} increases with increasing μ or L_d/t_d . On the other hand, in the total shear strain γ , there is a sudden

dip at the thickness location near surface under high friction condition. This behavior is attributed to the decrease in the reverse shear strain γ – induced by the shear stress after a neutral plane.

3. Under different values of the friction coefficient or roll bite geometry, the equivalent strain ε_{eq} had much better correlation with the ΔL and $\Delta\theta$ through sheet thickness, except near the surface, than the total shear strain γ . This is attributed to the fact that ΔL and $\Delta\theta$ are determined by a combination of strains in the x and y directions and the shear strain and their histories during rolling.
4. The L_d/t_d is expressed by equation including initial thickness, t_0 , roll diameter, d^ϕ , and nominal reduction, r . The L_d/t_d at the same r can be employed as a universal parameter on the equivalent strain in a rolled sheet, except for thickness locations near the surface under high friction conditions.
5. In ARB process, which half of the sheet-surface regions comes to the center in the next cycle and its procedure is repeated, a rolling condition with large roll bite geometry at a high friction condition is desirable for fabricating ultrafine-grained materials efficiently.

6. References

- ABAQUS/Explicit ver. 6.5 User's manual, Theory manual; (2006). Providence, ABAQUS Inc., CD-ROM.
- Ataka, M., (2006). *Handbook of Technology of Plasticity*, Corona Publishing Co., Ltd., Tokyo, pp. 11-131.
- Backofen, W.A., 1972. *Deformation Processing*, Addison-Wesley Pub. Co., Massachusetts, pp. 94-115.
- Cho, J.Y., Inoue, T., Yin, F. & Nagai, K. (2004). Effect of shear deformation on microstructural evolution of Ni-30Fe alloy during hot deformation. *Materials Transactions*, Vol. 45, pp. 2966-2973.
- Cui, Q. & Ohori, K. (2000). Grain refinement of high purity aluminium by asymmetric rolling. *Materials Science and Technology*, Vol. 16, pp. 1095-1101.
- Dieter, G.E., 1988. *Mechanical Metallurgy SI Metric Edition*, McGraw-Hill Book Co., 0-07-100406-8, Singapore, pp. 501-563.
- Hashimoto, S., Tsukatani, I., Kashima, T. & Miyoshi, T. (1998). Development of hot-rolled steel sheet with a high r-value by rolling in ferrite region with lubrication. *Kobe Steel Engineering Reports*, Vol. 48, pp. 14-18.
- Hidalgo, P., Cepeda-Jiménez, C.M., Ruano, O.A. & Carreño, F. (2010). Influence of the Processing Temperature on the Microstructure, Texture, and Hardness of the 7075 Aluminum Alloy Fabricated by Accumulative Roll Bonding. *Metallurgical and Materials Transactions A*, Vol. 41, pp. 758-767.
- Horita, Z., Fujinami, T., Nemoto, M. & Langdon, T.G. (2000). Equal-channel angular pressing of commercial aluminum alloys: grain refinement, thermal stability and tensile properties. *Metallurgical and Materials Transactions A*, Vol. 31, pp. 691-701.
- Inoue, T., Torizuka, S., Nagai, K., Tsuzaki, K. & Ohashi, T. (2001). Effect of plastic strain on grain size of ferrite transformed from deformed austenite in Si-Mn steel. *Materials Science and Technology*, Vol. 17, pp. 1580-1588.
- Inoue, T., Torizuka, S. & Nagai, K. (2002). Effect of shear deformation on refinement of crystal grains. *Materials Science and Technology*, Vol. 18, pp. 1007-1015.

- Inoue, T., Ochiai, T., Yin, F. & Nagai, K. (2007a). Test production of ultrafine-grained steel plate by large-scale forging press. *Tetsu-to-Hagane*, Vol. 93, pp. 693-702.
- Inoue, T., Yin, F. & Kimura, Y. (2007b). Strain distribution and microstructural evolution in multi-pass warm caliber rolling. *Materials Science and Engineering A*, Vol. 466, pp.114-122.
- Inoue, T., Horita, Z., Somekawa, H. & Ogawa, K. (2008). Effect of initial grain sizes on hardness variation and strain distribution of pure aluminum severely deformed by compression tests *Acta Materialia*, Vol. 56, pp. 6291-6303.
- Inoue, T., Todaka, Y. & Horita, Z. (2009a). Special issue on severe plastic deformation for production of ultrafine structures and unusual mechanics properties: understanding mechanisms. *Materials Transactions*, Vol. 50, pp. 1-116.
- Inoue, T. & Tsuji, N. (2009b). Quantification of strain in accumulative roll-bonding under unlubricated condition by finite element analysis. *Computational Materials Science*, Vol. 46, pp. 261-266.
- Inoue, T., Somekawa, H. & Mukai, T. (2009c). Hardness variation and strain distribution in magnesium alloy AZ31 processed by multi-pass caliber rolling. *Advanced Engineering Materials*, Vol. 11-8, pp. 654-658.
- Inoue, T., Yin, F., Kimura, Y., Tsuzaki, K. & Ochiai, S. (2010a). Delamination effect on impact properties of ultrafine-grained low carbon steel processed by warm calibre rolling. *Metallurgical and Materials Transactions A*, Vol. 41-2, pp. 341-355.
- Inoue, T., Yin, F. & Kimura, Y. (2010b). Effect of deformation mode on texture of ultrafine-grained low carbon steel processed by warm caliber rolling. *Materials Science Forum*, Vols. 638-642, pp 2793-2798.
- Kamikawa, N., Sakai, T. & Tsuji, N. (2007). Effect of redundant shear strain on microstructure and texture evolution during accumulative roll-bonding in ultralow carbon IF steel. *Acta Materialia*, Vol. 55, pp. 5873-5888.
- Kang, J.-H., Inoue, T. & Torizuka, S. (2010). Effect of shear strain on the microstructural evolution of a low carbon steel during warm deformation. *Materials Transactions*, Vol. 51, pp. 27-35.
- Kawabe, H., Matsuoka, S., Seto, K., Sakata, T., Furugimi, O. & Ohara, T. (1996). Effect of roll diameter on deep drawability of extra low C sheet steels with hot rolling at ferrite region. *CAMP-ISIJ*, Vol. 9, p. 1333.
- Kimura, Y., Inoue, T., Yin, F. & Tsuzaki, K. (2008). Inverse temperature dependence of toughness in an ultrafine grain structure steel. *Science*, Vol. 320, pp. 1057-1060.
- Kolahi, A., Akbarzaden, A. & Barnett, M.R. (2009). Electron back scattered diffraction (EBSD) characterization of warm rolled and accumulative roll bonding (ARB). *J. Materials Processing Technology*, Vol. 209, pp. 1436-1444.
- Lee, S.H., Saito, Y., Tsuji, N., Utsunomiya, H. & Sakai, T. (2002). Role of shear strain in ultragrain refinement by accumulative roll-bonding (ARB) process. *Scripta Materialia*, Vol. 46, pp. 281-285.
- Matsuo, M. (1989). Texture control in the production of grain oriented silicon steels. *ISIJ International*, Vol. 29, pp. 809-827.
- Matsuoka, S., Morita, M., Furukumi, O. & Obara, T. (1997). Structural variation along thickness direction of extra-low carbon sheet steels rolled in ferrite region. *Tetsu-to-Hagane*, Vol. 83, pp. 127-132.

- Mukai, T., Somekawa, H., Inoue, T. & Singh, A. (2010). Strengthening Mg-Al-Zn alloy by repetitive oblique shear strain with caliber roll. *Scripta Materialia*, Vol. 62, pp. 113-115.
- Mukhopadhyay, A., Higginson, R.L., Howard, I.C. & Sellars, C.M. (2007). Strain summation in finite element modeling of multipass hot rolling. *Materials Science and Technology*, Vol.23-1 pp.29-37.
- Saito, Y., Tsuji, N., Utsunomiya, H., Sakai, T. & Hong, R.G. (1998). Ultra-fine grained bulk aluminum produced by accumulative roll-bonding (ARB) process. *Acta Materialia*, Vol. 39, pp. 1221-1227.
- Sakai, T., Saito, Y., Hirano, K. & Kato, K. (1988). Recrystallization and texture formation in high speed hot rolling of austenitic stainless steel. *Trans ISIJ*, Vol. 28, pp. 1028-1035.
- Sakai, T. & Saito, Y. (1999). Effect of inhomogeneous deformation through the thickness on microstructure and texture of hot rolled sheet. *J. Japan Society Technology Plasticity*, Vol. 40, pp. 1158-1163.
- Segal, V.M. (1995). Materials processing by simple shear. *Materials Science and Engineering A*, Vol. 197, pp.157-164.
- Todaka, Y., Inoue, T. & Horita, Z. (2008). Special issue on severe plastic deformation for production of ultrafine structures and unusual mechanics properties: Investigating role of high-density lattice defects. *Materials Transactions*, Vol. 49, pp. 1-106.
- Um, K.-K., Jeong, H.-T., An, J.-K., Lee, D.N., Kim, G. & Kwon, O. (2000). Effect of initial sheet thickness on shear deformation in ferritic rolling of IF-steel sheets. *ISIJ International*, Vol. 40, pp. 58-64.
- Underwood, L.R. (1952). *The Rolling of Metals*. Chapman&Hall Ltd., London, pp. 57-93, 203-268.
- Utsunomiya, H., Tanda, K., Saito, Y., Sakai, T. & Tsuji, N. (1999). *J. Japan Society Technology Plasticity*, Vol. 40, pp. 1187-1191.
- Xing, Z.P., Kang, S. B. & Kim, H.W. (2002). Microstructural evolution and mechanical properties of the AA8011 alloy during the accumulative roll-bonding process. *Metallurgical and Materials Transactions A*, Vol. 33, pp. 1521-1530.
- Zhang, X.J., Hodgson, P.D. & Thomson, P.F. (1996). The effect of through-thickness strain distribution on the static recrystallization of hot rolled austenitic stainless steel strip. *Materials Processing Technology*, Vol. 60, pp. 615-619.

Finite element analysis of wall deflection and ground movements caused by braced excavations

Gordon Tung-Chin Kung
*National Cheng Kung University
Taiwan*

1. Introduction

Excavation is usually employed in the construction of basement of buildings and underground space of the mass rapid transit and the subway system in urban areas (see Fig. 1). Braced excavation is a complicated soil-structure interaction problem and it is a challenge to engineers to simultaneously ensure the safety of excavation system as well as the integrity of properties adjacent to the excavation site including buildings, structures, and life pipes. In a routine excavation design, engineers usually pay much attention to reaching the safety requirements of excavation design, namely, the stability of excavation system. Many empirical, semi-empirical and numerical methods were developed for evaluating the stability of excavation. The past practical experiences indicated that the accuracy of the existing methods to evaluate the stability of excavation is generally acceptable although the ground conditions at various excavation sites are non-homogeneous and have highly spatial variation. This may be attributed to the employment of a relatively conservative factor of safety in the stability analysis of excavation.



Fig. 1. Photos of constructing the mass rapid transit and subway system in Taiwan

According to the practical experiences, it is found that the damage to buildings adjacent to an excavation was reported occasionally even though the stability of excavation can be

ensured. The economic loss of damage to buildings is considerable and such incidents usually cause the procrastination of construction time limit for the excavation project. As such, the serviceability of structure adjacent to the excavation is usually the key factor and plays the predominant role in the performance-based excavation design.

In practice, any of empirical, semi-empirical, or numerical methods can be adopted to evaluate the serviceability of buildings adjacent to an excavation. For empirical and semi-empirical methods, the procedures for estimating the potential of damage to adjacent buildings generally includes three main elements: 1) prediction of excavation-induced ground movements, 2) evaluation of building deformation caused by excavation-induced ground movements estimated, and 3) evaluation of damage potential of buildings based on the building deformation estimated. In the first element, the estimated free-field ground surface settlement caused by excavation is usually employed to evaluate the building deformation. In practice, the empirical and semi-empirical methods, as shown in Fig.2 and Fig. 3, are often selected to estimate the excavation-induced ground movements.

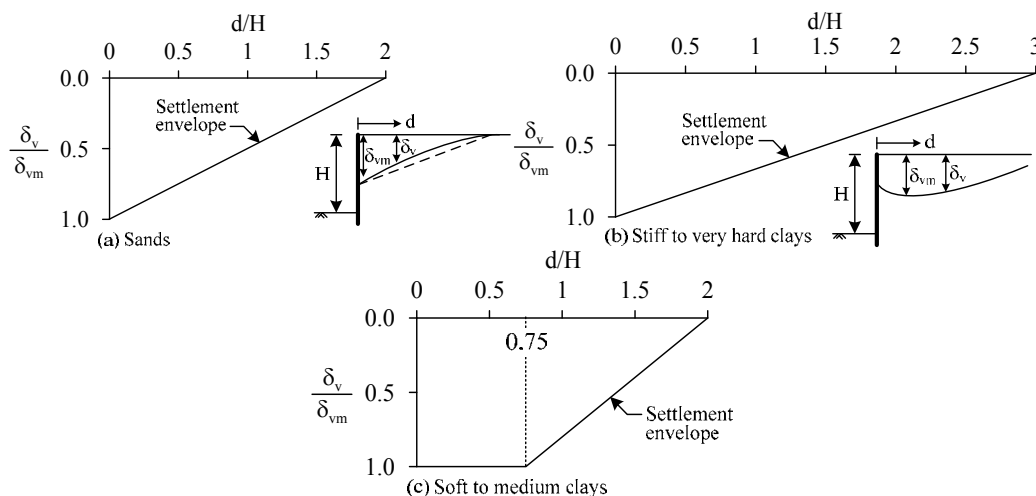


Fig. 2. Design charts for estimating the distribution of surface settlement adjacent to excavation in different soil types (Clough and O'Rourke, 1990)

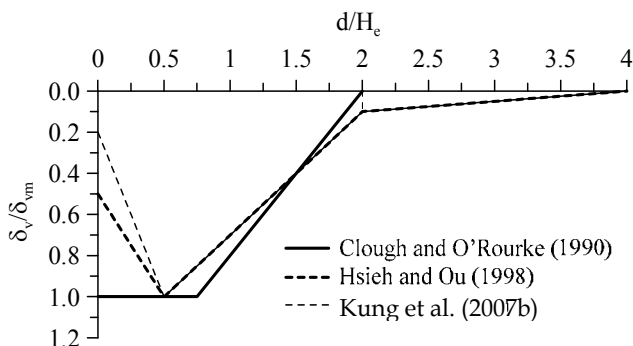


Fig. 3. The ground surface settlement distributions proposed by previous studies for excavations in soft-to-medium clay

It should be noted that these settlement distributions/profiles are established based primarily on the ground surface settlement points monitored in excavation case histories. The applicability of these distributions in the settlement analysis of a new excavation case might be questionable in light of the significant variation of influence factors, such as ground condition and workmanship, between the existing excavation cases and the intended new one. In addition, use of free-field ground surface settlement to estimate the differential settlement of the building may not fully reflect the effect of soil-structure interaction behaviour. It is believed that the ground settlement computed without considering the presence of buildings would be different from that computed with directly incorporating a building into the numerical analysis. Such difference may reduce the accuracy of evaluating damage potential of buildings.

The finite element method (FEM) is often employed to model complex soil-structure interaction problems such as braced excavations. Many investigators (Finno & Harahap, 1991; Hashash & Whittle, 1996; Ng & Yan, 2000; Kung et al., 2007a; Kung et al., 2009) have verified the reliability of FEM applied to analyses of deep excavations even though the uncertainty of soil models and parameters adopted is concerned. Although the deflection of the braced wall can generally be predicted well using a routine FEM analysis, the prediction of surface settlement is usually not as accurate (e.g., Burland, 1989). Previous studies (Simpson, 1993; Whittle et al., 1993; Stallebrass and Taylor, 1997; Kung, 2003; Kung 2007a) have shown that the accuracy of surface settlement predictions by FEM can be significantly improved if the soil behaviour at small strain levels can be properly modelled. Based on the aspects mentioned above, it is desirable to utilize the finite element method to analyze the excavation-induced ground movements and building responses simultaneously and further conduct the evaluation of building damage caused by excavation.

In this chapter, the possible factors that may affect ground movements caused by excavation are first discussed. The background and development of numerically predicting excavation-induced wall deflection and ground movement are reviewed and introduced especially on the monitoring of excavation case histories. The development of small strain triaxial testing is also reviewed. Subsequently, a simplified soil model, the Modified Pseudo-Plasticity model, developed by the author and his colleagues is described in details. A series of small-strain triaxial tests conducted on the undisturbed Taipei silty clay are used to examine the performance of the Modified Pseudo-Plasticity model. Finally, two well-documented excavation case histories, one located in soft-to-medium clay and the other located in stiff clay, are analyzed numerically using the Modified Pseudo-Plasticity model, and then the conclusions are drawn.

2. Possible factors affecting excavation-induced ground movements

The complicated excavation system may be affected by a large number of factors such as wall stiffness, wall length, ground conditions, groundwater, excavation geometry, construction sequences, strut stiffness, workmanship and so on. It is essential to investigate the factors that may affect the excavation-induced wall deflection and ground movements when studying the topic: "Finite element analysis of wall deflection and ground movements caused by braced excavations". As reported in the previous studies (e.g., Mana and Clough, 1981; O'Rourke, 1981; Wong and Broms, 1989; Hashash and Whittle, 1996; Kung et al.,

2007b), the wall deflection and ground movements are affected by many factors, which may be grouped into three major categories (Kung, 2009):

A) Inherent factors

1. Stratigraphy: such as soil strength, soil stiffness, stress history of soil, and groundwater conditions. In general, larger wall deflection could be induced for an excavation in soils with lower strength and stiffness.
2. Site environment: such as adjacent buildings and traffic conditions. High-rise buildings and heavy traffic adjacent to the excavation site may cause extra wall deflection.

B) Design-related factors

1. Properties of retaining system: including wall stiffness, strut stiffness, and wall length. Larger wall deflection may be expected when using low-stiffness wall.
2. Excavation geometry: including width and depth of excavation. Generally, the wall deflection is approximately proportioned to the excavation depth.
3. Strut prestress: the strut prestress is aimed at making good connection between the strut and the wall. However, the prestress might reduce the wall deflection.
4. Ground improvement: such as jet grouting method, deep mixing method, compaction grouting method, electro-osmosis method, and buttresses. The soil strength and stiffness could be strengthened by ground improvement, which may reduce wall deflection.

C) Construction-related factors

1. Construction methods: such as the top-down method and bottom-up method.
2. Over-excavation: Over-excavation prior to installation of strut may cause larger wall deflection.
3. Prior construction: such as the effect of trench excavation prior to the construction of the diaphragm wall.
4. Construction of concrete floor slab: the thermal shrinkage of the concrete floor slab may result in an increase in the wall deflection.
5. Duration of the construction sequence: the duration of the strut installation or the floor construction. For an excavation in clay, longer duration for installing the strut or constructing the floor slab may cause larger wall deflection due to the occurrence of consolidation or creep of clay.
6. Workmanship: poorer workmanship may cause higher wall deflection.

The reasonable predictions of wall deflection and ground movement may be obtained provided most of factors can be adequately considered in the process of finite element analysis of braced excavation.

3. Literature review on simulation of excavation-induced ground movements

3.1 Discrepancy in ground movement between observation and prediction

The finite element method has been extensively used in the deformation analysis of retention system and ground caused by excavation over the past decades. According to past studies, it is generally acknowledged that wall deflection is relatively easier to predict than the ground movement using the finite element method with the conventional soil constitutive model such as the Modified Cam-clay model (e.g., Atkinson, 1993). As such, with the conventional soil constitutive model, the FEM simulation of the ground movement is often not as accurate as that of the wall deflection. A significant discrepancy in the ground surface settlement between the in-situ measurements and numerical simulations has been

observed and reported. Figure 4 displays the comparison of wall deflection and ground surface settlement at the final excavation stage between observations and predictions. In this figure, the Taipei National Enterprise Center case was analyzed using various soil models, including the hyperbolic model, the Mohr-Coulomb model, and the Modified Cam-clay model. As shown in Fig. 4, the prediction of settlement distribution is not satisfactory, while the predicted wall deflection is considered relatively reasonable. Specifically, the concave distribution of settlement at distances from 0 to 25m predicted using each of the three soil models is significantly less than the observed. The difference in the value and location of maximum settlement between observation and prediction is evident. For the zone far away from the wall (the distance larger than 25m), the predicted settlement is greater than the observed. It should be noted that the wall deflection at various construction stages can be accurately predicted using the three soil models although only the wall deflection simulated at the final stage is shown herein.

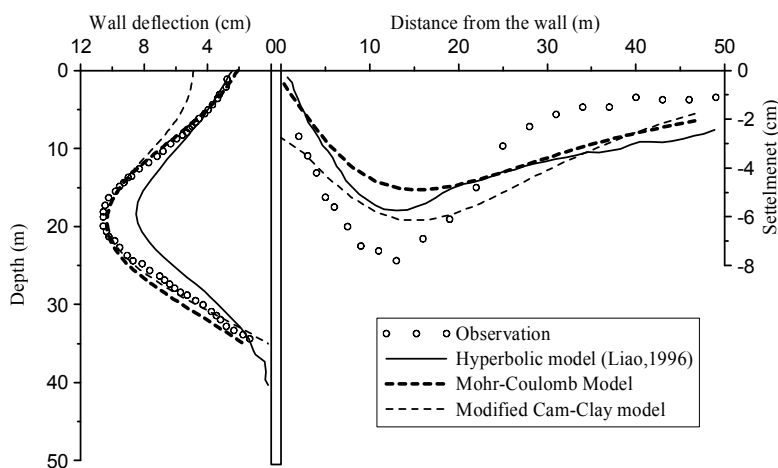


Fig. 4. Performance of various soil models on the simulation of wall deflection and ground surface settlement on the Taipei National Enterprise Center case

The reader may think why the accuracy of predicting the settlement distribution is that important. Essentially, the inaccurate predictions of ground surface settlement by the finite element method could cause the result that the prevention of building damage near the excavation can not be effectively reached. Figure 5 schematically illustrates the effect of discrepancy in the settlement distribution on the evaluation of damage potential of buildings between actual observations and finite element predictions. For example, if a building is assumed to have no rigidity, finite element predictions tend to underestimate the angular distortion, β , of the building (see Fig. 5) and thus underestimate the damage level of building caused by the excavation-induced settlement.

$$\beta_{prediction} = \delta_{ab} / L \square \delta_{a'b'} / L = \beta_{observation} \quad (1)$$

where δ represents the differential settlement between adjacent footings and L represents the distance between adjacent footings.

As shown in Fig. 5, use of the surface settlement distribution analyzed by finite element method could lead to improper building protection, which might cause damage to adjacent buildings although the prediction of wall deflection is satisfactory. It would be desirable to further study how to improve the accuracy of numerically predicting the ground surface settlement.

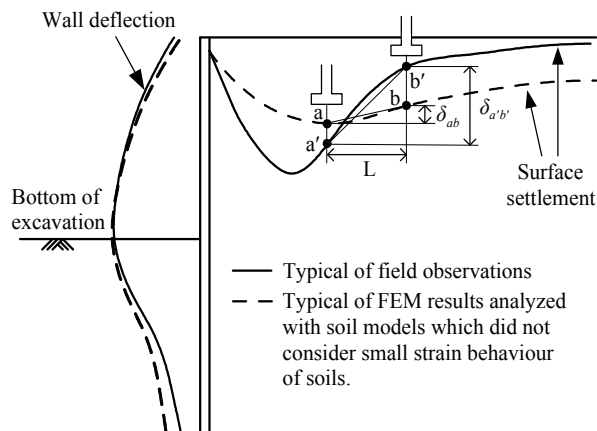


Fig. 5. The inaccurate evaluation of differential settlement of building based on finite element predictions (Kung et al., 2009)

In this section, the in-situ ground observations are utilized to explore the possible factors causing the inaccurate prediction of settlement distribution. Figure 6 shows the excavation-induced shear strain of ground behind the retaining wall observed in the Taipei National Enterprise Center excavation case. The induced shear strain of ground increases with increase of excavation depth. The maximum shear strain of soils at the final excavation depth of 19.7 m is approximately equal to 0.6%. The shear strain of ground far away from the wall is generally less than 0.1%. The stress-strain behaviour of soils at strains less than 0.6 % (or conservatively 1%) for a routine excavation have a dominate effect on the excavation-induced deformation. That is, the capability of the soil model in describing the stress-strain-strength characteristics of soil at small strain levels must be considered when the soil-structure interaction problem is analyzed. Otherwise, the accuracy of predicting excavation-induced ground movements would be significantly reduced.

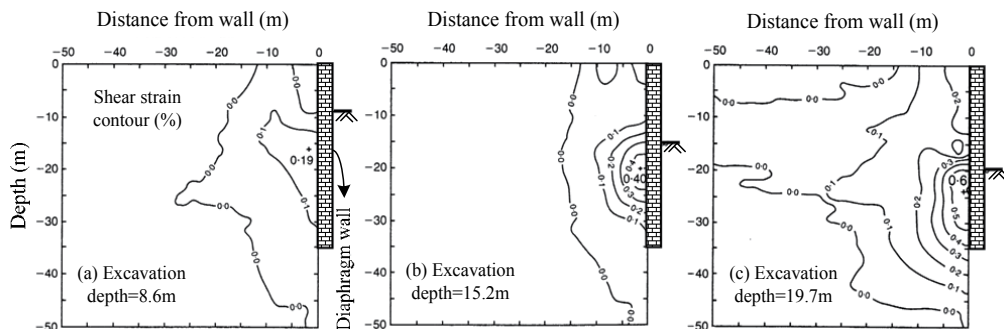


Fig. 6. Excavation-induced shear strain of ground observed in the TNEC case

3.2 Small-strain triaxial testing

The Bishop and Wesley triaxial cell is extensively used in the conventional triaxial test to measure the stress-strain-strength characteristics of soil. The triaxial test is a good method to investigate the soil behaviour, but some studies indicated that results of conventional triaxial tests are reliable only when the measurements of strain are approximately larger than 0.1% (e.g., Atkinson, 1993). Figure 7 displays the typical variation of soil stiffness with the strain from $10^{-4}\%$ to larger than 10%. The degradation of soil stiffness especially at the small strain level is significant. As such, it is believed that the capability of soil model in describing the stress-strain-strength characteristics of soil at a wide range of strain (e.g., 10^{-5} to 10^{-2}) would significantly affect the accuracy in the prediction of excavation-induced ground movements. In other words, the capability of a test in measuring the stress-strain-strength characteristics of soil at a wide range of strain plays a crucial role in the prediction of excavation-induced ground movements. Jardine et al. (1984) indicated that the method of measuring soil deformation in the conventional triaxial test using an LVDT installed outside the triaxial cell was inadequate. Also, many studies (e.g., Burland, 1989; Simpson, 1993; Kung 2003) indicated that the stiffness of soil measured in the conventional triaxial test would be significantly smaller than that measured by in-situ tests or triaxial tests equipped with transducers that can measure very small deformation of soil. Thus, a method to directly measure small deformation of soil locally on the sample was proposed. To date, several kinds of small-strain instruments have been developed (e.g., Clayton and Khatrush, 1986; Goto et al., 1991).

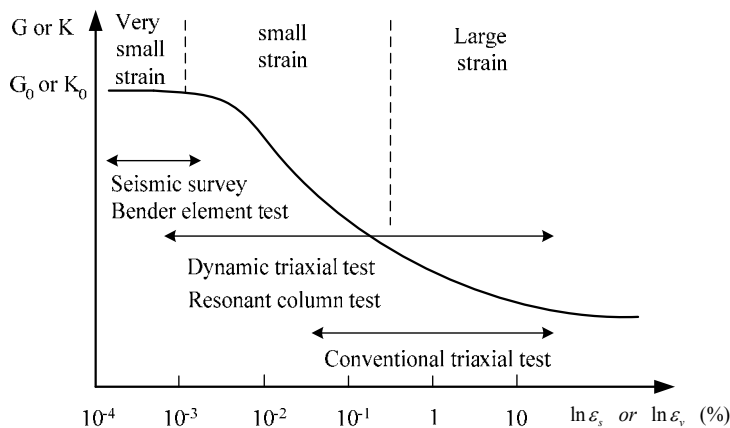


Fig. 7. Variation of soil stiffness with a wide range of strain

Over the past decades, few advanced instruments/transducers were developed for measuring the characteristics of soil at small strain levels (see Fig. 8 and Fig. 9). Figure 8 shows the appearance of bender element and the installation in the top cap and the base pedestal. The bender element composed of two piezoelectric-ceramic chips combined together can be used to measure the shear wave velocity through the sample during the consolidation and shearing stages of triaxial tests. When conducting a bender element test, a function generator is generally used to provide the excitation voltage to one of the bender elements (called the transmitter). The excitation voltage would cause the element to vibrate

and bend, so that a shear pulse is sent through the sample and received by another bender element, i.e. the receiver.

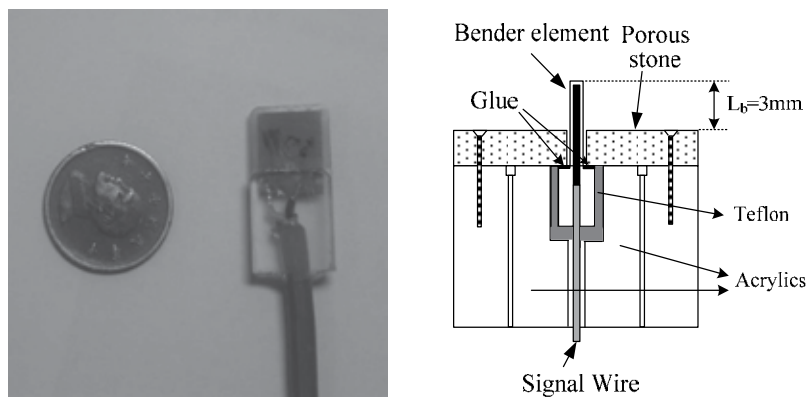


Fig. 8. Appearance and installation of the bender element (Kung, 2007)

Figure 9 exhibits four advanced small strain transducers, including (a) Eletrolevel gauge (Jardine, et al., 1984), (b) Hall effect transducer (Clayton and Khatrush, 1986), (c) Local displacement transducer (Goto, et al., 1991), and (d) Proximity transducer. In addition, the small, light linear variable differential transformer (LVDT) is also available and its installation method is similar to the proximity transducer (Fig. 9d). It should be noted that the principle of measuring the local deformation on a sample with each of the four instruments is not the same. The reader can be referred to references for additional information. It should be noted that some errors may be induced as measuring the stress-strain behavior of soils at small strain using the conventional triaxial apparatus (see Fig. 10). The main errors include manmade errors during preparation of a sample, low precision of instruments, inaccurate measurement methods, and improper arrangement of the conventional triaxial apparatus. Specifically, six kinds of errors may be induced, including (1) rotation of the top cap (2) measurement of loading (3) measurement of deformation (4) environmental temperature (5) disturbance during the preparation of a sample (6) tilting and bedding errors of a sample. Errors No. (1) and (6) are schematically shown in Fig. 11. To date, we can easily purchase the small strain instruments such as the hall effect transducer, the proximity transducer, or the small LVDT, and install them in the existing triaxial apparatus to locally measure the deformation of soils at small strain levels during the triaxial tests. However, it would make a futile effort provided we only substitute a small strain transducer for the original LVDT in a conventional triaxial testing system due to other errors may not be eliminated and the actual behaviour of soil may not be accurately measured (see Fig. 10).

Simply speaking, the force and deformation of a sample should be directly measured on or close to the sample. Use of the local strain transducer and the submerible load cell is a necessity. Obviously, the triaxial cell should be modified to avoid the rotation. Also, the temperature of water during the triaxial test must be measured for calibrating the recorded data. The procedure of preparation and installation of a sample required additional care to prevent the errors (5) and (6). Figure 12 shows an example of the small-strain triaxial testing system, which was developed by the author (Kung 2007).

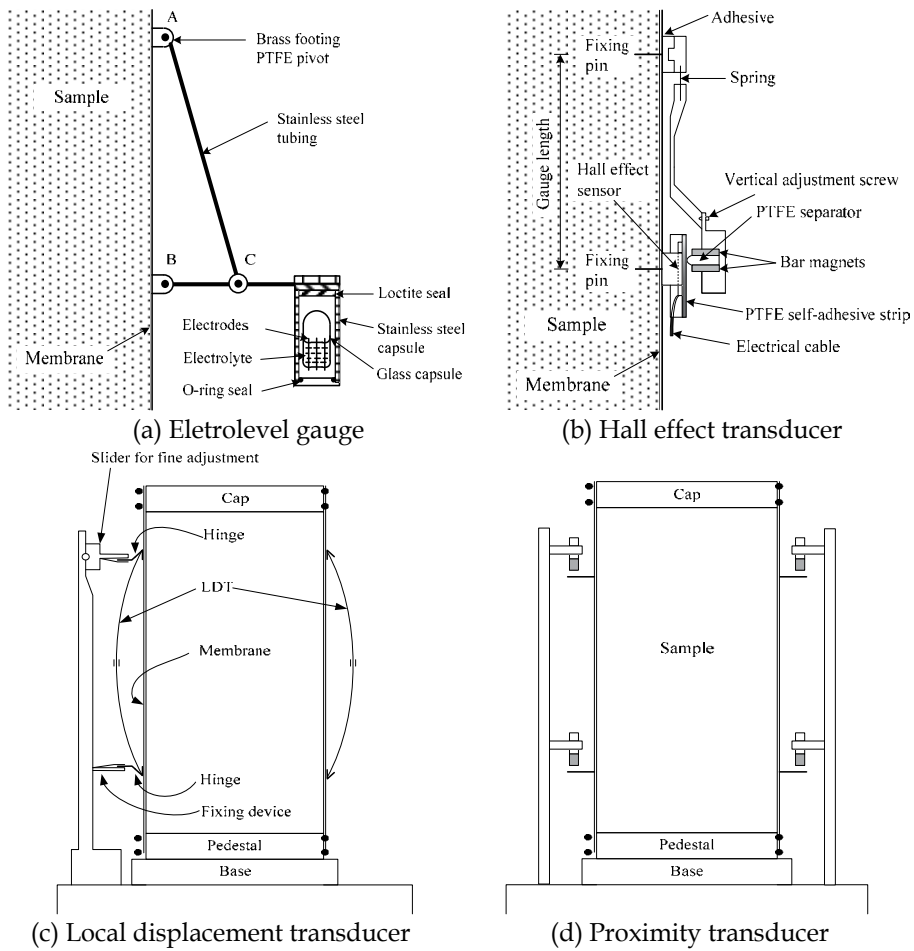


Fig. 9. The installation of various small strain transducers on the sample (Kung, 2007)

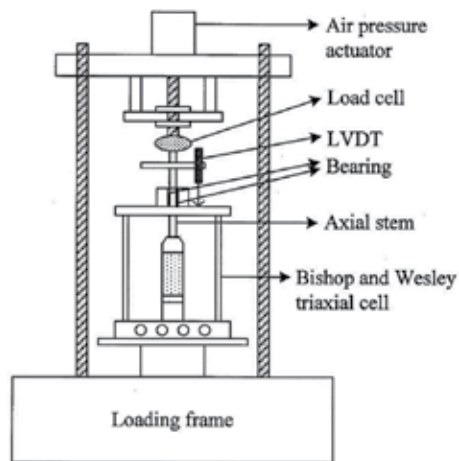
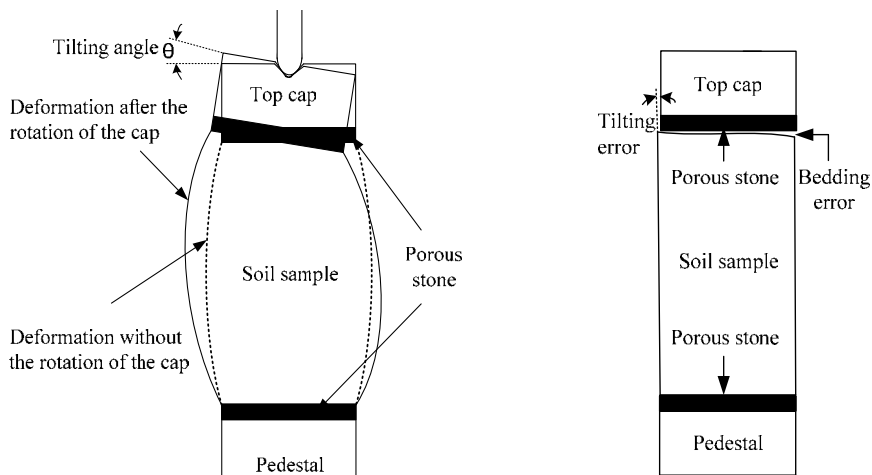


Fig. 10. Layout of the conventional triaxial apparatus (Kung, 2007)



(a) The uneven deformation of a sample due to rotation of the top cap

(b) Tilting and bedding errors caused by set-up of a sample

Fig. 11. Various sources of errors that may be induced by the conventional triaxial apparatus

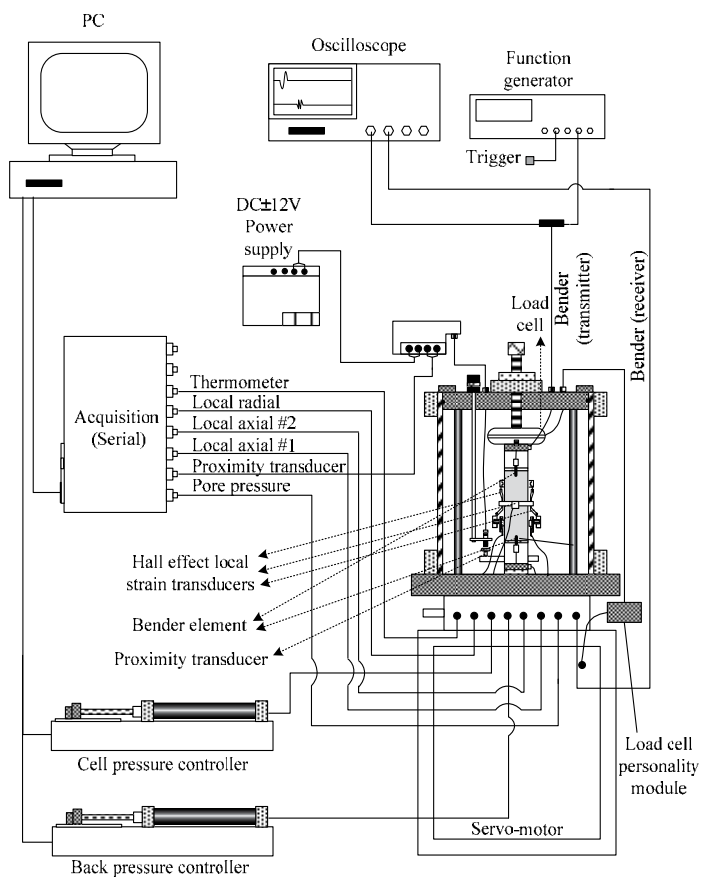


Fig. 12. Layout of the developed small-strain triaxial testing system (Kung, 2007)

4. Development of Modified Pseudo-Plasticity soil model

In recent years, several soil models have been improved to be capable of describing the stress-strain-strength characteristics of soil at small strain levels based primarily on the results of small strain triaxial tests. For instance, Stallebrass and Taylor (1997) incorporated the small strain nonlinear behaviour of soil within the initial yielding surface into the Modified Cam-clay model. The finite element solutions estimated by using this model can improve the discrepancy in the surface settlement between observation and prediction. Although the accuracy of excavation-induced ground movement predictions may be improved by employing a small-strain constitutive model, it could be difficult for engineers to employ such complicated constitutive models in a routine analysis of braced excavation. A simple soil model capable of capturing the small-strain behaviors of soils would be of great interest to practitioners. In this section, a simplified small strain soil model, called the Modified Pseudo-Plasticity model, developed by the author and his colleague are introduced herein and examined.

4.1 Modified pseudo plasticity model

For the Modified Pseudo-Plasticity model, the stress-strain relationship is expressed as follows:

$$\frac{\sigma_1 - \sigma_3}{s_{u\beta}} = \frac{\varepsilon}{\frac{s_{u\beta}}{E_i} + \frac{R_f \varepsilon}{2}} \quad (2)$$

where E_i is the initial Young's modulus; $s_{u\beta}$ is the undrained shear strength in the direction of the principal stress axis with an angle of β under the plane-strain condition; R_f is the failure ratio $(= (\sigma_1 - \sigma_3)_f / (\sigma_1 - \sigma_3)_{ult})$, where $(\sigma_1 - \sigma_3)_f$ is the deviator stress at failure and $(\sigma_1 - \sigma_3)_{ult}$ is the ultimate deviator stress. The term $s_{u\beta}$ is determined by (see Fig. 13):

$$s_{u\beta} = \left(\sqrt{R^2 - (q_c \sin 2\beta)^2} + q_c \cos 2\beta \right) s_{uc} \quad (3)$$

where $q_c = (1 - K_s)/2$ and $R = [(3K_s + 1)(K_s + 3)/12]^{1/2}$, in which K_s is the ratio of undrained shear strength ($K_s = s_{ue}/s_{uc}$); s_{uc} and s_{ue} are the undrained shear strengths in the triaxial compression and extension tests, respectively; β represents the angle between the direction of the maximum principal stress axis and the vertical direction.

In an incremental analysis, the tangential Young's modulus E_t can be expressed as:

$$E_t = E_{ur} (1 - R_f SL)^2 \quad (4)$$

where E_{ur} is the unloading-reloading stiffness; SL is the stress level.

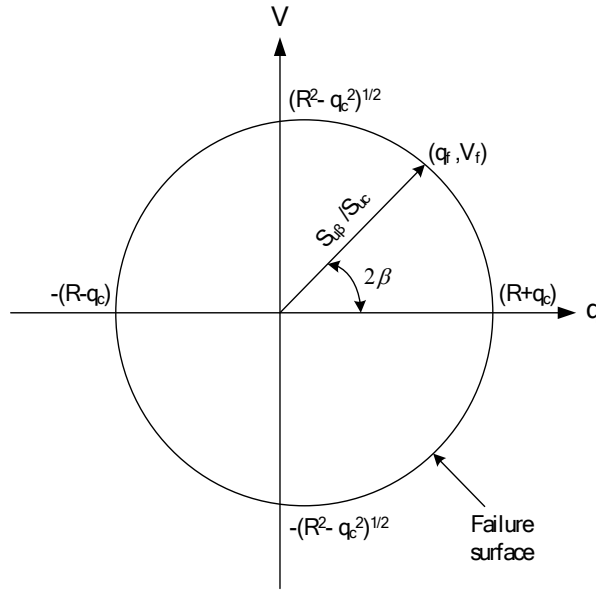


Fig. 13. Undrained shear strength in different directions of the maximum principal stress axis

In the Modified Pseudo-Plasticity model, E_{ur} can be determined by:

$$E_{ur} = E_i \quad \text{for } \varepsilon \leq 10^{-5}, \text{ and}$$

$$E_{ur} = E_i \left[1 - \frac{\varepsilon - 0.00001}{a + b(\varepsilon - 0.00001)} \right] \quad \text{for } \varepsilon > 10^{-5} \quad (5)$$

where ε is the axial strain; the coefficients a and b are parameters that control the degradation of stiffness.

The stress level criterion by Prevost (1979), in which the yield surface (S) and the failure surface (F) expressed as Equations 6 and 7 are used to differentiate the states of primary loading and unloading/reloading condition of soil (see Fig. 14), is incorporated into the Modified Pseudo-Plasticity model.

$$F = \frac{\left(\frac{q - q_o(1 - SL)}{SL} - q_c \right)^2 + \left(\frac{V}{SL} \right)^2}{R^2} - \Lambda^2 = 0 \quad (6)$$

$$S = \frac{[q - q_c SL - q_o(1 - SL)]^2 + V^2}{R^2} - \lambda^2 = 0 \quad (7)$$

$$SL = \frac{V}{V_f} = \frac{\left(\frac{[q - q_c SL - q_o(1 - SL)]^2 + V^2}{R^2} \right)^{\frac{1}{2}}}{\Lambda} \quad (8)$$

Failure surface: F
Yield surface: S

— SL=1.0
- - - SL=0.5
... SL=0.1

Fig. 14. Undrained shear strength in various directions of the maximum principal stress axis

The performance of the Modified Pseudo-Plasticity model on describing the stress-strain characteristics of clay is evaluated in this section. A number of the undisturbed soil samples were taken from the TNEC case (sites 1 and 2) located in Taipei to conduct the small strain triaxial tests for validating the Modified Pseudo-Plasticity model. Specifically, a total of five K_0 -consolidation undrained shearing test (CK₀AC test) as well as two multiple

unloading/reloading shearing tests were conducted. The Modified Pseudo-Plasticity model is then used to simulate the results of the CK_0AC tests and multiple unloading/reloading shearing tests. The samples used in these tests are classified as low plasticity, soft-to-medium silty clay and the basic properties obtained by the test are shown in Table 1. The liquid limit for the samples ranges from 33 to 40, and the plasticity index from 11 to 18. The water contents are in the range of 33%-35%.

Site	Soil types	γ_t	ω (%)	LL (%)	PI (%)
Site 1	CL	18.93	33-35	33-37	12-17
Site 2	CL	18.52	31-33	34-40	11-18

Table 1. Basic properties of soil tested

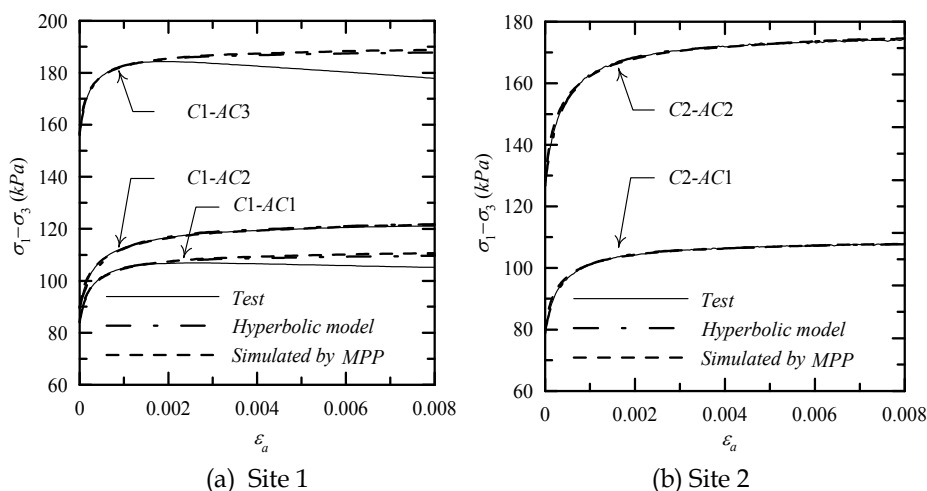


Fig. 15. Comparison of stress-strain curves between test results and simulations

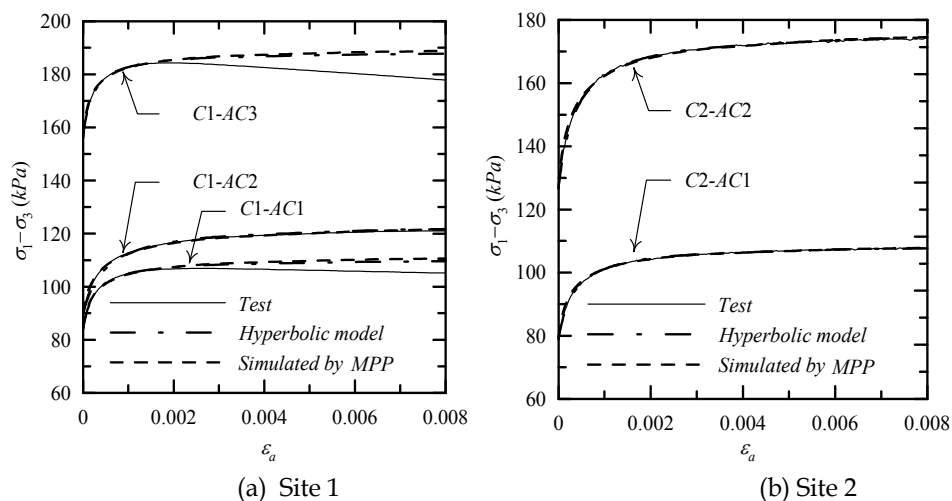


Fig. 16. Comparison of stiffness-strain relationship between test results and simulations

Figure 15 shows the comparison of stress-strain curves between test results and simulations by the Modified Pseudo-Plasticity model and the hyperbolic model. In general, the stress-strain curves at strains ranging from 0 to 0.008 can be reasonably simulated by the hyperbolic model and the Modified Pseudo-Plasticity model. It can be observed that the softening behaviour of soil (e.g., C1-AC1, C1-AC3) measured at strain larger than 0.002 cannot be appropriately simulated due to the fact that both the hyperbolic model and the Modified Pseudo-Plasticity model are only valid for describing the hardening behaviour of soil. The stiffness-strain relationship of each of five CK₀AC tests is compared and illustrated in Fig. 16. The initial normalized secant stiffness (E_{sec}/s_u) of clay measured by the small strain triaxial tests falls into the range of 2000 to 2200. The values of E_{sec}/s_u obtained by the hyperbolic model are significantly lower than the measured, while the estimation of E_{sec}/s_u by the Modified Pseudo-Plasticity model is satisfactorily consistent with the measured. Overall, the tendency of stiffness degradation of clay can be accurately represented by the Modified Pseudo-Plasticity model.

Finally, Fig. 17 compares the degradation of unloading/reloading modulus (E_e) between test results and simulations. Test results display that the undrained unloading/reloading modulus (E_e) of the Taipei clay decreases with the increase of the axial strain. The simulations shown in this figure exhibit that the variation of unloading/reloading stiffness of clay can be accurately captured by the Modified Pseudo-Plasticity. Compared with the hyperbolic model and the Pseudo-Plasticity model, the unloading/reloading modulus (E_e) assumed by the two soil models is a constant and would be inadequate to be used to predict the excavation-induced ground movements.

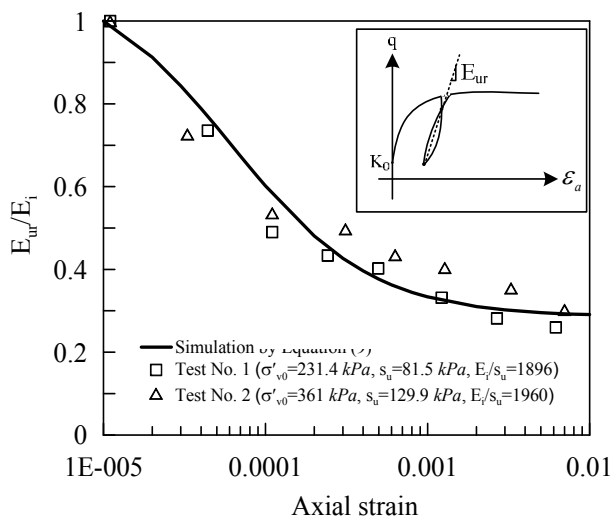


Fig. 17. Comparison of degradation of elastic modulus between test results and simulations

5. Analysis of braced excavation using the Modified Pseudo-Plasticity model

Two well-documented excavation case histories, the Taipei National Enterprise Center case (Ou et al., 1998) and the Post Office Square Garage case (Whittle et al., 1993), are selected to conduct the numerical analyses of excavation for examining the applicability of the Modified Pseudo-Plasticity model.

5.1 Taipei National Enterprise Center excavation case

As shown in Fig. 18, the shape of the Taipei National Enterprise Center case was slightly irregular. The width was 43 m, while the lengths of the southern and northern edges were 106 m and 61 m, respectively. A diaphragm wall, which was 0.9 m thick and 35 m deep, was used as the earth-retaining structure. The foundation of the Taipei National Enterprise Center case was constructed using the Top-down construction method, in which the wall was supported by 150 mm thick solid concrete floor slabs. The maximum excavation depth was 19.7 m. The excavation-induced wall deflection and ground movement were observed through five inclinometers (WI is installed in the wall; SI-1 to SI-4 are installed in the soil), three extensometers and a number of settlement points along the main observation section. Three pairs of inclinometer casings, SI-1, SI-2 and SI-3, and rod-type multipoint extensometers were installed to measure the vertical and horizontal deformation of soil simultaneously.

The Taipei National Enterprise Center case is located in the Taipei Basin, which is generally formed by a thick alluvium formation (the Sungshan Formation) lying above the Chingmei gravel Formation. The thickness of the Sungshan Formation is around 40 to 50 m. Essentially, the Sungshan Formation has six alternating silty sand (SM) and silty clay (CL) layers and mainly consists of low-plasticity and slightly over-consolidated soft to medium clay. Typical soil properties of the Sungshan Formation are shown in Table 2. In this case, the soft to medium clay at depths from 8 m to 33 m has the predominant effect in the excavation-induced deformation behaviour. The Chingmei gravel Formation can be found at the depth of 46 m. The depth of ground water table is around 2 m.

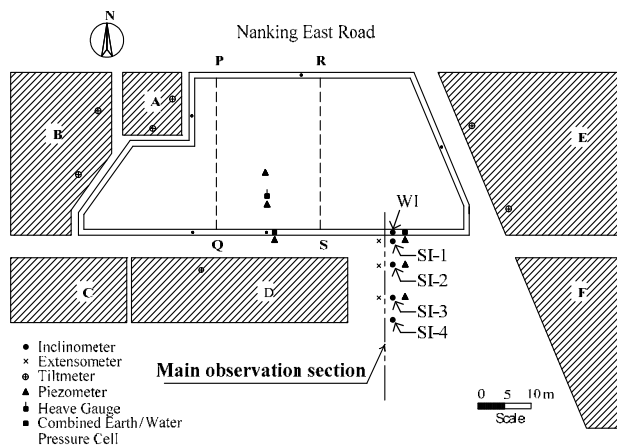


Fig. 18. Plan view of the TNEC case and the instrumentation plan

Formation	Layer	Depth (m)	Soil Type	SPT-N (Blows/ft)	ω_n (%)	LL	PI	ϕ' (°)
Sungshan	VI	0-5.6	CL	3	32	34	23	33
	V	5.6-8.0	SM	11	25	-	-	31
	IV	8.0-33.0	CL	3-10	25-40	29-39	9-19	30
	III	33.0-35.0	SM	20	24	-	-	31
	II	35.0-37.5	CL	14	28	33	21	32
	I	37.5-46.0	SM	30	30	-	-	-
Chingmei		>46.0	GP	>100	-	-	-	-

Table 2. Basic soil properties at the TNEC excavation site.

The considerations and analytical procedures in the analysis of the TNEC case history are described in detail as follows.

(1) Element types

The 8-node rectangular isoparametric quadrilateral element (Q8) was used for soil elements and diaphragm wall elements. The strut during excavation, either steel member or concrete slab, is normally subjected to axial force, and therefore the bar element was used to simulate the behavior of the strut or concrete slab.

(2) Modeling of soil and structure

The diaphragm wall was assumed to behave as a linearly-elastic material, for which both Young's modulus and Poisson's ratio were assumed constant. The clayey soil and sandy soil were assumed to behave as elasto-plastic materials as described by the MPP Model and hyperbolic model, respectively. The undrained analysis for the clayey layers and drained analysis for the sandy layers were employed in the finite element analyses of the TNEC case.

(3) Initial conditions

The effective horizontal stress is equal to the effective vertical stress multiplied by the coefficient of the at-rest lateral earth pressure (K_0) at initial conditions. In the TNEC case, K_0 was determined from the triaxial tests on the undisturbed Taipei clay. The pore water pressure, which was slightly lower than the hydrostatic pressure, was determined from the measurement of the in-situ pore water pressure using piezometers. The total stresses are equal to the sum of effective stress and pore water pressure.

(4) Determination of soil parameters

The MPP model for clay layers and hyperbolic model for sand layers are employed in the analysis. The determination of parameters for hyperbolic model is referred to the original definitions (Duncan and Chang, 1970). As mentioned previously, six soil parameters, E_i / s_u , s_u / σ'_v , R_f , a , b and K_s , are required to define the MPP model. The values of E_i / s_u , s_u / σ'_v and K_s were directly determined by the small-strain triaxial tests on the undisturbed Taipei clay sampled from the TNEC case using the local strain instruments and bender element tests simultaneously. Test results showed that E_i / s_u fell in the range from 1600 to 2500 and s_u / σ'_v varied from 0.30 to 0.35. According to Kung (2003), K_s for the Taipei clay is approximately equal to 0.75. For the term R_f , the value of $R_f = 0.9$ is considered adequate to simulate the stress-strain characteristics in the analysis. The results of the multiple unloading/reloading tests can be adequately represented by Equation 5 with $a = 0.0001$ and $b = 1.4$. The values of the six soil parameters used in the analysis are listed in Table 3.

Depth (m)	γ_t (kN/m ³)	K_0	E_i / s_u	s_{uc} / σ'_v	R_t	a	b	v	K_s
0-5.6	18.3	1.0	2100	0.32	0.9	0.0001	1.4	0.499	0.75
8-33	18.9	0.51	2100	0.32	0.9	0.0001	1.4	0.499	0.75
35-37.5	18.2	0.51	2100	0.34	0.9	0.0001	1.4	0.499	0.75

(a) Parameters of clayey layers (MPP model)

Depth (m)	γ (kN/m ³)	K_0	c' (kPa)	ϕ' (°)	R_f	$K = K_{ur}$	n	ν
5.6-8	18.9	0.49	0	31	0.9	750	0.5	0.3
33-35	19.6	0.49	0	31	0.9	2500	0.5	0.3
37.5-46	19.6	0.47	0	32	0.9	2500	0.5	0.3

(b) Parameters of sandy layers (Hyperbolic model)

Table 3. Soil parameters used in finite element analyses of the TNEC case

(5) Determination of structural parameters

The nominal Young's modulus of diaphragm wall (E_c) can be calculated by:

$$E_c = 4700\sqrt{f'_c} \quad (9)$$

where f'_c is the compressive strength of concrete (MPa).

For the FEM analysis of braced excavation, the nominal E_c is reduced to account for the effect of underwater construction of the diaphragm wall. In this study, 80% of nominal E_c is taken to conduct the FEM analysis. In the TNEC case, f'_c is equal to 27.44 MPa. The stiffness of struts or floor slabs, k , is determined by:

$$k = EA/LS \quad (10)$$

where E is Young's modulus of steel or concrete; A is the cross-section area; L is the length; S is the horizontal span. The stiffness of struts and floor slabs adopted are shown in Table 4. Figure 19(a) shows the comparison of the wall deflection between field observations and FEM predictions. The cantilever-type wall deflection at first and second stages can be accurately estimated. After the construction of concrete floor slabs, the deep-inward movements of wall deflection were induced at subsequent stages. The calculated maximum wall deflections are very close to the observations at stages 3 to 7, while the locations where the maximum wall deflection occurred can be accurately estimated except stages 6 and 7, where the estimated position is slightly deeper than the observations. The calculated maximum wall deflection is 109 mm, which is practically identical to the measured.

Stage No.	TNEC case			POSG case		
	H_e (m)	H_p (m)	k (kN/m/m)	H_e (m)	H_p (m)	k (kN/m/m)
1	2.8	N/A	N/A	2.8	-	-
2	4.9	2.0	8240	6.1	0.4	482623
3	8.6	3.5 & 0	125568	8.3	3.1	168918
4	11.8	7.1	125568	11.6	6.4	168918
5	15.2	10.3	125568	14.5	9.5	168918
6	17.3	13.7	125568	17.1	12.5	168918
7	19.7	16.5	24035	20.1	15.6	180984
8	-	-	-	23.2	18.6	180984

Note: H_e is the excavation depth; H_p is the depth where the strut is installed; k denotes the stiffness of strut and floor slab

Table 4. Propping arrangements for the excavation case histories and stiffness of struts and floor slab used in FEM analyses

Figure 19(b) shows the comparison of ground surface settlement. The observations reveal that the concave shape of surface settlement was induced mainly at distances of 0 to 25 m away from the wall. The maximum surface settlement after the completion of the final excavation stage (stage 7, excavation depth = 19.7 m) is around 74 mm and the location where the maximum surface settlement occurred is 13 m away from the wall. The results show that the trend of the settlement profile is fairly accurately estimated.

The predicted surface settlement at the range of 0 to 25 m away from the wall compare well to the observations, but the predictions of the surface settlement at the range of 25 to 40 m away from the wall are slightly larger than the observations. Generally, predictions of the maximum surface settlement at each stage are considered satisfactory except that at stages 5, 6, and 7, the maximum surface settlement is slightly underestimated.

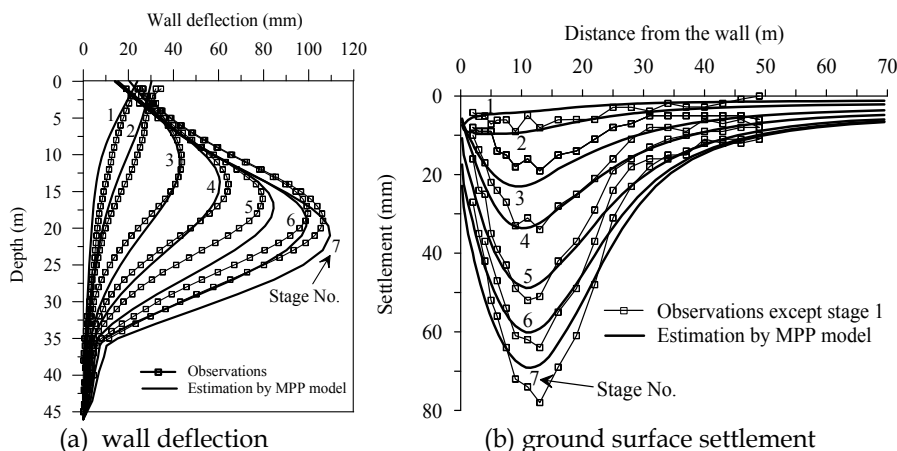


Fig. 19. Comparison of wall deflection and ground surface settlement

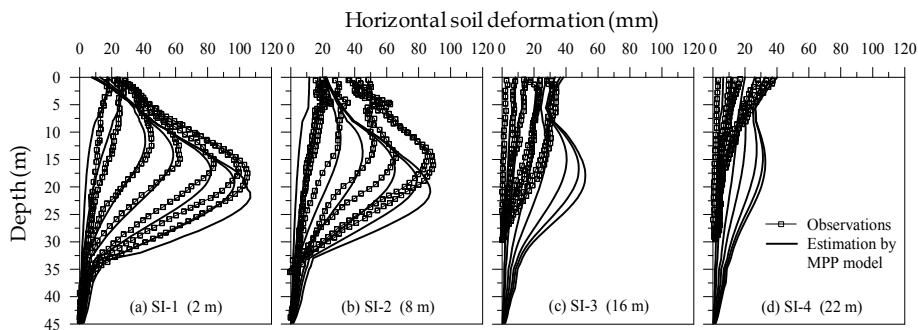


Fig. 20. Comparison of the horizontal soil deformation in the TNEC case

Figure 20 shows the comparison of the horizontal soil deformation. The results show that the computed maximum horizontal soil deformation and profiles along the SI-1 and SI-2 sections are generally close to the observations, although the difference is observed for the location where the maximum horizontal soil deformation occurred at the final two stages. However, the soil deformation along the SI-3 and SI-4 sections are accurately predicted at depths smaller than 10 m and over-predicted at depths larger than 10 m. Moreover,

compared with predictions of the vertical settlement, both the vertical and horizontal soil deformation were overestimated by the MPP model in the distance range of 25 to 40 m. In view of the difficulty of obtaining satisfactory predictions of all three responses (wall deflection, surface settlement, and horizontal soil deformation) simultaneously with a finite element analysis, the results obtained using the Modified Pseudo-Plasticity soil model are considered satisfactory.

5.2 Post Office Square Garage case

The observations during construction in the Post Office Square Garage (POSG) excavation case history were performed by Whittle et al. (1993). As shown in Fig. 21(a), this case occupies a plan area of 6880 m² (approximately 116×61 m) in the heart of financial district of Boston. Existing buildings up to 40 stories tall are located adjacent to the site. The diaphragm wall (25.6 m deep and 0.9 m thick) extending down into the bedrock was used as the permanent lateral earth pressure support. The POSG case was performed by the Top-down construction method. The maximum excavation depth was 23.2 m. The detailed construction sequences, including the excavation depth and the depth where the floor slab was constructed, are listed in Table 4. .

Figures 21(b) and 21(c) show the plan view of inclinometers and surface settlement points installed around the site, respectively. There are 13 inclinometers cast within the wall and 11 inclinometers located either in front of adjacent buildings or in close proximity to the diaphragm wall. Besides, Whittle et al. (1993) also installed six multiple position borehole extensometers to measure the relative vertical displacement of the clay, till and rock and 5 observation wells and 30 piezometers to measure the ground-water and piezometric level.

Based on an averaged profile of subsurface stratigraphy interpreted from a series of 15 borings conducted at the site, the stratigraphy mostly consists of the fill layer, clay, sand, till and bedrock. The fill layer comprises a heterogeneous mixture of sand, sandy gravel and construction debris. Underlying the fill is a deposit of low plasticity ($I_p=20-30\%$), moderately sensitive ($s_t=3-6$) clay containing numerous lenses of sand layers. The results of oedometer tests show that the clay has an in-situ over-consolidation ratio decreasing with depth in the deposit and ranging from OCR = 2-6. The soil deposits overlying the bedrock are classified as glacial till comprising a very heterogeneous mixture of particles, ranging from silt-size to cobbles and boulders. Finally, the bedrock is a moderately to severely weathered argillite deposit containing discontinuous layers of sandstone and quartzite.

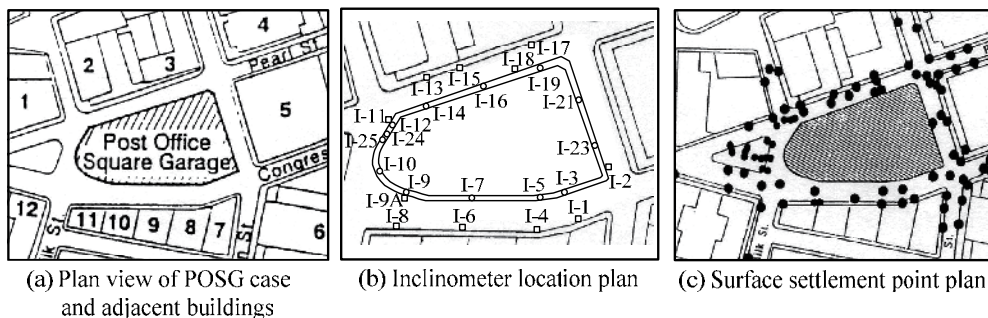


Fig. 21. Plan view of the POSG case and the instrumentation plan (after Whittle et al., 1993)

The values of the six soil parameters used in the analysis of the POSG case are listed in Table 5. According to the results of FEM analyses on the POSG case history using the Modified Pseudo-Plasticity soil model, the results of simulations of the wall deflection, ground surface settlement, and horizontal soil deformation behind the wall are extracted to compare with the observations. Before comparing the analysis results with observations, it is necessary to identify if the two-dimensional FEM analyses conducted under the plane strain condition are comparable to the complicated three-dimensional excavation observations. As shown in Fig. 21(b), a total of 13 inclinometers were installed in the wall to observe the wall deflection during excavation. Considering the arrangement of inclinometers and plane-strain condition, this study carried out the FEM analysis with a cross-section perpendicular to the long side and compared results of analysis with measurements of I-5, I-7 and I-16 because those inclinometers are located at the positions close to the central zone of diaphragm wall. According to the PSR concept proposed by Ou et al. (1996), the locations of I-5, I-7 and I-16 fall into the range of plane strain condition in light of values of PSR very close to 1.0. Therefore, the two-dimensional FEM analysis results of the POSG case would be comparable to the measurements of I-5, I-7 and I-16.

Figure 22 shows the comparison of wall deflection between the estimations by the MPP model and measurements of I-5, I-7 and I-16. At the first stage, the cantilever-type wall deflections can be appropriately simulated and the estimations are very close to the measurements of I-7 but slightly less than those of I-5 and I-16. At the second stage, the estimated wall deflection behaves the deep-inward behavior due to the fact that the high-stiffness concrete floor slab at depth of 0.4 m was constructed prior to the second-stage excavation. However, such cantilever-type wall deflection analyzed doesn't agree with the measurements of I-5, I-7, and I-16. The maximum wall deflection estimated at the second stage is practical equal to that observed by I-5 and I-7, but significantly smaller than that observed by I-16. Also, the depth where the maximum wall deflection occurred at the second stage ($\approx 7\text{m}$) for the FEM simulations is obviously different from the observations, in which the maximum wall deflection is induced at the ground surface level. This difference in the wall deflection profile between estimations and observations may be caused by effect of long time required to construct the concrete floor slab at ground surface level and its thermal shrinkage. The deep-inward wall deflection was observed at stages 3 to 7. Essentially, the estimated wall deflection profiles, the maximum wall deflection, and the location where the maximum wall deflection occurred generally fall into the range of measurements of I-5, I-7, and I-16. Overall, the estimated wall deflections using the MPP soil model are satisfactory.

Figure 23 shows the comparison of ground surface settlement between estimations and observations. In this figure, all the surface settlement observations recorded around the site were collected to compare with the estimations since the POSG case is located in the heart of financial district of Boston and there is not enough space to measure the profile of ground surface settlement along a specific section perpendicular to the diaphragm wall.

Depth (m)	γ (kN/m ³)	K_0	E_t / s_u	s_{uc} / σ'_v	R_f	a	b	v	K_s
0-2.4	19.2	1.0	2100	0.70	0.9	0.00001	1.2	0.499	0.55
2.4-15.6	19.6	1.0	2100	0.70	0.9	0.00001	1.2	0.499	0.55

(a) For clayey layers

Depth (m)	γ (kN/m ³)	K_0	c' (kPa)	ϕ' ($^\circ$)	R_f	$K = K_{ur}$	n	ν
15.6-17.1	19.6	0.4	0	37	0.9	1500	0.5	0.3
17.1-23.2	20.4	0.5	0	43	0.9	1000	0.5	0.3
23.2-25.6	22.0	1.0	175	32	0.9	3000	0.5	0.3
25.6-45.6	22.0	1.0	300	32	0.9	4000	0.5	0.3

(b) For sandy layers

Table 5. Soil parameters used in FEM analyses of the POSG case

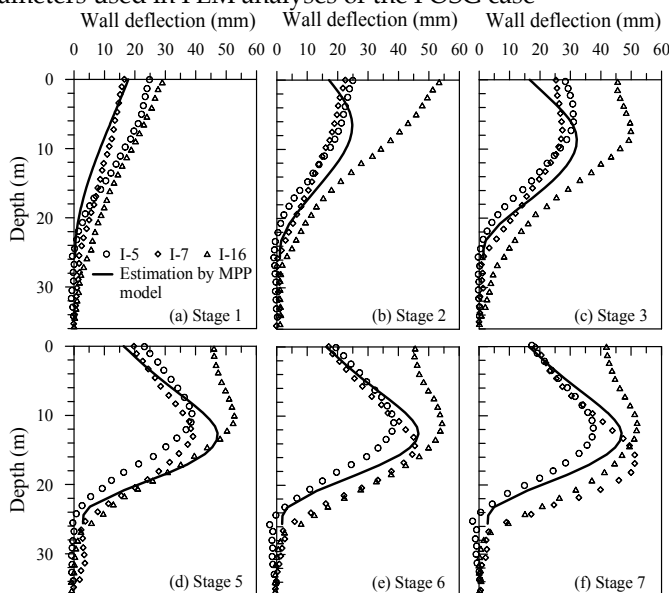


Fig. 22. Comparison of wall deflection (Kung, 2010)

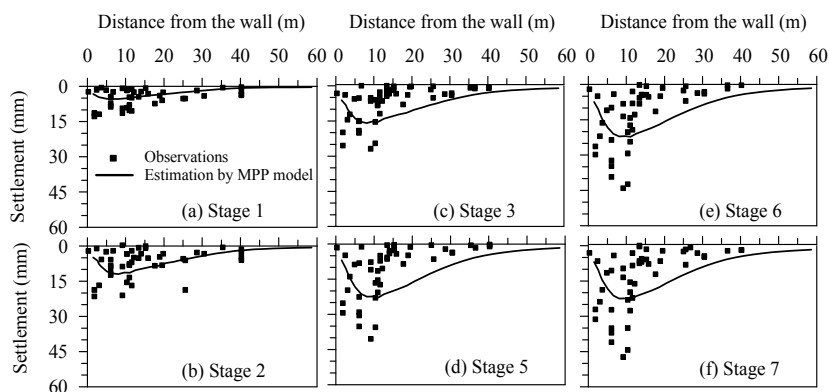


Fig. 23. Comparison of ground surface settlement (Kung, 2010)

Theoretically, the ground surface settlement distribution obtained by the two-dimensional FEM analysis under the plane strain condition should be an envelope; namely, all surface settlement observations should not exceed the envelope. Past studies indicated that for a practical excavation case, it is common to receive the settlement observations that are larger

than the estimated distribution analyzed by the two-dimensional FEM due to the fact that many factors such as the uncertainty of stratigraphy, variation of construction sequences, and traffic condition are not usually consistent with the design conditions. As shown in Fig. 22, amounts of part of surface settlement observations, especially in the range of 0-10 m behind the wall, are larger than that estimated by the MPP soil model, irrespective of excavation stages. Although the maximum surface settlement was underestimated at each stage, most of surface settlement measurements are smaller than the estimations, while the variation of settlement with the distance from the wall is consistent with the estimated settlement profile.

Figure 24 compares estimated horizontal soil deformations with in-situ data observed from inclinometers located at distance of 4, 9, and 14 m behind the wall. Specifically, the horizontal soil deformation recorded by I-2 and I-18 (4m behind the wall), I-13, I-15, and I-17 (9m behind the wall), and I-1 and I-8 (14m behind the wall) are extracted to compare with the estimations. The locations of those inclinometers can be referred to Fig. 21(b). For the condition of 4 m behind the wall, observations of I-2 and I-18 at stage 1 can be reasonably estimated (solid line). For stages 4 and 7, the estimated horizontal soil deformation would slightly overestimate the observations of I-18 but significantly overestimate those of I-2. The possible reason is due to the fact that I-2 is located at the corner and the deformation would be reduced by the corner effect. It may not be appropriate to directly compare the estimations with the observations of inclinometers, which are not located at the middle area of wall. Accordingly, this study employed PSR concept suggested by Ou et al. (1996) to further compare the horizontal soil deformation.

Briefly, PSR can be defined as the ratio of the maximum deflection in an arbitrary section to that computed under plane strain conditions (the same excavation width). For the scenario of 4 m behind the wall, the value of PSR is determined to be approximately equal to 0.8 and used to modify the FEM estimations. As shown in Fig. 24, the scaling estimations are closer to the observations of I-18 at later stages. The estimation of the maximum horizontal soil deformation at stage 7 is satisfactory. For the condition of 9m behind the wall, the horizontal soil deformation at stage 1 can be accurately estimated, while the maximum horizontal soil deformation at stage 7 can be accurately estimated but the estimated profile is not satisfactory.

For the condition of 14 m behind the wall, the horizontal soil deformation would be significantly overestimated because I-1 and I-8 are located at the zone closer to the corner. The value of PSR for I-1 and I-8 approximately equal to 0.2 is determined based on Ou et al. (1996). Then, the horizontal soil deformation profiles are modified with $PSR=0.2$ and the results are comparable to the observations. The corner effect on excavation behavior is significant and the additional attention should be paid when comparing the observations with the analysis results, which are obtained using two dimensional finite element analysis.

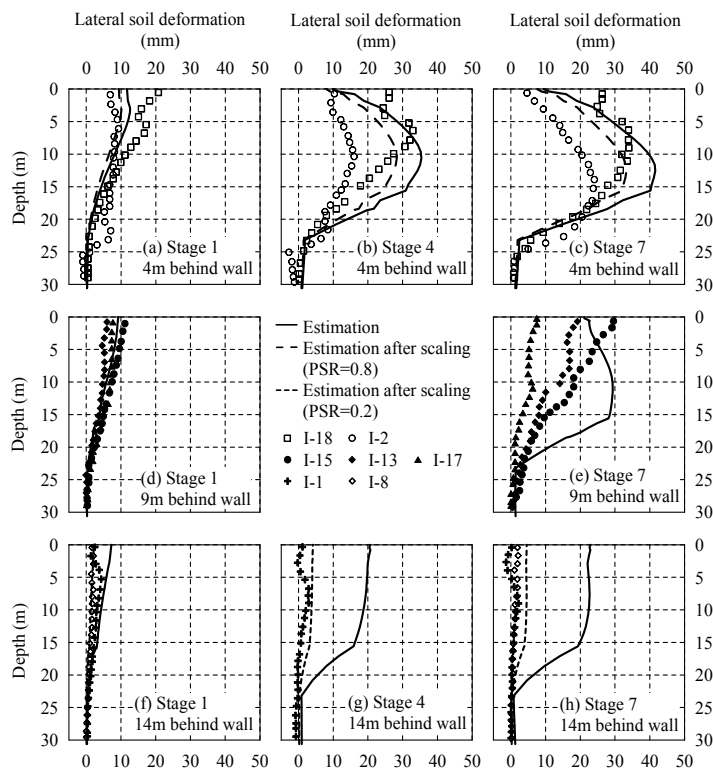


Fig. 24. Comparison of lateral soil deformations (Kung, 2010)

6. Conclusions

To accurately predict excavation-induced ground movements is a complicated but essential task in a routine excavation design for achieving the goal to prevent the damage to buildings adjacent to excavation. Use of numerical methods, such as the finite element method, to predict the ground movements caused by excavation is advantageous due to the stress and strain of the retention system and ground can be provided in the numerical analysis. The analysis results show that the capability of the soil model adopted in describing the stress-strain-strength of characteristics of soils at a wide range of strain, especially at small strain ranging from 10^{-5} to 10^{-2} , plays the crucial role in accurately predicting the excavation-induced ground movements. In addition, the engineer also has to realize the importance of small strain triaxial tests, which can be employed to be a basis for developing the above-mentioned small strain soil models and to measure the soil parameters of small strain soil models for deformation analysis of excavation. Indeed, it is not a simple work to perform such numerical analysis of excavation using small strain soil models but it would significantly benefit the excavation design. The Modified Pseudo-Plasticity model developed is merely one of qualified soil models. The engineer is strongly encouraged to study such numerical analysis of excavation using the small strain soil model and employ in the future design of excavation. Of course, use of small strain soil models to develop new simplified methods for the prediction of excavation-induced ground movements and building responses is desirable.

7. References

- Atkinson, J.H. (1993). *An Introduction to the Mechanics of Soils and Foundations:/ through critical state soil mechanics*, McGraw-Hill, ISBN:007707713X , London.
- Burland JB. (1989).Ninth Laurits Bjerrum memorial lecture: small is beautiful-the stiffness of soils at small strain. *Canadian Geotechnical Journal*, Vol. 26, 499–516, ISSN:1208-6010.
- Clayton, C.R.I. & Khatrush, S.A. (1986). A New Device for Measuring Local Axial Strains on Triaxial Specimens. *Geotechnique*, Vol. 36, 593-597, ISSN: 0016-8505.
- Clough, G.W. & O'Rourke, T.D. (1990). Construction-induced movements of in-situ walls. *Proceeding of Specialty Conference on Design and Performance of Earth Retaining Structure*, pp. 439-470, American Society of Civil Engineers, ISBN: 0872627616, Cornell University, June 1990, Ithaca, New York.
- Duncan, J.M. & Chang, C.Y. (1970). Nonlinear Analysis of Stress and Strain in Soils. *Journal of the Soil Mechanics and Foundations Division*, Vol. 96, No. 5, 637-659, ISSN: 1090-0241.
- Finno, R.J. & Harahap, I.S. (1991). Finite element analysis of HDR-4 excavation. *Journal of Geotechnical and Geoenvironmental Engineering*, Vol. 117, No. 10, 1590-1609, ISSN: 1090-0241.
- Goto, S.; Tatsuoka, F.; Shibuya, S.; Kim, Y.S. & Sato, T. (1991). A Simple Gauge for Local Small Strain Measurements in the Laboratory. *Soil and Foundations*, Vol. 31, No. 1, 169-180, ISSN: 0038-0806.
- Hashash, Y.M.A. & Whittle, A.J. (1996). Ground movement prediction for deep excavations in soft clay. *Journal of the Geotechnical Engineering*, Vol. 122, No. 6, 474–486, ISSN: 1090-0241.
- Hsieh, P.G. & Ou, C. Y. (1998). Shape of ground surface settlement profiles caused by excavation. *Canadian Geotechnical Journal*, Vol. 35, No. 6, 1004–1017, ISSN:1208-6010.
- Jardine, R.J.; Symes, M.J. & Burland, J.B. (1984). The measurement of Soil Stiffness in the Triaxial Apparatus. *Geotechnique*, Vol. 34, No. 3, 323-340, ISSN: 0016-8505.
- Kung, G.T.C. (2003). Surface settlement induced by excavation with consideration of small strain behavior of Taipei silty clay. PhD Dissertation, Department of Construction Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan.
- Kung, G.T.C. (2009). "Comparison of excavation-induced wall deflection using top-down and bottom-up construction methods in Taipei silty clay." *Computers and Geotechnics*, Vol. 36, No. 3, 373-385, ISSN: 0266-352X.
- Kung, G.T.C. (2010). Modeling small-strain nonlinearity of soils for numerical simulation of braced excavation in stiff clay. *Advances in Computer Science and Engineering*, Vol. 4, No. 1, 1-21, ISSN: 0973-6999.
- Kung, G.T.C.; Hsiao, E.C.L. & Juang, C.H. (2007a). Evaluation of a simplified small strain soil model for estimation of excavation-induced movements. *Canadian Geotechnical Journal*, Vol. 44, 726–736, ISSN:1208-6010.
- Kung, G.T.C.; Juang, C.H.; Hsiao, E.C.L. & Hashash, Y.M.A. (2007b). Simplified model for wall deflection and ground surface settlement caused by braced excavation in clays. *Geotechnical and Geoenvironmental Engineering*, Vol. 133, No. 6, 731-747, ISSN: 1090-0241.

- Kung, G.T.C.; Ou, C.Y. & Juang, C.H. (2009). Modeling small-strain behaviour of Taipei clays for finite element analysis of braced excavations. *Computers and Geotechnics*, Vol. 36, No. 1-2, 304-319, ISSN: 0266-352X.
- Liao, J.T. (1996). Performances of a Top Down Deep Excavation, Ph.D. Dissertation, Department of Construction Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan.
- Mana, A.I. & Clough, G.W. (1981). Prediction of movements for braced cut in clay. *Journal of the Geotechnical Engineering*, Vol. 107, No. GT8, 759-777, ISSN: 1090-0241.
- Ng, C.W.W. & Yan, W.M. (2000). A true three-dimensional numerical analysis of diaphragm walling. *Geotechnique*, Vol. 49, No. 6, 825-834, ISSN: 0016-8505.
- O'Rourke, T.D. (1981). Ground movements caused by braced excavations. *Journal of the Geotechnical Engineering*, Vol. 107, No. 6, 1159-1177, ISSN: 1090-0241.
- Ou, C.Y., Chiou, D.C. & Wu, T.S. (1996). Three-dimensional finite element analysis of deep excavations, *Journal of the Geotechnical Engineering*, Vol. 122, No. 5, 337-345, ISSN: 1090-0241.
- Ou, C.Y.; Liao, J.T. & Lin, H.D. (1998). Performance of Diaphragm Wall Constructed Using Top-down Method. *Journal of Geotechnical and Geoenvironmental Engineering*, Vol. 124, No. 9, 798-808, ISSN: 1090-0241.
- Prevost, J.H. (1979). Undrained shear tests on clay. *Journal of the Geotechnical Engineering*, Vol. 105, No. 1, pp. 49-64, ISSN: 1090-0241.
- Simpson, B. (1993). Development and application of a new soil model for prediction of ground movements, *Proceedings of the Wroth Memorial Symposium*, pp. 628-643, ISBN:0727719165, St Catherine's College, Oxford, July 1992, Oxford, London.
- Stallebrass, S.E. & Taylor, R.N. (1997). The development and evaluation of a constitutive model for the prediction of ground movements in overconsolidation clay. *Geotechnique*, Vol. 47, No. 2, 235-253, ISSN: 0016-8505.
- Whittle, A.J.; Hashash, Y.M.A. & Whitman, R.V. (1993). Analysis of deep excavation in Boston. *Journal of the Geotechnical Engineering*, Vol. 119, No. 1, 69-90, ISSN: 1090-0241.
- Wong, K.S. & Broms, B.B. (1989). Lateral wall deflections of braced excavation in clay. *Journal of the Geotechnical Engineering*, Vol. 115, No. 6, 853-870, ISSN: 1090-0241.

Vehicle-Bridge Dynamic Interaction Using Finite Element Modelling

Arturo González
University College Dublin
Ireland

1. Introduction

First investigations on the dynamic response of bridges due to moving loads were motivated by the collapse of the Chester railway bridge in the UK in the middle of the 19th century. This failure made evident the need to gain some insight on how bridges and vehicles interact, and derived into the first models of moving loads by Willis (1849) and Stokes (1849). These models consisted of a concentrated moving mass where the inertial forces of the underlying structure were ignored. The latter were introduced for simple problems of moving loads on beams in the first half of the 20th century (Jeffcott, 1929; Inglis, 1934; Timoshenko & Young, 1955). Although Vehicle-Bridge Interaction (VBI) problems were initially addressed by railway engineers, they rapidly attracted interest in highway engineering with the development of the road network and the need to accommodate an increasing demand for heavier and faster vehicle loads on bridges. In the 1920's, field tests carried out by an ASCE committee (1931) laid the basis for recommendations on dynamic allowance for traffic loading in bridge codes, and further testing continued in the 50's as part of the Ontario test programme (Wright & Green, 1963). However, site measurements are insufficient to cover all possible variations of those parameters affecting the bridge response, and VBI modelling offers a mean to extend the analysis to a wide range of scenarios (namely, the effect of road roughness or expansion joints, the effect of vehicle characteristics such as suspension, tyres, speed, axle spacing, weights, braking, or the effect of bridge structural form, dimensions and dynamic properties). A significant step forward took place in the 50's and 60's with the advent of computer technology. It is of particular relevance the work by Frýba (1972), who provides an extensive literature review on VBI and solutions to differential equations of motion of 1-D continuous beam bridge models when subjected to a constant or periodic force, mass and sprung vehicle models. At that time, VBI methods were focused on planar beam and vehicle models made of a limited number of degrees of freedom (DOFs). From the decade of the 70's, the increase in computer power has facilitated the use of numerical methods based on the Finite Element Method (FEM) and more realistic spatial models with a large number of DOFs. This chapter reports on the most widely used finite element techniques for modelling road vehicles and bridges, and for implementing the interaction between both.

The complexity of the mathematical models used to describe the dynamic response of a structure under the action of a moving load varies with the purpose of the investigation and the desired level of accuracy. 1-D models can be used in preliminary studies of the bridge vibration, although they are more suited to bridges where the offset of the vehicle path with respect to the bridge centreline is small compared to the ratio of bridge length to bridge width. Bridge models can be continuous or made of discretized finite elements. Vehicle models can consist of moving constant forces, masses or sprung masses. The simplest vehicle models are made of constant forces that ignore the interaction between vehicle and bridge, and thus, they give better results when the vehicle mass is negligible compared to the bridge mass. The mass models allow for inertial forces of the moving load, but they are unable to capture the influence of the road irregularities on the vehicle forces and subsequently on the bridge response. The sprung mass models allow for modelling frequency components of the vehicle and they vary in complexity depending on the assumptions adopted for representing the performance of tyres, suspensions, etc. Once the equations of motion of bridge and vehicle models have been established, they are combined together to guarantee equilibrium of forces and compatibility of displacements at the contact points. The fundamental problem in VBI modelling is that the contact points move with time and for each point in time, the displacements of the vehicle are influenced by the displacements of the bridge, which affect the vehicle forces applied to the bridge which in turn again alter the bridge displacements and interaction forces. This condition makes the two sets of equations of motion coupled, prevents the existence of a closed form solution (except for simple models, i.e., moving constant forces), and makes the use of numerical methods necessary. It is the aim of this chapter to show how to formulate and solve the VBI problem for a given road profile, and sophisticated vehicle and bridge FE models. The chapter is divided in four sections: the definition of the dynamic behaviour of the bridge model (Section 2), those equations describing the response of the vehicle model (Section 3), the road profile (Section 4), and the algorithm used to implement the interaction between vehicle, road and bridge models (Section 5).

2. The Bridge

The response of a discretized FE bridge model to a series of time-varying forces can be expressed by:

$$[M_b]\{\ddot{w}_b\} + [C_b]\{\dot{w}_b\} + [K_b]\{w_b\} = \{f_b\} \quad (1)$$

where $[M_b]$, $[C_b]$ and $[K_b]$ are global mass, damping and stiffness matrices of the model respectively, $\{w_b\}$, $\{\dot{w}_b\}$ and $\{\ddot{w}_b\}$ are the global vectors of nodal bridge displacements and rotations, their velocities and accelerations respectively, and $\{f_b\}$ is the global vector of interaction forces between the vehicle and the bridge acting on each bridge node at time t . In Equation (1), damping has been assumed to be viscous, i.e., proportional to the nodal velocities. Rayleigh damping is commonly used to model viscous damping and it is given by:

$$[C_b] = \alpha[M_b] + \beta[K_b] \quad (2)$$

where α and β are constants of proportionality. If ζ is assumed to be constant, α and β can be obtained by using the relationships $\alpha = 2\zeta\omega_1\omega_2/(\omega_1+\omega_2)$ and $\beta = 2\zeta/(\omega_1+\omega_2)$ where ω_1 and ω_2 are the first two natural frequencies of the bridge, although ζ can also be varied for each mode of vibration (Clough & Penzien, 1993). The damping mechanisms of a bridge may involve other phenomena such as friction, but they are typically ignored because the levels of damping of a bridge are small, and a somewhat more complex damping modelling would not change the outcome significantly. Cantieni (1983) tests 198 concrete bridges finding an average viscous damping ratio (ζ) of 1.3% (minimum of 0.3%), Billing (1984) reports on an average value of 2.2% (minimum of 0.8%) for 4 prestressed concrete bridges of spans between 8 and 42 m, and 1.3% (minimum of 0.4%) for 14 steel bridges with spans between 4 and 122 m, and Tilly (1986) gives values of 1.2% (minimum 0.3%) for 213 concrete bridges of span between 10 and 85 m, and 1.3% (minimum 0.9%) for 12 composite, steel-concrete bridges of span between 28 and 41 m (Green, 1993). Damping usually decreases as the bridge length increases, and it is smaller in straight bridges than in curved or skew bridges. Regarding the first natural frequency (in Hz) of the bridge, Cantieni (1983) finds a relationship with bridge span length L (in meters) given by $95.4L^{-0.933}$ based on 224 concrete bridges (205 of them prestressed). Nevertheless, the scatter is significant due to the variety of bridges, and when focusing the analysis on 100 standard bridges of similar characteristics (i.e., relatively straight), a regression analysis resulted into $90.6L^{-0.923}$. Tilly (1986) extended Cantieni's work to 874 bridges (mostly concrete) leading to a general expression of $82L^{-0.9}$. Heywood et al (2001) suggests a relationship $100/L$ for a preliminary estimation of the main frequency of the bridge, although there could be significant variations for shorter spans and singular structures going from $80/L$ to $120/L$ (i.e., timber and steel bridges exhibit smaller natural frequencies than reinforced or prestressed concrete bridges). The theoretical equation for the first natural frequency of a simply supported beam given by $f_1 = \pi\sqrt{EI/\mu}/(2L^2)$ where μ is mass per unit length and E is modulus of elasticity, has been found to be a good approximation for single span simply supported bridges (Barth & Wu, 2007). For bridges 15 m wide, span to depth ratio of 20 and assuming $E = 35 \times 10^9$ N/m², the theoretical equation of the beam leads to a frequency given by the relationship $85/L$ for solid slab decks made of inverted T beams ($L < 21$ m) and by $84.7L^{-0.942}$ (approximately $102/L$) for beam-and-slab sections ($17 \text{ m} < L < 43 \text{ m}$). Single spans with partially restrained boundary conditions or multi-span structures lead to higher first natural frequencies than the one obtained with the theoretical equation of the simply supported beam. In this case, Barth & Wu (2007) suggests multiplying the frequency obtained using the equation of the beam by a correction factor λ^2 that depends on the maximum span length, average section stiffness and number of spans.

2.1 Types of FE Models for Bridges

The size and values of $[M_b]$, $[C_b]$ and $[K_b]$ are going to depend on the type of elements employed in modelling the bridge deck. The coefficients of these matrixes are established using the FEM by: (a) applying the principal of virtual displacements to derive the elementary mass, damping and stiffness matrixes and then, assembling them into the global matrixes of the model, or (b) simply constructing the model based on the built-in code of a FE package such as ANSYS (Deng & Cai, 2010), LS-DYNA (Kwasniewski et al., 2006), NASTRAN (Baumgärtner, 1999; González et al., 2008a), or STAAD (Kirkegaard et al., 1997).

Given that a simple 1-D beam model is unable to accurately represent 2-D or 3-D bridge behaviour, the most common techniques for modelling bridge decks can be classified into: 2-D plate modelling (Fig. 1), 2-D grillage modelling (Fig. 2) and 3-D FE models. A 3-D FE model can be made of 3-D solid elements (Kwasniewski et al., 2006; Deng & Cai, 2010) or a combination of 1-D, 2-D and/or 3-D finite elements. The global matrixes of the 2-D grillage and plate bridge models are the result of assembling 1-D beam and 2-D plate elementary matrixes respectively. The number, location and properties of the elements employed in 2-D FE models are discussed here.

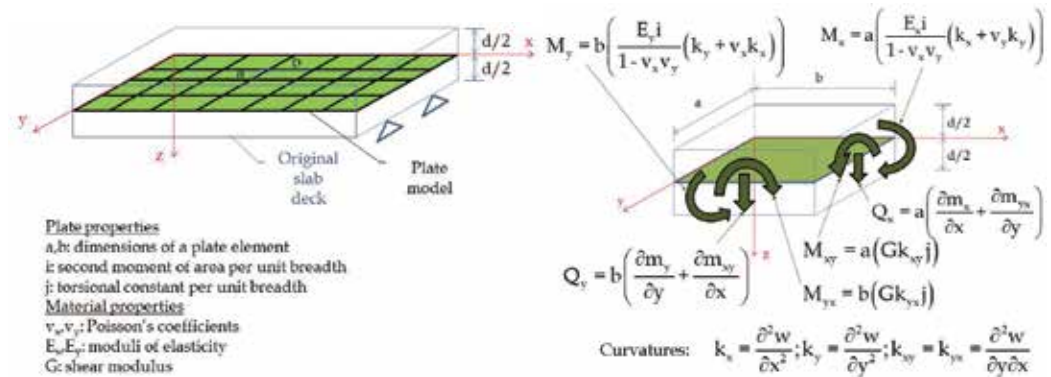


Fig. 1. Plate model of a bridge deck

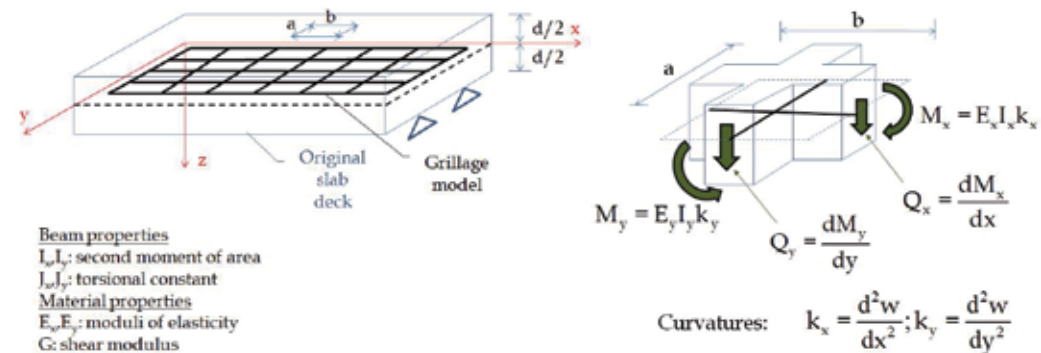


Fig. 2. Grillage model of a bridge deck

Plate bridge models have been used to investigate VBI by Olsson (1985), Kirkegaard et al (1997), Henchi et al (1998), Zhu & Law (2002), Cantero et al (2009) and González et al (2010). They are based on thin plate theory, which assumes the normal and shear strains in the 'z' direction (perpendicular to the plate plane) to be negligible. Using kinematic, constitutive and equilibrium equations, it is possible to derive the expressions for internal moment (M_x, M_y, M_{xy} are total moments acting on each face and m_x, m_y, m_{xy} are per unit breadth) and shear (Q_x, Q_y) shown in Fig. 1 as a function of nodal vertical displacement in the 'z' direction ($w=w(x,y)$), material (E_x, E_y, v_x, v_y, G) and section properties (a, b, i, j).

Grillage bridge models are also often found in VBI literature (Huang et al., 1992; Wang et al., 1996; Tan et al., 1998; Liu et al., 2002; Nassif & Liu, 2004). While a plate model of a bridge deck is a mesh made of 2-D plate elements, a grillage model of a bridge is a skeletal

structure consisting of a mesh of 1-D beams. The static equations that provide the bending moments (M_x, M_y) and shear (Q_x, Q_y) for a series of beams in the 'x' and 'y' directions are given in Fig. 2. A large investigation by West (1973) concluded that grillage modelling is very well suited for analysis of slab and beam-and-slab bridge decks and it has obvious advantages regarding simplicity, accuracy, computational time and ease to interpretate the results. It must be noted that there are a number of inaccuracies associated to grillage modelling (O'Brien & Keogh, 1999). I.e., the diagram of internal forces is discontinuous at the grillage nodes and it is difficult to ensure the same twisting curvature in two perpendicular directions ($k_{xy} = k_{yx}$), although these limitations can be greatly reduced by the use of a fine mesh. Therefore, the influence of curvature in the direction perpendicular to the moment being sought (products $v_x k_x$ and $v_y k_y$ in Equations of bending moment in Fig. 1) is ignored in grillage calculations, but the low value of Poisson's ratio and the fact that this effect is ignored in both directions reduces the implications of this inconsistency. Finally, the influence of the twisting moment is neglected on the calculation of shear, which may be significant in the case of very high skew bridges.

2.2 Guidelines for Deck Modeling

The properties of the 1-D beam elements of a grillage model or the 2-D elements of a plate model will depend of the type of deck cross-section being modelled: a solid slab, a voided slab, a beam-and-slab, or a cellular section, with or without edge cantilevers. When modelling a bridge deck using a plate FE model, O'Brien & Keogh (1999) recommend to: (a) use elements as regularly shaped as possible. In the case of a quadrilateral element, one dimension should not be larger than twice the other dimension; (b) avoid mesh discontinuities (i.e., in transitions from a coarse mesh to a fine mesh all nodes should be connected); and (c) place element nodes at bearings and use elastic springs if compressibility of bearings or soil was significant. It must also be noted there are inaccuracies associated to the calculation of shear near supports which can result in very high and unrealistic values.

The idealisation of a uniform solid slab deck as in Fig. 3 by a plate FE model is straightforward, i.e., the plate elements are given the same depth and material properties as the original slab. If using a grillage to model a solid slab deck, each grillage beam element (longitudinal or transverse) must have the properties (second moment of area and torsional constants) that resemble the longitudinal and transverse bending and twisting behaviour of the portion of the bridge being represented. By comparison of equations in Figs. 1 and 2, and neglecting the terms $v_y k_y$ and $v_x k_x$, the second moments of area for the grillage beams in the 'x' and 'y' directions that imitate the behaviour of a thin plate of width a , length b and depth d are given by:

$$I_x = a \frac{i}{1 - \nu_x \nu_y} \approx ai; \quad I_y = b \frac{i}{1 - \nu_x \nu_y} \approx bi \quad (3)$$

where i is the second moment of area per unit breadth of a thin plate or $d^3/12$. The torsional constants of the beams in the longitudinal and transverse directions will be given by:

$$J_x = aj; \quad J_y = bj \quad (4)$$

where j is the torsional constant per unit breadth of a slab or $d^3/6$. Hambly (1991) provides the following recommendations to decide on the number of beams and location of nodes for a grillage model: (a) place grillage beams along load paths within the structure (i.e., where longitudinal or transverse beams are located in the bridge, above bearings, etc.); (b) consider how stresses distribute within the structure (i.e., in the case of a solid thin plate, since the vertical shear flow due to a twisting moment stops at $0.3d$ from the section edge - d is the section depth -, an edge longitudinal grillage beam should be placed at this location); (c) the number of longitudinal beams can go from 1 to 20 and the spacing between longitudinal beams typically varies between $2d$ and $L/4$ where L is the span length. A grillage model is unable to capture the local load dispersion throughout the depth, therefore, smaller spacings than $2d$ do not necessarily improve accuracy; (d) the spacing between transverse grillage beams should be approximately equal to the spacing between longitudinal grillage beams; (e) in straight bridges, transverse grillage beams are perpendicular to longitudinal grillage beams except in the case of skew reinforcement; (f) when modelling high skew bridges, very close bearings or compressible soil, springs may be necessary to allow for possible vertical displacements of the grillage nodes at the support locations and; (g) a fine mesh should be employed in those areas with large variations of load effect, i.e., over internal supports. The load effect in a beam resulting from a grillage analysis represents the total moment acting on a portion of the bridge. Hence, small spacings between grillage beams are associated to small portions of the bridge and allow a better definition of the load effect distribution than a coarse mesh.

Fig. 3 shows plan, cross-section and elevation view of a grillage model associated to a solid slab deck of width W , span length L and depth d (an overhang of length V is allowed at both end supports). Placing the longitudinal edge beams at $0.3d$ from the edge and initially assuming the longitudinal beams are spaced by $2d$, the number of longitudinal beams, n_L , is given by the integer part, typically rounded to an odd number, of $((W-2 \times 0.3d)/2d+1)$. Therefore, the spacing between longitudinal beams, S_L , will be given by $(W-2 \times 0.3d)/(n_L-1)$. This value should be adjusted to facilitate the existence of longitudinal grillage beams across the path of the vehicle wheels. Similarly, the number of transverse beams, n_T , is initially estimated by assuming they have the same spacing as longitudinal beams ($n_T = \text{integer}(L/S_L+1)$), and their final spacing will be given by $S_T = L/(n_T-1)$. In the grillage model of Fig. 3, there are 4 types of beams: longitudinal interior (thick dashed lines), longitudinal edge (thin dashed), transverse interior (thick dotted) and transverse edge (thin dotted). For a uniform slab deck, all beams bend about the neutral axis of the cross-section they represent (mid-depth). Therefore, the total second moment of area, I , and torsional constant, J , will be the result of multiplying i ($= d^3/12$) and j ($= d^3/6$) respectively by the width of the portion of the bridge that each grillage beam is associated to (Equations (3) and (4)). When calculating the beam torsional constant, only the breadth where shear stresses act needs to be taken into account (i.e., there will be no shear stresses within $0.3d$ from the bridge edge).

In the case of modelling a slab with edge cantilever (Fig. 4(a)), three regions can be distinguished: the edge elements that constitute the cantilever, the interior elements that form the main bridge deck and the transition elements separating the other two. When calculating the properties of the longitudinal beams, it must be taken into account that: (a) the edge portion of the bridge will bend about the neutral axis of the cantilever (i.e., $i = d_c^3/12$ where d_c is the cantilever thickness); (b) the interior elements will bend about the neutral axis of the bridge (i.e., $i = [d_m((d_m/2)-z_b)^2 + d_m^3/12]$ where z_b and d_m are the location

of the neutral axis of the cross-section of the bridge and the depth of the main deck respectively); and (c) the transition elements will bend about some axis in-between (i.e., inertia is obtained subtracting the inertia of the cantilever about its own axis from the inertia of both the transition and cantilever elements with respect to the bridge neutral axis. This adjustment allows for a gradual variation of the neutral axis of the bridge that rises from the interior elements towards the edge elements). The thickness of the edge (t_e), interior (t_m) and transition (t_t) plate elements can be obtained from $t = \sqrt[3]{12i}$ using the value of second moment of area per unit breadth, i , corresponding to the element under investigation.

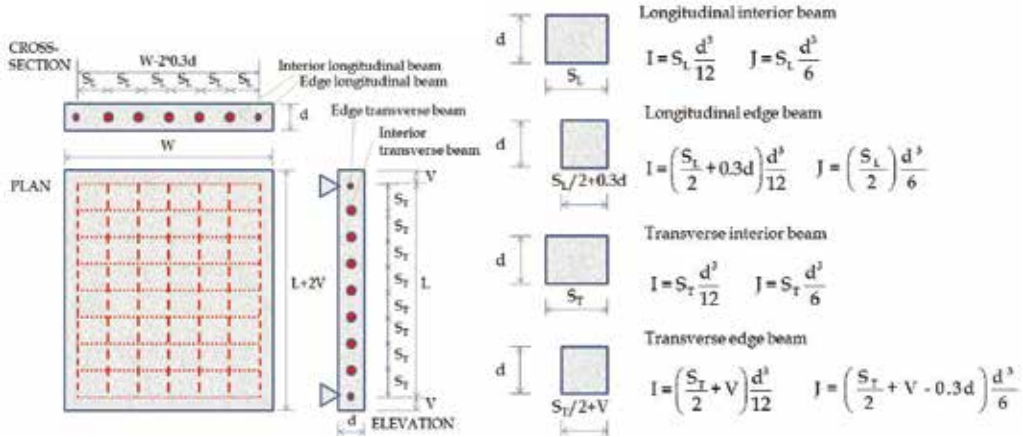


Fig. 3. Grillage model of a solid slab deck

When modelling orthotropic decks (different second moment of area per unit breadth in the 'x' and 'y' directions), the second moment of area of longitudinal and transverse beams of a grillage model can be adjusted accordingly. However, in the case of a plate model, the plate elements have a unique thickness, which means the second moment of area per unit breadth in the 'x' and 'y' directions must be the same. Therefore, flexural rigidities per unit breadth in the 'x' and 'y' direction (D_x and D_y) of a thin plate are given by:

$$D_x = \frac{E_x i}{1 - \nu_x \nu_y} \approx E_x i; \quad D_y = \frac{E_y i}{1 - \nu_x \nu_y} \approx E_y i \quad (5)$$

D_x and D_y relate bending moments to curvatures in the 'x' and 'y' directions respectively (Fig. 1). If $E_x = E_y$, then $D_x = D_y$ (isotropic). For the case of a voided slab deck (Fig. 4(b)), the flexural rigidities per unit breadth in the 'x' and 'y' directions, derived in Equation (6) using theory of continuum mechanics (O'Brien & Keogh, 1999), can be clearly different depending on the ratio void diameter to bridge depth (i.e., highly orthotropic).

$$D_x = E \left[\frac{d^3}{12} - \frac{d_v^4}{64s_v} \right]; \quad D_y = \frac{Ed^3}{12} \left[1 - 0.95 \left(\frac{d_v}{d} \right)^4 \right] \quad (6)$$

where E is the modulus of elasticity, d is the full depth of the voided slab, d_v is the void diameter and s_v is the distance between void centres. In order to be able to imitate the response of an orthotropic deck as closely as possible, the flexural rigidities of a plate FE model (Equation (5)) must be adjusted to match those in the original voided slab deck (Equation (6)). For this purpose, two different modulus of elasticity are assumed, \tilde{E}_x and \tilde{E}_y . Then, by doing $\tilde{E}_x = E$, it is possible to obtain the value of the plate thickness replacing the flexural rigidity of the original bridge in the 'x' direction (from Equation (6)) into Equation (5), i.e., $t = \sqrt[3]{D_x/E}$. Similarly, the modulus of elasticity in the 'y' direction, \tilde{E}_y , is adjusted to give the correct D_y (from Equation (6)) through $\tilde{E}_y = D_y/i = 12D_y/t^3$.

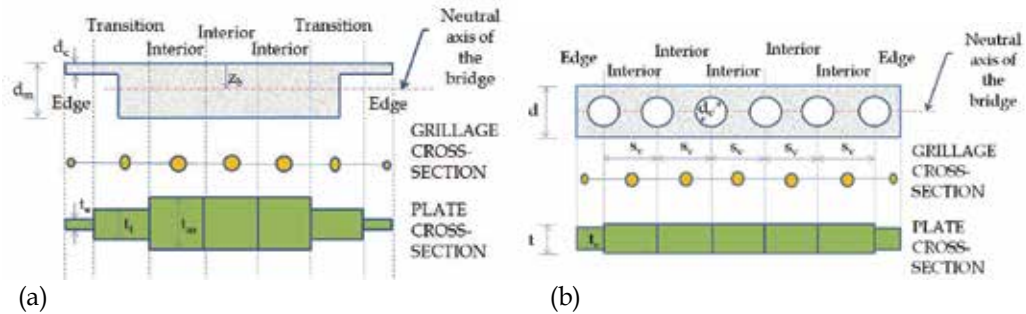


Fig. 4. Grillage and plate FEM: (a) Slab with edge cantilever; (b) Voided slab

When using a grillage to model a beam-and-slab deck, the girders and diaphragms of the deck are associated to longitudinal and transverse grillage beams respectively. Additional transverse grillage beams representing only slab will be necessary to cover for the entire bridge span. When calculating the properties of the longitudinal grillage beams, it must be taken into account that the portion of the deck being represented will bend about its own axis (and not about the neutral axis of the entire bridge cross-section) due to the poor load transfer of this type of construction. The torsional constant of a longitudinal grillage beam will be equal to the torsional constant of the bridge beam plus the torsional constant of the thin slab. Therefore, in the case of a composite section made of different materials for beam and slab, properties will be obtained using an equivalent area in one of the two materials, adjusting the width of the other material based on the modular ratio. When transverse grillage beams are not placed at the location of diaphragm beams, and they only represent slab, they will bend about the neutral axis of the slab and their torsional constant will be equal to the torsional constant of a thin slab. Alternatively, a beam-and-slab bridge deck can be modelled using a combination of 1-D beams and 2-D plates. The guard rails, longitudinal and transverse beams in the bridge will be idealised with beam elements while the deck will be modelled with plate elements (Chompooming & Yener, 1995; Kim et al., 2005; González et al., 2008a).

In the case of using a grillage to model a cellular deck, a longitudinal grillage beam can be placed at the location of each web making the cross-section. These grillage beams will bend about the neutral axis of the full section. The properties of edge longitudinal grillage members representing edge webs must take into account the contribution of edge

cantilevers to torsion and bending stiffness. The sum of the torsional constants of all grillage longitudinal beams will be equal to the sum of the torsional constants of the individual cells (i.e., a closed section) plus the torsional constants of the edge cantilevers (i.e., a thin slab). Caution must be placed upon the determination of the area of the transverse grillage beams, which should be reduced to allow for transverse shear distortion, typical of cellular sections. Models of bridges with cables attached to the deck can be found in Wang & Huang (1992), Chatterjee et al (1994a), Guo & Xu (2001), Xu & Guo (2003) and Chan et al (2003).

3. The Vehicle

While the equations of motion of the bridge are obtained using the FEM, there are three alternative methods to derive the equations of motion of the vehicle: (a) imposing equilibrium of all forces and moments acting on the vehicle and expressing them in terms of their DOFs (Hwang & Nowak, 1991; Kirkegaard et al., 1997; Tan et al., 1998; Cantero et al., 2010), (b) using the principle of virtual work (Fafard et al., 1997) or a Lagrange formulation (Henchi et al., 1998), and (c) applying the code of an available FE package. The equations of equilibrium deal with vectors (forces) and they can be applied to relatively simple vehicle models, while an energy approach has the advantage of dealing with scalar amounts (i.e., contribution to virtual work) that can be added algebraically and are more suitable for deriving the equations of complex vehicle models. Similarly to the bridge, the equations of motion of a vehicle can be expressed in matrix form as:

$$[M_v]\{\ddot{w}_v\} + [C_v]\{\dot{w}_v\} + [K_v]\{w_v\} = \{f_v\} \quad (7)$$

where $[M_v]$, $[C_v]$ and $[K_v]$ are global mass, damping and stiffness matrices of the vehicle respectively, $\{w_v\}$, $\{\dot{w}_v\}$ and $\{\ddot{w}_v\}$ are the vectors of global coordinates, their velocities and accelerations respectively, and $\{f_v\}$ is the vector of forces acting on the vehicle at time t .

The modes of vibration of the theoretical model should resemble the body pitch/bounce and axle hop/roll motions of the true vehicle. Body oscillations are related to the stiffness of suspensions and sprung mass (vehicle body) and they have frequencies between 1 and 3 Hz for a heavy truck and between 2 and 5 Hz for a light truck. Axle oscillations are mainly related to the unsprung masses (wheels and axles) and have higher frequencies, in a range from 8 Hz to 15 Hz. In the presence of simultaneous vehicle presence, frequency matching and resonance effects on the bridge response are unlikely. But in the case of a single truck, Cantieni (1983) has measured average dynamic increments on bridges with normal pavement conditions of 30-40% (maximum of 70%) when their first natural frequency fell within the range 2 to 5 Hz, that decreased to an average dynamic increment between 10 and 20% when the natural frequency fell outside that frequency range. Green et al (1995) analyse the decrease in bridge response when the vehicles are equipped with air-suspensions compared to steel suspensions due to the low dynamic forces they apply to the bridge.

3.1 Types of FE Models for Vehicles

There is a wide range of sprung vehicle models used in VBI investigations. Planar vehicle models have been found to provide a reasonable bridge response for ratios bridge width to

vehicle width greater than 5 (Moghimi & Ronagh, 2008a). A single-DOF model can be used for a preliminary study of the tyre forces at low frequencies due to sprung mass bouncing and pitching motion (Chatterjee et al., 1994b; Green & Cebon, 1997) and a two-DOF model (i.e., a quarter-car) can be employed to analyse main frequencies corresponding to body-bounce and axle hop modes (Green & Cebon, 1994; Chompooming & Yener, 1995; Yang & Fonder, 1996; Cebon, 1999). If the influence of axle spacing was investigated, then a rigid walking beam (Hwang & Nowak, 1991; Green & Cebon, 1994; Chompooming & Yener, 1995) or an articulated multi-DOF model (Veletsos & Huang, 1970; Hwang & Nowak, 1991; Green et al., 1995; Harris et al., 2007) will become necessary. The vehicle model can be extended to three dimensions to allow for roll and twisting motions. Most of these spatial models consist of an assemblage of 1-D elements, but they can also be made of 2-D and 3-D FEs for a detailed representation of the vehicle aerodynamic forces and deformations (Kwasniewski et al., 2006). However, it seems unlikely such a degree of sophistication could affect the bridge response significantly.

So, spatial FE vehicle models are typically composed of mass, spring, bar and rigid elements that are combined to model tyre, suspension, axle/body masses and connections between them. The equations of motion of a spatial vehicle model can be established for rigid (Tan et al., 1998; Zhu & Law, 2002; Kim et al., 2005) or articulated configurations (EIMadany, 1988; Fafard et al., 1997; Kirkegaard et al., 1997; Nassif & Liu, 2004; Cantero et al., 2010). In the latter, equations of motion can be formulated for the tractor and trailer separately, and the DOFs of both parts can be related through a geometric condition that takes into account the hinge location. A series of lumped masses are employed to represent axles, tractor and trailer body. The body masses are connected to the frame by rigid elements, and the frame is connected to the axles by spring-dashpot systems that model the response of the suspension. Each axle is typically represented as a rigid bar with lumped masses at both ends that correspond to the wheel, axle bar, brakes and suspension masses (Alternatively, the axle mass can be assumed to be concentrated at the local centre of gravity of the rigid bar connecting both wheels with two DOFs: vertical displacement and rotation about the longitudinal axis). Then, the lumped mass at each wheel is connected to the road surface by a spring-dashpot system simulating the response of the tyre. Fig. 5 illustrates the forces acting on the tractor and trailer masses of an articulated vehicle travelling at constant speed. Equations of motion for a vehicle braking or accelerating can be found in Law & Zhu (2005) and Ju & Lin (2007).

The system of time-varying forces acting on each lumped mass consists of inertial forces (i.e., $m_{i-j}\ddot{w}_{i-j}$), gravity forces ($m_{i-j}g$), suspension forces ($f_{s,i-j}$) and tyre forces ($f_{t,i-j}$).

Equilibrium of vertical forces acting on the sprung mass of the tractor and equilibrium of moments of all forces about a 'y' axis going through its centroid (Fig. 5(a)) lead to Equations (8) and (9) respectively.

$$\sum F_z = 0 \rightarrow m_s (\ddot{w}_s - g) + \sum_{i=1}^p (f_{s,i-left} + f_{s,i-right}) + f_H = 0 \quad (8)$$

where f_H is the interaction force at the hinge and p is the number of axles supporting the sprung mass.

$$\Sigma M_y = 0 \rightarrow I_y \ddot{\theta}_y - \sum_{i=1}^p (f_{S,i-right} + f_{S,i-left}) x_i - f_H x_H = 0 \quad (9)$$

where x_i is the distance from axle i to the centroid of the sprung mass, and it can be positive (to the right of the centroid) or negative (to the left of the centroid). Similar equations of equilibrium can be obtained for the sprung mass of the trailer. From equilibrium of moments of all forces acting on the body mass about an 'x' axis going through its centroid (Fig. 5(b)), it is possible to obtain:

$$\Sigma M_x = 0 \rightarrow I_x \ddot{\theta}_x + \sum_{i=1}^p f_{S,i-left} a_{i-left} - \sum_{i=1}^p f_{S,i-right} a_{i-right} = 0 \quad (10)$$

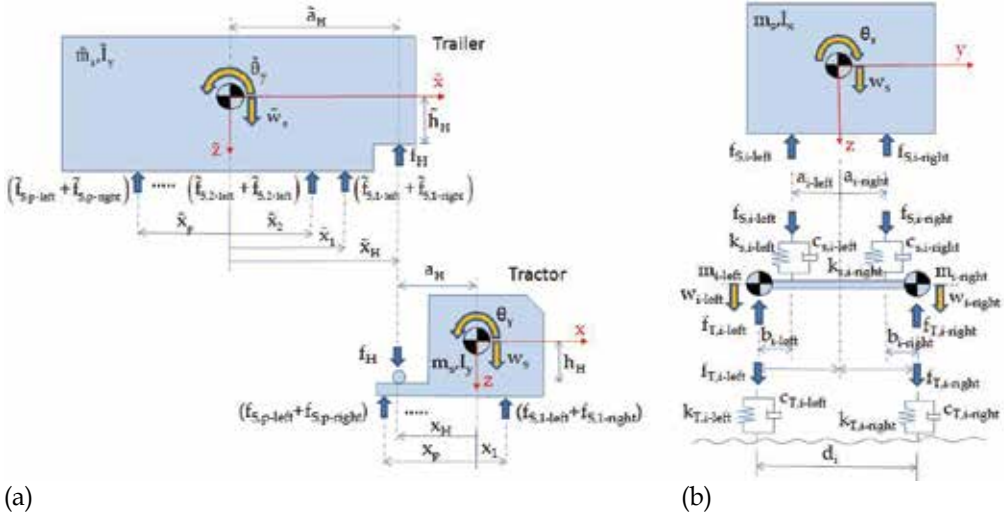


Fig. 5. Sprung lumped mass models: (a) Forces acting on tractor and trailer sprung masses (side view), (b) Forces acting on the sprung and unsprung masses (front view)

Wheel, axle and suspension mass are assumed to be concentrated at both ends of an axle i (m_{i-left} and $m_{i-right}$), and equilibrium at these lumped masses is given by Equation (11). Identical equations can be obtained for the unsprung masses of the trailer.

$$\Sigma F_z = 0 \rightarrow \begin{cases} m_{i-left} (\ddot{w}_{i-left} - g) - \frac{d_i - b_{i-left}}{d_i} f_{S,i-left} - \frac{b_{i-right}}{d_i} f_{S,i-right} + f_{T,i-left} = 0 \\ m_{i-right} (\ddot{w}_{i-right} - g) - \frac{d_i - b_{i-right}}{d_i} f_{S,i-right} - \frac{b_{i-left}}{d_i} f_{S,i-left} + f_{T,i-right} = 0 \end{cases} ; i=1, \dots, p \quad (11)$$

The two equations of equilibrium of vertical forces of the tractor (Equation (8)) and trailer body masses can be combined into one by cancelling out f_H . Therefore, the suspension and tyre forces can be expressed as a function of the DOFs of the vertical displacement of the unsprung and sprung masses. Tyre forces can be defined by viscous damping elements (proportional to the damping constant and the relative change in velocity between the two

end points of the damper), and spring elements (proportional to the stiffness constant and the relative change in displacement between the two end points of the spring) as given in Equation (12) for the wheel forces in the left side.

$$f_{T,i-left} = k_{T,i-left} (w_{i-left} + u_{i-left} - r_{i-left}) + c_{T,i-left} (\dot{w}_{i-left} + \dot{u}_{i-left} - \dot{r}_{i-left}) \quad (12)$$

where u_{i-left} and r_{i-left} represent the bridge displacement and the height of the road irregularities respectively under the left wheel of axle i at a specified point in time. u_{i-j} is related to the nodal displacements and rotations $\{w_b\}^{(e)}$ of the bridge element where the wheel $i-j$ is located through the displacement interpolation functions $\{N(\bar{x}, \bar{y})\}$ of the element. The values of $\{N(\bar{x}, \bar{y})\}$ are a function of the coordinates (\bar{x}, \bar{y}) of the wheel contact point with respect to the coordinates of the bridge element.

$$u_{i-j} = \{N(\bar{x}, \bar{y})\}^T \{w_b\}^{(e)} \quad (13)$$

In Figure 5(b), suspension systems have also been defined by a viscous damping element in parallel with a spring element. In this case of linear suspension elements, the forces f_s are given by Equation (14) for the suspension forces in the left side of the vehicle.

$$f_{S,i-left} = k_{S,i-left} (w_s + x_i \theta_y + (a_{i-left} + b_{i-left}) \theta_x - w_{i-left}) + c_{S,i-left} (\dot{w}_s + x_i \dot{\theta}_y + (a_{i-left} + b_{i-left}) \dot{\theta}_x - \dot{w}_{i-left}) \quad (14)$$

The equations of tractor and trailer are not independent, as a compatibility condition can be defined between both rigid bodies rotating about the hinge (also known as fifth wheel point). Therefore, the number of DOFs of the system can be reduced by considering the following relationship between the displacements of the centroids of tractor and trailer:

$$\tilde{w}_s = w_s + x_H \theta_y + \tilde{x}_H \tilde{\theta}_y \quad (15)$$

All equations of equilibrium can be expressed as a function of the vertical displacements of the wheels $(w_{1-left}, w_{1-right}, \dots, w_{p-left}, w_{p-right}, \tilde{w}_{1-left}, \tilde{w}_{1-right}, \dots, \tilde{w}_{p-left}, \tilde{w}_{p-right})$, the three DOFs of the tractor $(w_s, \theta_x, \theta_y)$ and the two DOFs of the trailer $(\tilde{\theta}_x, \tilde{\theta}_y)$, by replacing Equations (12), (14) and (15) into the equations of equilibrium of the lumped masses. The equations can be expressed in a matrix form which will lead to the mass, stiffness and damping matrices of the vehicle (Equation (7)). The coefficients of these matrixes can be found in Cantero et al (2010) for an arbitrary number of tractor and trailer axles. The forcing vector $\{f_v\}$ will be a combination of tyre properties (stiffness/damping) and height of the road/bridge profile. Values for parameters of suspension and tyre systems are available in the literature (Wong, 1993; Kirkegaard et al., 1997; Fu & Cebon, 2002; Harris et al., 2007). In the case of a 5-axle articulated truck, typical magnitudes are 7000 kg for a tractor sprung mass, 700 kg for a steer axle, 1000 kg for a drive axle, 800 kg each trailer axle, tyre stiffness of 735×10^3 kN/m and damping of 3×10^3 kNs/m, suspension stiffness of 300×10^3 kN/m for a steer axle, 500×10^3 and 1000×10^3 kN/m for a drive axle if air- and steel-suspension respectively, 400×10^3 and

1250x10³ kN/m for a trailer axle if air- and steel-suspension respectively, and suspension viscous damping of 3x10³ kNs/m.

The general form of the equations of motion described in this section is only valid for linear tyre and suspension systems and it needs to be adapted to introduce different degrees of complexity within the vehicle behaviour, such as tyre or suspension non-linearities. Green & Cebon (1997) propose the following equation that facilitates the incorporation of non-linear elements:

$$[M_v]\{\ddot{w}_v\} = [S_v]\{f_s\} + [T_v]\{f_T\} + \{G\} \quad (16)$$

where $\{f_s\}$ and $\{f_T\}$ are vectors of suspension and tyre forces respectively, $[S_v]$ and $[T_v]$ are constant transformation matrices relating the suspension and tyre forces respectively to the global coordinates of the vehicle, and $\{G\}$ is the vector of gravitational forces applied to the vehicle. The relationship between Equations (7) and (16) is given by $\{f_v\} = [T_v][K_T]\{v\} = [C_T]\{\dot{v}\}$ where $[K_T]$ and $[C_T]$ are stiffness and damping matrices respectively for tyre elements and $\{v\}$ is the height of road irregularities (or road irregularities plus bridge deflections), $[C_v] = [S_v]\{C_s\}[S_v]^T + [T_v][C_T][T_v]^T$ and $[K_v] = [S_v]\{K_s\}[S_v]^T + [T_v][K_T][T_v]^T$ where $[K_s]$ and $[C_s]$ are stiffness and damping matrices respectively for suspension elements.

4. The Road

The road profile can be measured or simulated theoretically. When simulating a profile $r(x)$, it can be generated from power spectral density functions as a random stochastic process:

$$r(x) = \sum_{i=1}^N \sqrt{2G_d(n_k)} \Delta n \cos(2\pi n_k x - \theta_i) \quad (17)$$

where $G_d(n_k)$ is power spectral density function in m²/cycle/m; n_k is the wave number (cycle/m); θ_i is a random number uniformly distributed from 0 to 2π ; Δn is the frequency interval ($\Delta n = (n_{\max} - n_{\min})/N$ where n_{\max} and n_{\min} are the upper and lower cut-off frequencies respectively); N is the total number of waves used to construct the road surface and x is the longitudinal location for which the road height is being sought. The road class is based on the roughness coefficient a (m³/cycle), which is related to the amplitude of the road irregularities, and determines $G_d(n_k)$ ($G_d(n_k)$ is equal to $a/(2\pi n_k)^2$). ISO standards specify 'A' ($a < 2 \times 10^{-6}$), 'B' ($2 \times 10^{-6} \leq a < 8 \times 10^{-6}$), 'C' ($8 \times 10^{-6} \leq a < 32 \times 10^{-6}$), 'D' ($32 \times 10^{-6} \leq a < 128 \times 10^{-6}$) and other poorer road classes depending on the range of values where a is located (ISO 8608, 1995). For a given roughness coefficient, different road profiles can be obtained varying the random phase angles θ_i . When using two parallel tracks of an isotropic surface, a coherence function needs to be employed to produce a second random profile correlated with the first profile (Cebon & Newland, 1983; Cebon, 1999; Nassif & Liu, 2004). The coherence function guarantees good and poor correlation between two parallel tracks for long and short wavelengths respectively. The contact between the bridge and the vehicle is typically assumed to be at a single point rather than the area corresponding to the tyre

patch. Therefore, the lengths of randomly generated road profile are passed through a moving average filter to simulate the envelope of short wavelength disturbances by the tyre contact patch (i.e., 0.2 m). The magnitude of the bridge response depends strongly not only on the general unevenness of the bridge surface, but also on the velocity of the vehicles (González et al., 2010), the condition of the road leading to the bridge, and the effects of occasional large irregularities such as potholes, misalignments at the abutments or expansion joints that are often found on the bridge approach. Chompooming & Yener (1995) show how certain combinations of bumps and vehicle speed can originate a high dynamic excitation of the bridge.

5. Vehicle-Bridge Interaction Algorithms

When analysing the VBI problem, two sets of differential equations of motion can be established: one set defining the DOFs of the bridge (Equation (1)) and another set for the DOFs of the vehicle (Equation (7)). It is necessary to solve both subsystems while ensuring compatibility at the contact points (i.e., displacements of the bridge and the vehicle being the same at the contact point of the wheel with the roadway). The algorithms to carry out this calculation can be classified in two main groups: (a) those based on an uncoupled iterative procedure where equations of motion of bridge and vehicle are solved separately and equilibrium between both subsystems and geometric compatibility conditions are found through an iterative process (Veletsos & Huang, 1970; Green et al., 1995; Hwang & Nowak, 1991; Huang et al., 1992; Chatterjee et al., 1994b; Wang et al., 1996; Yang & Fonder, 1996; Green & Cebon, 1997; Zhu & Law, 2002; Cantero et al., 2009), and (b) those based on the solution of the coupled system, i.e., there is a unique matrix for the system that is formed by eliminating the interaction forces appearing in the equations of motion of bridge and vehicle, and updated at each point in time (Olsson, 1985; Yang & Lin, 1995; Yang & Yau 1997; Henchi et al., 1998; Yang et al., 1999, 2004a; Kim et al., 2005; Cai et al., 2007; Deng & Cai, 2010; Moghimi & Ronagh, 2008a). The use of Lagrange multipliers can also be found in the solution of VBI problems (Cifuentes, 1989; Baumgärtner, 1999; González et al., 2008a).

A step-by-step integration method must be adopted to solve the uncoupled or coupled differential equations of motion of the system. These numerical methods break the time down into a number of steps, Δt , and calculate the solution $w(t+\Delta t)$ from $w(t)$ based on assumed approximations for the derivatives that appear in the differential equations. They are different from methods for single-DOF systems because most FE models with lots of DOFs poorly idealise the response of the higher modes, and the integration method should have optimal dissipation properties for the removal of those non-reliable high frequency contributions.

Fourth-order Runge-Kutta is a popular integration method in the solution of large multi-DOF VBI systems (Frýba 1972; Huang et al., 1992; Wang & Huang, 1992; Cantero et al., 2009; Deng & Cai, 2010). Acceleration is expressed as a function of the other lower derivatives and a change of variable transforms the second order equation of motion into two first order equations (Equation (18)). Then, the recurrence formulae of fourth-order Runge-Kutta is employed to approximate the derivatives according to a weighted average of four estimates of the slope in the interval Δt .

$$\{\dot{Z}\} = \{\dot{w}\} ; \quad \{\dot{Z}\} = [M]^{-1} \{f\} - [C]\{Z\} - [K]\{w\} \quad (18)$$

Fourth-order Runge-Kutta is easy to implement, however, it is conditionally stable and it requires evaluating several functions per time step which can be time-consuming. For these reasons, some authors prefer implicit unconditionally stable integration methods such as Newmark- β (Olsson, 1985; Hwang & Nowak, 1991; Chompooming & Yener, 1992; Yang & Fonder, 1996; Fafard et al., 1997; Yang & Yau, 1997; Zhu & Law, 2002; Kim et al., 2005) or Wilson- θ (Tan et al., 1998; Nassif & Liu, 2004). The Newmark family of integration methods is based on a truncation of Taylor's series that assumes a linear variation of acceleration from time t to time $(t+\Delta t)$ (a constant acceleration with parameters $\beta = 0.25$ and $\gamma = 0.5$ is typically used to avoid instability problems). Wilson- θ is a modified version of the Newmark method, where acceleration is assumed to vary linearly from time t to time $(t+\theta\Delta t)$ ($\theta \geq 1.37$, usually 1.4). After the conditions at time $(t+\theta\Delta t)$ are known, they are referred back to produce a solution at time $(t+\Delta t)$. Clough & Penzien (1993) recommend a time step $\Delta t \leq 1/(10f_{\max})$ where f_{\max} is the maximum frequency of the system to ensure convergence. Typical values of Δt lie between 0.001 and 0.0001 s.

In the case of non-linear equations of motion, i.e., the non-linear relationship between the force and nodal displacements in leaf-spring suspensions, an iterative procedure needs to be implemented for each point in time (Veletsos & Huang, 1970; Hwang & Nowak, 1991; Wang & Huang, 1992; Chatterjee et al., 1994a,b; Tan et al., 1998; Nassif & Liu, 2004). The initial values of the force and nodal displacement of the suspension are assumed to be those of the preceding time step. However, once the equations of the system are solved for the current time, the suspension nodal displacement will be somehow different from the value in the previous time step and the associated suspension force may change with respect to the assumed initial value. The calculations need to be repeated with the updated value of suspension force, and treated in an iterative procedure until reaching an acceptable tolerance that indicates negligible difference between displacements/forces of two successive iterations.

5.1 Algorithms based on an Uncoupled Iterative Procedure

These algorithms treat the equations of motion of the vehicle and the bridge as two subsystems and solve them separately using a direct integration scheme. The compatibility conditions and equilibrium equations at the interface between the vehicle tyres and bridge deck are satisfied by an iterative procedure. The basic idea behind this procedure consists of assuming some initial displacements for the contact points, which can be replaced in the equations of motion of the vehicle to obtain the interaction forces. Then, these interaction forces are employed in the equations of motions of the bridge to obtain the bridge displacements that will represent improved estimates of the initial displacements assumed for the contact points. The process is repeated until differences in two successive iterations are sufficiently small. These algorithms typically employ implicit schemes of integration such as Newmark- β or Wilson- θ methods to solve each subsystem and to achieve convergence after a number of iterations. The time step Δt here is larger than in the coupled solution, although the convergence rate may be slow.

One possible algorithm within the group of iterative procedures is illustrated in Fig. 6. The vehicle and the bridge interact through the tyre forces imposed on the bridge deck. The profile, $v(x,t)$, that is used to excite the vehicle is the sum of the original road profile ($r(x)$) and the deflection of the bridge ($w_b(x,t)$) (In the first step of the iterative procedure, no bridge displacements have been calculated yet and $v(x,t)$ can be taken to be $r(x)$). The

interaction forces (tyre forces $\{f_T\}$) are obtained solving for the DOFs of Equation (7) and replacing into Equation (12). These interaction forces are converted to equivalent bridge forces acting on the bridge nodes in the vicinity of the contact point using a location matrix ($\{f_b\}=[L]\{f_T\}$ where $[L]$ is a location matrix that relates the tire forces to the DOFs of the bridge). Then, $\{f_b\}$ is employed to obtain a new set of displacements $w_b(x,t)$ using Equation (1). This procedure is repeated for a number of iterations at each time step until some convergence criteria is met (i.e., difference between the bridge deflection $w_b(x,t)$ of two successive iterations being sufficiently small). Green & Cebon (1997) suggest to average bridge displacements in two successive iterations to facilitate convergence which they define as the relative difference in displacement of two successive iterations with respect to the maximum bridge deflection to be smaller than 2%. At that point, the vehicle is moved forward and the iterative procedure is repeated for the new location of the forces on the bridge.

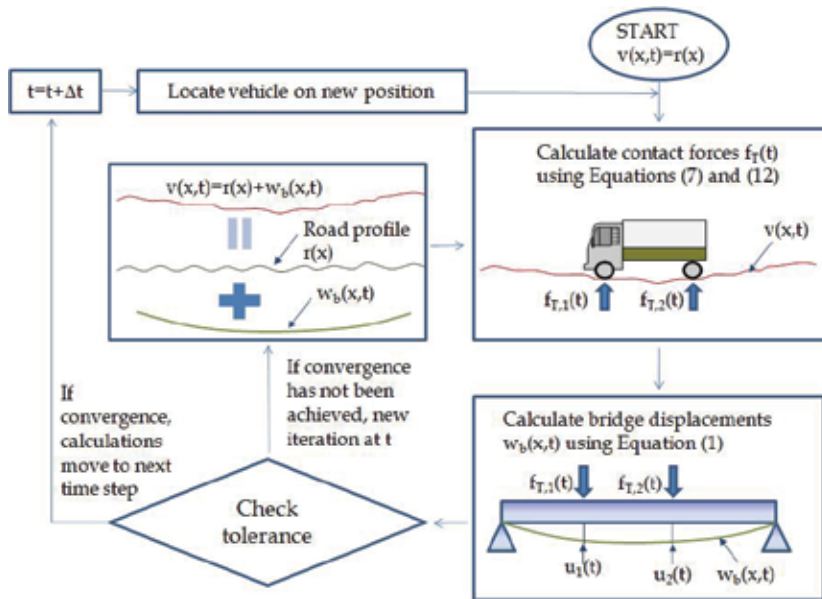


Fig. 6. VBI iterative procedure

Cantero et al (2009) suggest an alternative iterative procedure, where rather than calculating the final interaction forces at each time step, an initial estimate of the entire force history $\{f_T\}^{(1)}$ is obtained using Equations (7) and (12) with only the road profile $r(x)$ ($=v(x,t)^{(1)}$) as excitation source. Equation (1) is employed to calculate the bridge deflections, $w_b(x,t)^{(1)}$, due to the equivalent nodal forces $\{f_b\}^{(1)}$ derived from $\{f_T\}^{(1)}$. The bridge deflection $w_b(x,t)^{(1)}$ is then added to the road profile $r(x)$ to form $v(x,t)^{(2)}$, and a new estimated of the force history $\{f_T\}^{(2)}$ is obtained using the equations of motion of the vehicle and the profile $v(x,t)^{(2)}$ as excitation source. These time-varying forces $\{f_T\}^{(2)}$ are converted into bridge nodal forces $\{f_b\}^{(2)}$ that will result into a new bridge deflection history $w_b(x,t)^{(2)}$ using Equation (1). The process is repeated until convergence is achieved.

While the previous algorithms have been formulated in the time domain, Green & Cebon (1994,1997), Green et al (1995) and Henchi et al (1997) propose to solve the uncoupled

system of equations in the frequency domain. The equations of motion are solved by convolution of modal impulse response functions and modal excitation forces through the FFT, and application of modal superposition.

It must be noted that many DOFs are involved in the FE model of the bridge subsystem, but only the first modes of vibration make a significant contribution to the dynamic response of a VBI system. Therefore, the modal superposition method (MSM) is typically employed to solve the equations of motion of the bridge which reduces the computation effort considerably (Clough & Penzien, 1993). The basis of the MSM is the transformation of the original system of coupled equations (Matrixes in Equation (1) with non-zero off-diagonal terms) into a smaller set of uncoupled independent modal coordinate equations (i.e., zero off-diagonal terms). The total dynamic response will be obtained by superposition of the response obtained for each modal coordinate. The displacement vector of the bridge $\{w_b\}$ can be expressed as a function of the modal coordinates $\{q_b\}$ as follows:

$$\{w_b\} = [\Phi_b] \{q_b\} \quad (19)$$

where $[\Phi_b] = [\{\Phi_1\} \{\Phi_2\} \dots \{\Phi_m\}]$ is the normalized modal shape matrix containing a number m of mode shapes. The mode shapes $\{\Phi_i\}$ and frequencies associated to these mode shapes ω_i can be found using eigenvalue analysis:

$$\left[[K_b] - \omega_i^2 [M_b] \right] \{\Phi_i\} = \{0\} \quad (20)$$

Therefore $[\Phi_b]$ is normalized such that:

$$[\Phi_b]^T [M_b] [\Phi_b] = [I] \quad ; \quad [\Phi_b]^T [K_b] [\Phi_b] = [\omega_b^2] \quad (21)$$

where $[\omega_b^2]$ is a diagonal matrix containing the squares of the natural frequencies. Equation (1) can be written in modal coordinates as:

$$[M_b] [\Phi_b] \{\ddot{q}_b\} + [C_b] [\Phi_b] \{\dot{q}_b\} + [K_b] [\Phi_b] \{q_b\} = \{f_b\} \quad (22)$$

Assuming that the damping matrix $[C_b]$ satisfies modal orthogonality conditions (e.g., Rayleigh damping) and premultiplying both sides of the equation by $[\Phi_b]^T$, the following simplified system with m differential equations in modal coordinates results:

$$\{\ddot{q}_b\} + 2[\xi_b] [\omega_b] \{\dot{q}_b\} + [\omega_b^2] \{q_b\} = [\Phi_b]^T \{f_b\} \quad (23)$$

where $[\xi_b]$ and $[\omega_b]$ are modal damping and modal frequency matrixes of the bridge respectively and their size is related to the number of modes of vibration considered; $\{q_b\}$ is the modal coordinate vector and a dot means derivative with respect to time. Modal Equation (23) is a series of independent single-DOF equations, one for each mode of vibration, that can be solved by Newmark, Wilson- θ , Runge-Kutta or a piece-wise interpolation integration technique. The total response is obtained from the superposition of the individual modal solutions (Equation (19)). In some cases, the modal equations can

become coupled (i.e., damping matrix of the system not being diagonal after modal transformation) and the modal equations will need to be solved using a step-by-step integration method simultaneously rather than individually, but the reduction of the original number of equations where all DOFs were considered to a number m of modal equations will still be computationally advantageous. If the number of DOFs of the vehicle was considerably large, the MSM can also be applied to the equations of motion of the vehicle subsystem (Equation (7)).

5.2 Algorithms based on the Solution of the Coupled System

These algorithms are based on the solution of a unique system matrix at each point in time. The system matrix changes as the vehicle moves and its time-dependent properties can be derived using the principle of virtual work. These algorithms commonly use a step-by-step integration scheme such as Newmark- β (Kim et al., 2005) or fourth order Runge-Kutta (Deng & Cai, 2010) with a small time step Δt to solve the system matrix at each point in time. This procedure can be carried without any iteration if linear elements are employed, but if non-linear elements, such as friction, were present in the model, iterations will be necessary regardless of the vehicle-bridge system being solved as two subsystems or as one. The vehicle-bridge equations can be combined to form the system given below:

$$\begin{bmatrix} [M_b] & [0] \\ [0] & [M_v] \end{bmatrix} \begin{Bmatrix} \{\ddot{w}_b\} \\ \{\ddot{w}_v\} \end{Bmatrix} + \begin{bmatrix} [C_b+C_{b-b}] & [C_{b-v}] \\ [C_{v-b}] & [C_v] \end{bmatrix} \begin{Bmatrix} \{\dot{w}_b\} \\ \{\dot{w}_v\} \end{Bmatrix} + \begin{bmatrix} [K_b+K_{b-b}] & [K_{b-v}] \\ [K_{v-b}] & [K_v] \end{bmatrix} \begin{Bmatrix} \{w_b\} \\ \{w_v\} \end{Bmatrix} = \begin{Bmatrix} \{f_{b-r}\} \\ \{f_{b-r}\} + \{G\} \end{Bmatrix} \quad (24)$$

where $\{G\}$ are gravity forces and, $[C_{b-b}]$, $[C_{b-v}]$, $[C_{v-b}]$, $[K_{b-b}]$, $[K_{b-v}]$, $[K_{v-b}]$ and $\{f_{b-r}\}$ are time dependent matrixes/vectors that depend on the value and location of the interaction forces at each point in time. Equation (25) shows how MSM can be used to simplify the bridge subsystem of Equation (24).

$$\begin{aligned} & \begin{bmatrix} [I] & [0] \\ [0] & [M_v] \end{bmatrix} \begin{Bmatrix} \{\ddot{q}_b\} \\ \{\ddot{w}_v\} \end{Bmatrix} + \begin{bmatrix} 2[\xi_b][\omega_b] + [\Phi_b]^T [C_{b-b}] [\Phi_b] & [\Phi_b]^T [C_{b-v}] \\ [C_{v-b}] [\Phi_b] & [C_v] \end{bmatrix} \begin{Bmatrix} \{\dot{q}_b\} \\ \{\dot{w}_v\} \end{Bmatrix} \\ & + \begin{bmatrix} [\omega_b^2] + [\Phi_b]^T [K_{b-b}] [\Phi_b] & [\Phi_b]^T [K_{b-v}] \\ [K_{v-b}] [\Phi_b] & [K_v] \end{bmatrix} \begin{Bmatrix} \{q_b\} \\ \{w_v\} \end{Bmatrix} = \begin{Bmatrix} [\Phi_b]^T \{f_{b-r}\} \\ \{f_{b-r}\} + \{G\} \end{Bmatrix} \end{aligned} \quad (25)$$

The approach above is adopted by Deng & Cai (2010) and Henchi et al (1998) that combine DOFs of the bridge in the modal space and DOFs of the vehicle derived from a Lagrange formulation. Henchi et al (1998) use a central difference method to solve the coupled equations which finds computationally more efficient than an uncoupled approach solved using Newmark- β . However, some authors note some disadvantages of this method regarding computational effort as a result of properties of the vehicle and bridge matrixes such as symmetry being lost in the process, the need to update the coefficients matrices of the system at every time step with new positions of the vehicle on the bridge, or to carry new calculations of coefficients if the vehicle or bridge model changed (Yang & Lin, 1995; Yang & Fonder, 1996; Kirkegaard et al., 1997).

An alternative approach that leads to considerable savings in computational time is to implement the interaction on an element level rather than on a global level (Olsson, 1985; Yang & Lin, 1995; Yang & Yau, 1997; Yang et al., 1999). An interaction element is defined as that bridge element in

contact with a vehicle wheel. Those bridge elements that are not directly under the action of vehicle wheels remain unaltered in the global matrixes of the system. The interaction element is characterised by two sets of equations: those of the bridge element and those of the moving vehicle above the bridge element. The DOFs of the moving vehicle can be solved in time domain using finite-difference equations of the Newmark type, and then, the DOFs of the vehicle that are not in direct contact with the bridge are eliminated and condensed to the DOFs of the associated bridge element via the method of dynamic condensation. The interaction element has the same number of DOFs as the original bridge element and it can be directly assembled with the other bridge elements into the global matrixes, while retaining properties that are lost when condensation takes place on a global level.

5.3 Algorithms based on Lagrange Multipliers

The system of equations (24) can be extended with a number of additional coordinates to maintain symmetry. The differential equations of motion of the vehicle-bridge system are formed by the equations of the bridge and vehicle which appear without overlapping in the system matrix, and a number of additional constraints that enforce the compatibility condition at the contact points of vehicle and bridge. An advantage of this approach is the ease to accommodate it to a standard FE package (to qualify, the software should have the capability for solving a transient dynamic response, and to allow for direct inputs into the system matrix) while a drawback is the number of additional rows and columns in the system matrix that may require a significant computational effort.

Compatibility conditions between the vertical displacement $u_i(t)$ of each moving wheel j and the vertical displacement of the bridge at the contact point $w_b(x_j, t)$, are established for any time t with the use of time-dependent functions ($A_{ij}(t)$ and $B_{ij}(t)$ which vary for each node i and load j at each instant t). These auxiliary functions produce a constraint consistent with the shape functions of the bridge elements and they facilitate to formulate the compatibility condition at the contact point of the wheel j as:

$$u_j(t) = \sum_{i=1}^N A_{ij}(t)w_{b,i}(t) + \sum_{i=1}^N B_{ij}(t)\theta_{b,i}(t); \quad j = 1, 2, \dots, p \quad (26)$$

where $w_{b,i}(t) (=w_b(x_i, t))$ and $\theta_i(t)$ are the displacements and rotations of the node i in the bridge FE model, and N and p are the total number of bridge nodes and moving wheels respectively. $A_{ij}(t)$ and $B_{ij}(t)$ can be completely defined prior to the simulation once the velocity of the moving loads, the approach length, the spacings between loads and the coordinates of the bridge nodes on the wheel path are known. For illustration purposes only, the shape of these functions for bridge beam elements with nodes numbered 5, 6 and 7 and moving load 1 is shown in Fig. 7 (The values of $B_{ij}(t)$ will depend on the distance between consecutive nodes, Δx). They have zero value outside the interval between adjacent nodes, and for the location of a vehicle wheel j over a bridge node i , $A_{ij}(t)$ and $B_{ij}(t)$ become 1 and 0 respectively, which satisfies $u_j(t)=w_{b,i}(t)$. The numerical expression for these auxiliary functions can be found in Cifuentes (1989) and González et al (2008a).

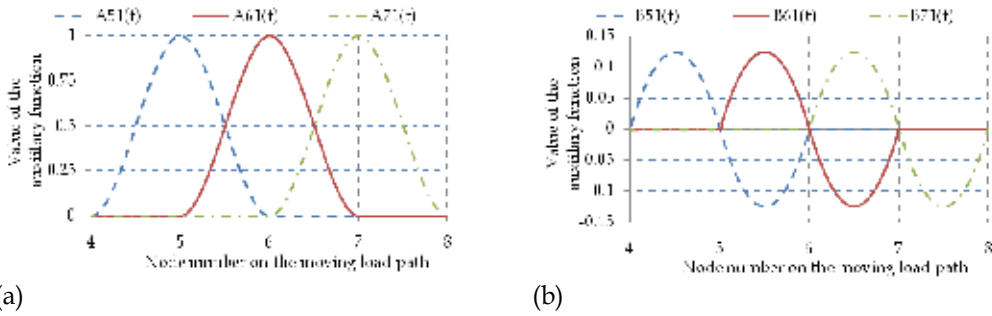


Fig. 7. Auxiliary Functions: (a) $A_{ij}(t)$, (b) $B_{ij}(t)$ ($\Delta x=1$)

The vector of equivalent nodal forces $\{f_b\}$ composed of forces $f_i(t)$ and moments $M_i(t)$ acting on a bridge node i at time t can also be expressed as a function of the p interaction forces using the auxiliary functions:

$$f_i(t) = \sum_{j=1}^p A_{ij}(t)(R_j + G_j) ; \quad M_i(t) = \sum_{j=1}^p B_{ij}(t)(R_j + G_j) ; \quad i = 1, 2, \dots, N \quad (27)$$

where R_j and G_j represent dynamic and static components respectively of the moving force j . The final system matrix is given in Equation (28) for a discretized beam bridge model with N nodes and p moving forces (Baumgärtner, 1999). The group of equations in the first row of the system matrix represents the motion of the bridge. The forcing term corresponding to this first row is expressed as a function of the auxiliary functions as defined by Equation (27). The second row assigns values of the auxiliary functions to scalar points ($\{s\}$) and it has a size equal to the number of bridge nodes in each wheel path multiplied by twice the number of wheels (A_{ij} and B_{ij}). The third and fourth rows of the system matrix represent the equations of motion of the vehicle. The equations of motion relating the DOFs of the wheel contact points ($=p$) to the interaction forces $\{R_j\}$ are given in the third row while the equations of motion of the remaining DOFs of the vehicle are given in the fourth row. In the fifth row, the only terms that are not zero are those that impose the constraint condition between deflections of the moving wheels and the bridge (Equation (26)). When comparing the system matrix of Equation (28) to the one in Equation (24), it can be observed that the system matrix is larger, but it is not time dependent. Therefore, in Equation (24) the forcing vector contains only forces and the vector of unknowns contains only displacements while in Equation (28) these vectors contain both forces and displacements. This coupling between vectors requires an iteration procedure at each point in time until achieving convergence. Details of the implementation of this technique using a fully computerized approach where the bridge and vehicle models are built using NASTRAN are given by González (2010).

$$\left(\begin{array}{ccccc}
 [M_b] \frac{d^2}{dt^2} + [K_b] & 0 & 0 & 0 & 0 \\
 0 & [I] & 0 & 0 & 0 \\
 0 & 0 & [M_v] \frac{d^2}{dt^2} + [C_v] \frac{d}{dt} + [K_v] & [I] & 0 \\
 0 & 0 & & 0 & 0 \\
 0 & 0 & [I] & 0 & 0
 \end{array} \right) \left(\begin{array}{c}
 \left\{ \begin{array}{c} w_{b,1} \\ \dots \\ w_{b,N} \end{array} \right\} \\
 \left\{ \begin{array}{c} s_{11} \\ s_{12} \\ \dots \\ s_{pN} \end{array} \right\} \\
 \left\{ \begin{array}{c} u_1 \\ \dots \\ u_p \end{array} \right\} \\
 \left\{ \begin{array}{c} w_v \\ \dots \\ R_p \end{array} \right\}
 \end{array} \right) = \left(\begin{array}{c}
 \left\{ \begin{array}{c} \Sigma s_{1j} (R_j + G_j) \\ \dots \\ \Sigma s_{Nj} (R_j + G_j) \end{array} \right\} \\
 \left\{ \begin{array}{c} A_{11}(t) \\ A_{12}(t) \\ \dots \\ A_{pN}(t) \end{array} \right\} \\
 \left\{ \begin{array}{c} 0 \\ \dots \\ 0 \end{array} \right\} \\
 \{0\} \\
 \left\{ \begin{array}{c} \Sigma s_{1i} w_{b,i} \\ \dots \\ \Sigma s_{pi} w_{b,i} \end{array} \right\}
 \end{array} \right)
 \quad (28)$$

6. Conclusions

This chapter has reviewed the techniques used for simulating the response of a bridge to the passage of a road vehicle, from the initial stages of preparation of the FE models of bridge and vehicle to the implementation of the interaction between both. VBI simulations can be used to compare the performance of alternative bridge designs or retrofitting options when traversed by traffic and to identify those bridge solutions less prone to dynamic excitation. They can also be employed to quantify the increase in dynamic amplification due to a rougher profile or a deteriorated expansion joint. Other relevant topics in VBI research concern traffic properties, i.e., suspension and tyre types, speed, acceleration/deceleration, vehicle weights and configuration, frequency matching or multiple vehicle presence.

A direct application of VBI modelling is related to the characterization of the total traffic load on a bridge. The traffic load model specified in the bridge design codes is a function of the bridge span length (or natural frequency), number of lanes or load effect, and it is necessarily conservative due to the uncertainties governing the bridge response to a critical traffic loading event prior to its construction. For instance, the critical traffic loading condition is likely to consist of vehicle configurations with very closely spaced axles, which generally tend to produce relative dynamic increments smaller than configurations with longer axle spacings. Additionally, in the case of long span bridges, the critical traffic loading condition consists of a traffic jam situation where vehicles travel at low speed and typically cause a minor dynamic excitation of the bridge. In the case of short-span bridges, the critical situation will be composed of a reduced number of vehicles travelling at highway speed that cause dynamic amplification factors very sensitive to the condition of the road profile before and on the bridge. Thus, bridge codes recommend higher dynamic allowances for shorter spans. Nevertheless, a number of authors have shown that the site-specific dynamic amplification derived from VBI simulation models and field data can represent a significant reduction with respect to these recommendations and save an existing bridge from unnecessary rehabilitation or demolition (Gonzalez et al., 2008a; OBrien et al., 2009).

Other applications include the analysis of bridge-friendly truck suspensions (Green et al., 1995; Harris et al., 2007), investigations regarding the occupants of the vehicle and riding comfort (Esmailzadeh & Jalili, 2003; Yang et al., 2004a; Moghimi & Ronagh, 2008b), interaction with the ground/substructure and induced vibrations (Yang et al., 2004a; Zhang et al., 2005; Chen et al., 2007; Yau, 2009), the effect of structural deterioration (Law & Zhu, 2004; Zhu & Law, 2006) or environmental conditions (Xu & Guo, 2003; Cai & Chen, 2004) on the bridge response to traffic, development of structural monitoring techniques (Yang et al., 2004b), testing of bridge weigh-in-motion (González, 2010) or moving force identification algorithms (González et al., 2008b), fatigue (Chang et al., 2003), vibration control (Kwon et al., 1998), noise (Chanpheng et al., 2004) and ambient vibration analysis (Kim et al., 2003).

7. References

- ASCE (1931). *Impact on Highway Bridges*. Final report of the Special Committee on Highway Bridges, Transactions ASCE, Vol. 95, 1089-1117
- Barth, K.E. & Wu, H. (2007). Development of improved natural frequency equations for continuous span steel I-girder bridges. *Engineering Structures*, Vol. 29, 3432-3442, ISSN 0141-0296
- Baumgärtner, W. (1999). Bridge-vehicle interaction using extended FE analysis. *Heavy Vehicle Systems, International Journal of Vehicle Design*, Vol. 6, Nos. 1-4, 1-12, ISSN 1744-232X
- Billing, J.R. (1984). Dynamic loading and testing of bridges in Ontario. *Canadian Journal of Civil Engineering*, Vol. 11, 833-843, ISSN 1208-6029
- Cai, C.S. & Chen, S.R. (2004). Framework of vehicle-bridge-wind dynamic analysis. *Journal of Wind Engineering and Industrial Aerodynamics*, Vol. 92, 579-607, ISSN 0167-6105
- Cai, C.S.; Shi, X.M.; Araujo, M. & Chen, S.R. (2007). Effect of approach span condition on vehicle-induced dynamic response of slab-on-girder road bridges. *Engineering Structures*, Vol. 29, 3210-3226, ISSN 0141-0296
- Cantero, D.; O'Brien, E.J.; González, A.; Enright, B. & Rowley, C. (2009). Highway bridge assessment for dynamic interaction with critical vehicles, *Proceedings of 10th International Conference on Safety, Reliability and Risk of Structures*, ICOSSAR, pp. 3104-3109, ISBN 978-0-415-47557-0, Osaka, Japan, September 2009, Taylor & Francis
- Cantero, D.; O'Brien, E.J. & González, A. (2010). Modelling the vehicle in vehicle-infrastructure dynamic interaction studies. *Proceedings of the Institution of Mechanical Engineers, Part K, Journal of Multi-body Dynamics*, Vol. 224(K2), 243-248, ISSN 1464-4193
- Cantieni, R. (1983). Dynamic load testing of highway bridges in Switzerland, 60 years experience of EMPA. *Report no. 211*. Dübendorf, Switzerland
- Cebon, D. (1999). *Handbook of Vehicle-Road Interaction*, Swets & Zeitlinger B.V., ISBN 9026515545, The Netherlands
- Cebon, D. & Newland, D.E. (1983). Artificial generation of road surface topography by the inverse F.F.T. method. *Vehicle System Dynamics*, Vol. 12, 160-165, ISSN 0042-3114
- Chan, T.H.T.; Guo, L. & Li, Z.X. (2003). Finite element modelling for fatigue stress analysis of large suspension bridges. *Journal of Sound and Vibration*, Vol. 261, 443-464, ISSN 0022-460X

- Chanpheng, T.; Yamada, H.; Miyata, T. & Katsuchi, H. (2004). Application of radiation modes to the problem of low-frequency noise from a highway bridge. *Applied Acoustics*, Vol. 65, No. 2, 109-123, ISSN 0003-682X
- Chatterjee, P.K.; Datta, T.K. & Surana, C.S. (1994a). Vibration of suspension bridges under vehicular movement. *Journal of Structural Engineering*, Vol. 120, No. 3, 681-703, ISSN 0733-9445
- Chatterjee, P.K.; Datta, T.K. & Surana, C.S. (1994b). Vibration of continuous bridges under moving vehicles. *Journal of Sound and Vibration*, Vol. 169, No. 5, 619-632, ISSN 0022-460X
- Chen, Y.-J.; Ju, S.-H.; Ni, S.-H. & Shen, Y.-J. (2007). Prediction methodology for ground vibration induced by passing trains on bridge structures. *Journal of Sound and Vibration*, Vol. 302, Nos. 4-5, 806-820, ISSN 0022-460X
- Chompooming, K. & Yener, M. (1995). The influence of roadway surface irregularities and vehicle deceleration on bridge dynamics using the method of lines. *Journal of Sound and Vibration*, Vol. 183, No. 4, 567-589, ISSN 0022-460X
- Cifuentes, A.O. (1989). Dynamic response of a beam excited by a moving mass. *Finite Elements in Analysis and Design*, Vol. 5, 237-46, ISSN 0168-874X
- Clough, W.R. & Penzien, J. (1993). *Dynamics of Structures, 2nd Edition*, McGraw-Hill, ISBN 0-07-011394-7, New York
- Deng, L. & Cai, C.S. (2010). Development of dynamic impact factor for performance evaluation of existing multi-girder concrete bridges. *Engineering Structures*, Vol. 32, 21-31, ISSN 0141-0296
- ElMadany, M.M. (1988). Design optimization of truck suspensions using covariance analysis. *Computers and Structures*, Vol. 28, No. 2, 241-246, ISSN 0045-7949
- Esmailzadeh, E. & Jalili, N. (2003). Vehicle-passenger-structure interaction of uniform bridges traversed by moving vehicles. *Journal of Sound and Vibration*, Vol. 260, No. 4, 611-635, ISSN 0022-460X
- Fafard, M.; Bennur, M. & Savard, M. (1997). A general multi-axle vehicle model to study the bridge-vehicle interaction. *Engineering Computations*, Vol. 14, No. 5, 491-508, ISSN 072773539X
- Fryba, L. (1972). *Vibration of Solids and Structures under Moving Loads*, Noordhoff International Publishing, ISBN 9001324202, Groningen, The Netherlands
- Fu, T.T. & Cebon, D. (2002). Analysis of a truck suspension database. *International Journal of Heavy Vehicle Systems*, Vol. 9, No. 4, 281-297, ISSN 1744-232X
- González, A.; Rattigan, P.; O'Brien, E.J. & Caprani, C. (2008a). Determination of bridge lifetime DAF using finite element analysis of critical loading scenarios. *Engineering Structures*, Vol. 30, No. 9, 2330-2337, ISSN 0141-0296
- González, A.; Rowley, C. & O'Brien, E.J. (2008b). A general solution to the identification of moving vehicle forces on a bridge. *International Journal for Numerical Methods in Engineering*, Vol. 75, No. 3, 335-354, ISSN 0029-5981
- González, A. (2010). *Development of a Bridge Weigh-In-Motion System*, Lambert Academic Publishing, ISBN 978-3-8383-0416-8, Germany
- González, A.; O'Brien, E.J.; Cantero, D.; Yingyan, L.; Dowling, J. & Žnidarič, A. (2010). Critical speed for the dynamics of truck events on bridges with a smooth surface. *Journal of Sound and Vibration*, Vol. 329, No. 11, 2127-2146, ISSN 0022-460X

- Green, M.F. (1993). Bridge dynamics and dynamic amplification factors – a review of analytical and experimental findings. Discussion. *Canadian Journal of Civil Engineering*, Vol. 20, 876-877, ISSN 1208-6029
- Green, M.F. & Cebon, D. (1994). Dynamic response of highway bridges to heavy vehicle loads: theory and experimental validation. *Journal of Sound and Vibration*, Vol. 170, 51-78, ISSN 0022-460X
- Green, M.F.; Cebon, D. & Cole, D.J. (1995). Effects of vehicle suspension design on dynamics of highway bridges. *Journal of Structural Engineering*, Vol. 121, No. 2, 272-282, ISSN 0733-9445
- Green, M.F. & Cebon, D. (1997). Dynamic interaction between heavy vehicles and highway bridges. *Computers and Structures*, Vol. 62, No. 2, 253-264, ISSN 0045-7949
- Guo, W.H. & Xu, Y.L. (2001). Fully computerized approach to study cable-stayed bridge-vehicle interaction. *Journal of Sound and Vibration*, Vol. 248, No. 4, 745-761, ISSN 0022-460X
- Hambly, E.C. (1991). *Bridge Deck Behaviour*, Taylor & Francis, ISBN 0-419-17260-2, New York
- Harris, N.K.; O'Brien, E.J. & González, A. (2007). Reduction of bridge dynamic amplification through adjustment of vehicle suspension damping. *Journal of Sound and Vibration*, Vol. 302, 471-485, ISSN 0022-460X
- Henchi, K.; Fafard, M.; Dhatt, G. & Talbot, M. (1997). Dynamic behaviour of multi-span beams under moving loads. *Journal of Sound and Vibration*, Vol. 199, No. 1, 33-50, ISSN 0022-460X
- Henchi, K.; Fafard, M.; Talbot, M. & Dhatt, G. (1998). An efficient algorithm for dynamic analysis of bridges under moving vehicles using a coupled modal and physical components approach. *Journal of Sound and Vibration*, Vol. 212, No. 4, 663-683, ISSN 0022-460X
- Heywood, R.; Roberts, W. & Bouilly, G. (2001). Dynamic loading of bridges. *Transportation Research Record, Journal of the Transportation Research Board*, Vol. 1770, 58-66, ISSN 0361-1981
- Huang, D.; Wang, T.-L. & Shahawy, M. (1992). Impact analysis of continuous multigirder bridges due to moving vehicles. *Journal of Structural Engineering*, Vol. 118, No. 12, 3427-3443, ISSN 0733-9445
- Hwang, E.-S. & Nowak, A.-S. (1991). Simulation of dynamic load for bridges. *Journal of Structural Engineering*, Vol. 117, No. 5, 1413-1434, ISSN 0733-9445
- Inglis, C.E. (1934). *A Mathematical Treatise on Vibration in Railway Bridges*, Cambridge University Press, Cambridge, UK
- ISO 8608 (1995). Mechanical vibration - road surface profiles - reporting of measured data, BS 7853:1996, ISBN 0-580-25617-0, London
- Jeffcott, H.H. (1929). On the vibrations of beams under the action of moving loads. *Phil. Magazine*, Series 7, No. 8(48), 66-97
- Ju, S.-H. & Lin, H.-T. (2007). A finite element model of vehicle-bridge interaction considering braking and acceleration. *Journal of Sound and Vibration*, Vol. 303, Nos. 1-2, 46-57, ISSN 0022-460X
- Kim, C.W.; Kawatani, M. & Kim, K.B. (2005). Three-dimensional dynamic analysis for bridge-vehicle interaction with roadway roughness. *Computers and Structures*, Vol. 83, 1627-1645, ISSN 0045-7949

- Kim, C.Y.; Jung, D.S.; Kim, N.-S.; Kwon, S.-D. & Feng, M.Q. (2003). Effect of vehicle weight on natural frequencies of bridges measured from traffic-induced vibrations. *Earthquake Engineering and Engineering Vibration*, Vol. 2, No. 1, 109-115, ISSN 1671-3664
- Kirkegaard, P.H.; Nielsen, S.R.K. & Enevoldsen, I. (1997). Heavy vehicles on minor highway bridges - Dynamic modelling of vehicles and bridges. Paper No. 172, Dept. of Building Technology and Structural Engineering, Aalborg University, Aalborg, Denmark, ISBN 1395-7953
- Kwasniewski, L.; Li, H.; Wekezer, J. & Malachowski, J. (2006). Finite element analysis of vehicle-bridge interaction. *Finite Elements in Analysis and Design*, Vol. 42, 950-959, ISSN 0168-874X
- Kwon, H.-C.; Kim, M.-C. & Li, I.-W. (1998). Vibration control of bridges under moving loads. *Computers and Structures*, Vol. 66, No. 4, 473-480, ISSN 0045-7949
- Law, S.S. & Zhu, X.Q. (2004). Dynamic behaviour of damaged concrete bridge structures under moving loads. *Engineering Structures*, Vol. 26, 1279-1293, ISSN 0141-0296
- Law, S.S. & Zhu, X.Q. (2005). Bridge dynamic responses due to road surface roughness and braking of vehicle. *Journal of Sound and Vibration*, Vol. 282, Nos. 3-5, 805-830, ISSN 0022-460X
- Liu, C.; Huang, D. & Wang, T.-L. (2002). Analytical dynamic impact study based on correlated road roughness. *Computers and Structures*, Vol. 80, 1639-1650, ISSN 0045-7949
- Moghim, H. & Ronagh H.R. (2008a). Impact factors for a composite steel bridge using non-linear dynamic simulation. *International Journal of Impact Engineering*, Vol. 35, No. 11, 1228-1243, ISSN 0734-743X
- Moghim, H. & Ronagh H.R. (2008b). Development of a numerical model for bridge-vehicle interaction and human response to traffic-induced vibration. *Engineering Structures*, Vol. 30, No. 12, 3808-3819, ISSN 0141-0296
- Nassif, H.H. & Liu, M. (2004). Analytical modeling of bridge-road-vehicle dynamic interaction system. *Journal of Vibration and Control*, Vol. 10, No. 2, 215-241, ISSN 1077-5463
- O'Brien, E.J. & Keogh, D. (1999). *Bridge Deck Analysis*, E & FN Spon, Taylor & Francis, ISBN 0-419-22500-5, London
- O'Brien, E.J.; Rattigan, P.; González, A.; Dowling, J. & Žnidarič, A. (2009). Characteristic dynamic traffic load effects in bridges. *Engineering Structures*, Vol. 31, No. 7, 1607-1612, ISSN 0141-0296
- Olsson, M. (1985). Finite element, modal co-ordinate analysis of structures subjected to moving loads. *Journal of Sound and Vibration*, Vol. 99, No. 1, 1-12, ISSN 0022-460X
- Stokes, G.G. (1849). Discussion of a differential equation relating to the breaking of railway bridges. *Transactions Cambridge Philosophic Society*, Part 5, 707-735
- Tan, G.H.; Brameld, G.H. & Thambiratnam, D.P. (1998). Development of an analytical model for treating bridge-vehicle interaction. *Engineering Structures*, Vol. 20, Nos. 1-2, 54-61, ISSN 0141-0296
- Tilly, G.P. (1986). Dynamic behaviour of concrete structures, in *Developments in civil engineering*, Vol. 13, Report of the Rilem 65MDB Committee. Elsevier, New York
- Timoshenko, S. & Young, D.H. (1955). *Vibration Problems in Engineering*, D. Van Nostrand Company, Inc., ISBN 978-1443731676, New York

- Veletsos, A.S. & Huang, T. (1970). Analysis of dynamic response of highway bridges. *ASCE Engineering Mechanics*, Vol. 96, No. EM5, Ref. 35, 593-620, ISSN 0733-9402
- Wang, T.-L. & Huang, D. (1992). Cable-stayed bridge vibration due to road roughness. *Journal of Structural Engineering*, Vol. 118, No. 5, 1354-1374, ISSN 0733-9445
- Wang, T.-L.; Huang, D.; Shahawy, M. & Huang, K. (1996). Dynamic response of highway girder bridges. *Computers and Structures*, Vol. 60, 1021-1027, ISSN 0045-7949
- West, R. (1973). Recommendations on the use of grillage analysis for slab and pseudo-slab bridge decks, Cement and Concrete Association and Construction Industry Research and Information Association, London, UK
- Willis, R. (1849). *Appendix Report of the Commissioners Appointed to Inquire into the Application of Iron to Railway Structures*, William Clowes, HM Stationary Office, London
- Wong, J.Y. (1993). *Theory of Ground Vehicles*, John Wiley & Sons, ISBN 978-0-471-52496-0, New York
- Wright, D.T. & Green, R. (1963). *Highway Bridge Vibrations, Part 2. Ontario Test Programme*, Report 5, Department of Civil Engineering, Queen's University, Kingston
- Xu, Y.L. & Guo, W.H. (2003). Dynamic analysis of coupled road vehicle and cable-stayed bridge systems under turbulent wind. *Engineering Structures*, Vol. 25, 473-486, ISSN 0141-0296
- Yang, F. & Fonder, G. A. (1996). An iterative solution method for dynamic response of bridge-vehicles systems. *Earthquake Engineering and Structural Dynamics*. Vol. 25, 195-215, ISSN 0098-8847
- Yang, Y.-B. & Lin, B.-H. (1995). Vehicle-bridge interaction analysis by dynamic condensation method. *Journal of Structural Engineering*, Vol. 121, No. 11, 1636-1643, ISSN 0733-9445
- Yang, Y.-B. & Yau, J.-D. (1997). Vehicle-bridge interaction element for dynamic analysis. *Journal of Structural Engineering*. Vol. 123, No. 11, 1512-1518, ISSN 0733-9445
- Yang, Y.-B.; Chang, C.H. & Yau, J.D. (1999). An element for analyzing vehicle-bridge systems considering vehicle's pitching effect. *International Journal for Numerical Methods in Engineering*, Vol. 46, 1031-1047, ISSN 0029-5981
- Yang, Y.-B.; Yau, J.D. & Wu, Y.S. (2004a). *Vehicle-Bridge Interaction Dynamics. with Applications to High-Speed Railways*, World Scientific Publishing Co., ISBN 981-283-847-8, Singapore
- Yang, Y.-B.; Lin, C.W. & Yau, J.D. (2004b). Extracting bridge frequencies from the dynamic response of a passing vehicle. *Journal of Sound and Vibration*, Vol. 272, Nos. 3-5, 471-493, ISSN 0022-460X
- Yau, J.D. (2009). Response of a train moving on multi-span railway bridges undergoing ground settlement. *Engineering Structures*, Vol. 31, 2115-2122, ISSN 0141-0296
- Zhang, N.; Xia, H. & Zhan, J.W. (2005). Dynamic analysis of vehicle-bridge-foundation interaction system, In *Environmental Vibrations*, Takemiya, H.; Kogakkai, J.; Gakkai, D. & Gkkai N.K., 137-144, Taylor & Francis Group, ISBN 0 415 39035 4, London
- Zhu, X.Q. & Law, S.S. (2002). Dynamic load on continuous multi-lane bridge deck from moving vehicles. *Journal of Sound and Vibration*, Vol. 251, No. 4, 697-716, ISSN 0022-460X
- Zhu, X.Q. & Law, S.S. (2006). Wavelet-based crack identification of bridge beam from operational deflection time history. *International Journal of Solids and Structures*, Vol. 43, 2299-2317, ISSN 0020-7683

Finite Element Modelling of Sound Transmission Loss in Reflective Pipe

Tomasz Strek
Poznan University of Technology
Institute of Applied Mechanics
Poland

1. Introduction

Acoustics is the physics of sound. Sound is the sensation, as detected by the ear, of very small rapid changes in the air pressure above and below a static value. This static value is atmospheric pressure (about 100,000 Pascals), which varies slowly. Associated with a sound pressure wave is a flow of energy. Physically, sound in air is a longitudinal wave where the wave motion is in the direction of the movement of energy. The wave crests are the pressure maxima, while the troughs represent the pressure minima (Comsol, 2007a).

Sound results when the air is disturbed by some source. An example is a vibrating object, such as a speaker cone in a hi-fi system. It is possible to see the movement of a bass speaker cone when it generates sound at a very low frequency. As the cone moves forward, it compresses the air in front of it, causing an increase in air pressure. Then it moves back past its resting position and causes a reduction in air pressure. This process continues, radiating a wave of alternating high and low pressure at the speed of sound.

Plane waves of constant frequency propagating through bulk materials have amplitudes that typically decrease exponentially with increasing propagation distance, such that the magnitude of the complex pressure amplitude varies as (Rossing, 2007)

$$|p(x)| = |p(0)|e^{-\alpha x}. \quad (1)$$

The quantity α is the plane wave attenuation coefficient and has units of nepers per meter (Np/m). It is an intrinsic frequency-dependent property of the material. This exponential decrease of amplitude is called attenuation or absorption of sound and is associated with the transfer of acoustic energy to the internal energy of the material. If $|p|^2$ decreases to a tenth of its original value, it is said to have decreased by 10 decibels (dB), so an attenuation constant of α nepers per meter is equivalent to an attenuation constant of $[20/(\ln 10)] \alpha$ decibels per meter, or 8.6859α decibels per meter.

The attenuation of sound due to the classical processes of viscous energy absorption and thermal conduction is derivable from the dissipative wave equation given previously for the acoustics mode (Rossing, 2007). In air, the relaxation processes that affect sound attenuation

are those associated with the (quantized) internal vibrations of the diatomic molecules O_2 and N_2 . The presence of the duct walls affects the attenuation of sound and usually causes the attenuation coefficient to be much higher than for plane waves in open space.

The problem of steady-state diffraction or transmission of sound energy by an aperture in a plane wall has attracted much attention in the literature for many years (Wilson and Soroka, 1965). Exact solutions are restricted to a few cases where the aperture geometry is simple and may be conveniently described in a coordinate system in which the wave equation becomes separable, or may allow the postulation of a set of velocity potentials that can be made to fit the boundaries. The theoretical methods available for approximating these solutions are valid for only a limited range of frequencies or wavelengths. The exact solutions are also of limited range in application because of practical difficulties in evaluating the infinite series occurring in the solutions. A frequency dependent model for the transmission loss of hard-walled pipe or duct is important for accurate estimates of more complicated acoustic systems (for instance, silencer-pipe connections).

Theoretical solutions of two- and three-dimensional acoustic wave propagation in hard-walled (Cummings, 1974; Rostafinski, 1974) and acoustically lined (Ko & Ho, 1977; Ko, 1979) ducts with rectangular cross-section have been presented by others researchers. In these studies, cylindrical coordinate system has been used to describe the curved duct. The theoretical solution for a circular section curved duct has not been achieved yet due to the mathematical difficulties encountered in the solution of the wave equation.

In paper Wilson and Soroka (1965) an approximate solution for the diffraction of a planes sound wave incident normally on a circular aperture in a plane rigid wall of finite thickness is obtained by postulating rigid, massless, infinitely thin plane pistons in each end of the aperture, whose motions simulate the movement of the air particles at these positions under acoustic excitation. In the paper (Chen et al., 2006) the improvement on the acoustic transmission loss of a duct by adding some Helmholtz resonator is discussed. Therefore, the calculation on the transmission loss of a duct in a rigid wall by modifying the formula derived by Wilson and Soroka were done.

In paper Sarigug (1999) acoustic surface pressures of various pipes in the shape of a quarter torus were calculated. Sound attenuation spectra of pipes possessing different bend sharpnesses were presented.

A predictive model for sound propagation in tubes with permeable walls was presented in paper Cummings and Kirby (1999). Tubes with permeable walls are used in various applications. The propagation model was verified by comparison to experimental data on a perforated metal tube and was applied to a practical type of permeable fabric tube.

The study of the refraction and transmission coefficients of acoustic waves in ducts of continuously varying cross-sectional area has been of interest to many researchers in the past. Various numerical methods for predicting these coefficients have been developed, such as the method of weighted residuals, the finite element method, the perturbation method, the boundary element method, and the matricial Riccati equation method (Utsumi, 2001). A method of solution based on a variational principle has been presented in paper (Utsumi, 2001) for the acoustic wave propagation in ducts having a continuous change in cross-sectional area.

Finite-element methods for time-harmonic acoustics governed by the reduced wave equation (Helmholtz equation) have been an active research area for nearly 40 years. Initial applications of finite-element methods for time-harmonic acoustics focused on interior

problems with complex geometries including direct and modal coupling of structural acoustic systems for forced vibration analysis, frequency response of acoustic enclosures, and waveguides (Thompson, 2006).

There are lot of papers presenting experimental and numerical methods to predict sound transmission loss in silencers. In general, only simple shaped silencers may be analyzed by conventional analytical approach. The silencers of complicated shape with or without acoustic internal structures have been analyzed by transfer matrix method, FEM (Finite Element Method), the transfer matrix method with BEM, multi-domain BEM (Boundary Element Method), multi-domain structural-acoustic coupling analysis, etc (Ju & Lee, 2005). Among these, the transfer matrix method with BEM has been used to calculate the transmission or insertion loss of the silencers. In this method, the particle velocities and sound pressures of each domain of the silencer are calculated by BEM and the four pole parameters of every domain are calculated.

The calculation of attenuation of sound in the acoustic systems with complicated internal structures by the conventional BEM combined with the transfer matrix method is incorrect at best or impossible for 3-dimensional domains due to its inherent plane wave assumption.

On this consideration, in this chapter it is proposed an efficient practical numerical method (based on FEM) for calculation of attenuation of sound within duct/pipe and the whole acoustic structure. The transmission loss estimation by the proposed numerical method is tested by comparison with the experimental one on an sound attenuation in reflective pipe. The method shows its viability by presenting the reasonably consistent anticipation of the experimental results.

2. Sound attenuation

A lot of devices with duct systems like air-condition, heating equipment, frig and so on are being used in people's life. External boundaries can channel sound propagation, and in some cases can create buildup or attenuation of acoustic energy within a confined space. Pipes or ducts acts as guides of acoustic waves, and the net flow of energy, other than that associated with wall dissipation, is along the direction of the duct. The general theory of guided waves applies and leads to a representation in terms of guided modes.

When sound impinges on a surface of external boundaries, some sound is reflected and some is transmitted and possibly absorbed within or on the other side of the surface. To understand the processes that occur, it is often an appropriate idealization to take the incident wave as a plane wave and to consider the surface as flat.

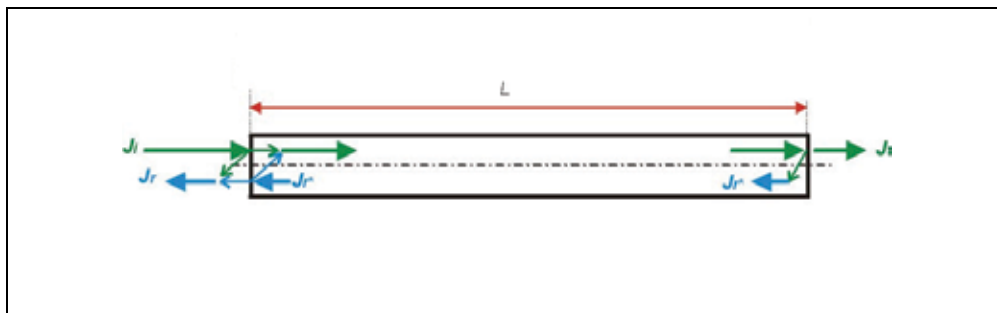


Fig. 1. A description of sound power transmission in the pipe with outlet to free space

When a plane wave reflects at a surface with finite specific acoustic impedance Z_s , a reflected wave is formed such that the angle of incidence Θ equals the angle of reflection (law of mirrors). Here both angles are reckoned from the line normal to the surface and correspond to the directions of the two waves.

In this paper the reflective pipe of length L and radius r with rigid walls is analysed. Fig. 1 shows sound power transmission, where J_i - sound power from source, J_r - sound power reflected from pipe inlet, J_r^* - sound power reflected from pipe outlet (if $L \rightarrow \infty$ then $J_r^* \rightarrow 0$) and J_t - sound power transmitted by outlet outside pipe.

The difference between the sound energy on one side of the pipe and that radiated from the second side (both expressed in decibels) is called the sound transmission loss. The larger the sound transmission loss (in decibels), the smaller the amount of sound energy passing through and consequently, less noise heard. Transmission loss depends on frequency.

For the ideal case, when there is no ambient velocity and when viscosity and thermal conduction are neglected, the energy density in considered domain at given time is described by (Rossing, 2007)

$$w = \frac{1}{2} \left(\rho v^2 + \frac{p^2}{\rho c^2} \right). \quad (2)$$

The first term in the above expression for energy density is recognized as the acoustic kinetic energy per unit volume, and the second term is identified as the potential energy per unit volume due to compression of the fluid.

The acoustic intensity is described by

$$\mathbf{I} = p \mathbf{v} \quad (3)$$

where \mathbf{I} is an intensity vector or energy flux vector, p is the pressure and \mathbf{v} is the fluid velocity.

For a plane wave the kinetic and potential energies are the same, and that the energy density is given by

$$w = \frac{p^2}{\rho c^2} \quad (4)$$

and the intensity becomes

$$\mathbf{I} = \mathbf{n} \frac{p^2}{\rho c}. \quad (5)$$

For such a case, the intensity and the energy density are related by

$$\mathbf{I} = c \mathbf{n} w. \quad (6)$$

This yields the interpretation that the energy in a sound wave is moving in the direction of propagation with the sound speed. Consequently, the sound speed can be regarded as an energy propagation velocity (Rossing, 2007).

The human ear can generally perceive sound pressures over the range from about 20 μ Pa up to about 200 Pa. Because the range of typical acoustic pressures is so large, it is convenient to work with a relative measurement scale rather than an absolute measurement scale. These scales are expressed using logarithms to compress the dynamic range (Rossing, 2007).

One can write the sound intensity level (SIL) as the logarithm of two intensities

$$SIL = 10 \text{Log} \left(\frac{I}{I_{ref}} \right) \quad (7)$$

where I is the intensity of sound wave and I_{ref} is the reference intensity. For the intensity of a sound wave in air, the reference intensity is defined to be $I_{ref} = 10^{-12} \text{W} / \text{m}^2$. Additionally, the sound pressure level (SPL) is defined as

$$SPL = 20 \text{Log} \left(\frac{p}{p_{ref}} \right) \quad (8)$$

where p is the acoustic pressure and p_{ref} is the reference pressure. For sound in air the reference pressure is defined as 20 μ Pa. Sound levels and parameters of the transmission properties are measured in units of decibels (dB).

There are a few measure parameters of the transmission properties of acoustic devices (i.e. muffler, duct) which are used by engineers. Insertion loss (IL), which is known to be a correct performance measure of a acoustic device, is defined as the difference between the sound powers without and with a device in place. By measuring the sound pressure levels SPL_1 and SPL_2 without and with device, respectively, the insertion loss of the device can be calculated using the following relationship

$$IL = SPL_1 - SPL_2 . \quad (9)$$

A useful measure of the acoustic performance which depends only on the device and not also on connected elements is the frequency dependent sound transmission loss (Byrne et al., 2006). It is defined as the ratio of the sound power incident on the device inlet to that of the sound power leaving the device at the outlet. Transmission loss (TL) also know as Sound Reduction Index (SRI) is given by

$$TL = 10 \log \frac{W_i}{W_t} , \quad (10)$$

where W_i denotes the incoming power at the inlet, W_t denotes the transmitted (outgoing) power at the outlet.

To determine the transmission loss in the above model, one must first calculate the incident and transmitted time-averaged sound intensities and the corresponding sound power values. The incoming and transmitted sound powers are given by equations

$$W_i = \int_A \frac{p_0^2}{2\rho c} dA, \quad W_t = \int_A \frac{|p|^2}{2\rho c} dA, \quad (11)$$

where p_0 represents the applied pressure source amplitude and A - the area of boundary where waves are incoming to the pipe and outgoing from the pipe.

Please notice that incoming sound power is expressed by applied pressure source amplitude and not by pressure at the inlet. Because calculated sound power at the inlet is sum of sound power from the source, sound power reflected from pipe inlet, pipe outlet and pipe boundaries. Effective incoming sound power at inlet is difference of sound power from the source and sound power reflected from pipe inlet. In sound transmission loss incoming sound power is expressed by effective incoming sound power.

Most of the energy entering and leaving the device is carried by plane waves. Hence, the sound energy incident on transmitted from the device can be found from the acoustic pressures associated with the incident and transmitted acoustic waves. If the acoustic pressure time histories of an appropriate transient incident wave and the corresponding transmitted wave are Fourier Transformed. The ratio of the moduli of their transforms are taken and expressed in logarithmic form, the required frequency dependent sound transmission loss of the device can be found. That is

$$TL = 20 \log \frac{FFT_i}{FFT_t} \text{ [dB]} \quad (12)$$

where FFT_i and FFT_t are the Fourier Transforms of the time histories of the incident and transmitted waves, respectively.

3. Theory background – mathematical model and computational method

3.1 Linear acoustic

Theory in this section is based on handbook of acoustic (Rossing, 2007). This book is the usual starting reference for learning the details of acoustic. If viscosity and thermal conductivity are neglected at the outset, then the Navier-Stokes equation reduces to the Euler equation

$$\rho \frac{D\mathbf{v}}{Dt} = -\nabla p + \rho \mathbf{g}. \quad (13)$$

The conservation of mass is described by the partial differential equation

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0. \quad (14)$$

where $\rho = \rho(\mathbf{x}, t)$ is the (possibly position- and time-dependent) mass density and \mathbf{v} is the local and instantaneous particle velocity.

The differential equation of state yields

$$\frac{Dp}{Dt} = c^2 \frac{D\rho}{Dt} \quad (15)$$

where $\frac{D}{Dt} = \left(\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla \right)$. Equation (14) with combination with equation (15) yields

$$\frac{Dp}{Dt} = -\rho c^2 \nabla \cdot \mathbf{v} . \quad (16)$$

The equations that one ordinarily deal with in acoustics are linear in the field amplitudes, where the fields of interest are quantities that depend on position and time.

Sound results from a time-varying perturbation of the dynamic and thermodynamic variables that describe the medium. For sound in fluids (liquids and gases), the quantities appropriate to the ambient medium (i.e. the medium in the absence of a disturbance) are customarily represented by the subscript 0, and the perturbation are represented by a prime on the corresponding symbol.

Thus one expresses the total pressure as

$$p = p_0 + p' \quad (17)$$

with corresponding expressions for fluctuations in specific entropy, fluid velocity and density.

The linearized version of the conservation of mass relation is

$$\frac{\partial \rho'}{\partial t} + \nabla \cdot (\mathbf{v}_0 \rho' + \mathbf{v}' \rho_0) = 0 . \quad (18)$$

Due to restrictions on using linear equations it should be noted that one regards $p' \ll \rho_0 c^2$, $|\mathbf{v}'| \ll c$. It is not necessary that p' be much less than p_0 , and it is certainly not necessary that $|\mathbf{v}'|$ be less than $|\mathbf{v}_0|$.

The equations for linear acoustics neglect dissipation processes and consequently can be derived from the equations for flow of a compressible ideal fluid. If one neglects gravity at the outset, and assumes the ambient fluid velocity is zero, then the ambient pressure is constant. In such a case (16) leads to

$$\frac{\partial p}{\partial t} + \rho c^2 \nabla \cdot \mathbf{v} = 0 . \quad (19)$$

The Euler equation leads to

$$\rho \frac{\partial \mathbf{v}}{\partial t} = -\nabla p. \quad (20)$$

Here a common notational convention is used to delete primes and subscripts. The density ρ here is understood to be the ambient density ρ_0 , while p and \mathbf{v} are understood to be the acoustically induced perturbations to the pressure and fluid velocity.

A single equation for the acoustic part of the pressure results when one takes time derivative of (19) and then uses (20) to express the time derivative of the fluid velocity in terms of pressure. The resulting equations for the case when the density varies with position is

$$\frac{\partial^2 p}{\partial t^2} = \rho c^2 \nabla \cdot \left(\frac{1}{\rho} \nabla p \right), \quad (21)$$

where $K = \rho c^2$ is called the adiabatic bulk modulus. When density is independent from position we have

$$\frac{\partial^2 p}{\partial t^2} = c^2 \nabla^2 p. \quad (22)$$

An important special case is a time-harmonic wave, for which the pressure varies with time as

$$p(\mathbf{x}, t) = p(\mathbf{x}) e^{i\omega t} \quad (23)$$

where $\omega = 2\pi f$ is the angular frequency, with f denoting the frequency. Assuming the same harmonic time dependence for the source terms, the wave equation for acoustic waves reduces to an inhomogeneous Helmholtz equation

$$-\frac{\omega^2 p}{\rho c^2} = \nabla \cdot \left(\frac{1}{\rho} \nabla p \right). \quad (24)$$

For the acoustic pressure in 2D axisymmetric geometries the wave equation becomes

$$\frac{\partial}{\partial r} \left(-\frac{r}{\rho} \frac{\partial p}{\partial r} \right) + r \frac{\partial}{\partial r} \left(-\frac{1}{\rho} \frac{\partial p}{\partial z} \right) - \left[\left(\frac{\omega}{c} \right)^2 - \left(\frac{m}{r} \right)^2 \right] \frac{rp}{\rho} = 0 \quad (25)$$

where m denotes the circumferential wave number and k_z is the out-of-plane wave number. In 2D axisymmetric geometries the independent variables are the radial coordinate r and the axial coordinate z .

Typical boundary conditions are: sound-hard boundaries (walls), sound-soft boundaries (zero acoustic pressure), specified acoustic pressure, specified normal acceleration, impedance boundary conditions and radiation boundary conditions (Comsol, 2007).

3.2 The dimensionless equations

For simplicity the preferred work choice is to work in non-dimensional frame of reference. Now some dimensionless variables will be introduced in order to make the system much easier to study. This procedure is very important so that one can see which combination of parameters is more important than the others.

Non-dimensionalised variables and scales are defined as follow: time $t' = \frac{c}{h}t$; distance

variable $\mathbf{x}' = \frac{1}{h}\mathbf{x}$; pressure $p' = \frac{p}{\rho c^2}$; angular frequency $\omega' = \frac{h}{c}\omega$; monopole source

$Q' = \frac{Q}{Q_0}$ are introduced, where h is characteristic distance (for instance pipe length).

The next step is to put these variables into the wave equation (Comsol, 2007)

$$\frac{1}{\rho c^2} \frac{\partial^2 p}{\partial t^2} + \nabla \cdot \left(-\frac{1}{\rho} \nabla p \right) = Q, \quad (26)$$

what gives

$$\frac{1}{\rho c^2} \frac{\rho c^2}{(h/c)^2} \frac{\partial'^2 p'}{\partial t'^2} + \frac{1}{h} \nabla' \cdot \left(-\frac{1}{\rho} \frac{\rho c^2}{h} \nabla' p' \right) = Q_0 Q, \quad (27)$$

or after mathematical manipulations with $\rho = \text{const}$

$$\left(\frac{c}{h} \right)^2 \frac{\partial'^2 p'}{\partial t'^2} + \left(\frac{c}{h} \right)^2 \nabla'^2 p' = Q_0 Q \quad (28)$$

where Q [$1/s^2$] monopole source.

Then multiply through by $\frac{h^2}{c^2}$ what gives

$$\frac{\partial'^2 p'}{\partial t'^2} + \nabla'^2 p' = \frac{h^2 Q_0}{c^2} Q'. \quad (29)$$

Putting dimensionless variables into the time-harmonic wave equation we get

$$-\omega'^2 p' - \nabla'^2 p' = 0, \quad (30)$$

and into the 2D axisymmetric time-harmonic wave equation we get

$$\frac{\partial'}{\partial'r'}\left(-r'\frac{\partial p'}{\partial'r'}\right)+r'\frac{\partial'}{\partial'r'}\left(-\frac{\partial p'}{\partial'z'}\right)-\left[\omega'^2-\left(\frac{m'}{r'}\right)^2\right]\frac{r'p'}{\rho}=0. \quad (31)$$

Since now primes will not be written (old variables symbols will be used) but it's important to remember that they are still there.

3.3 The basic of Finite Element Method

Many CFD practitioners prefer finite volume methods because the derivation of the discrete equations is based directly on the underlying physical principles, thus resulting in “physically sound” schemes. From a mathematical point of view, finite volume, difference, and element methods are closely related, and it is difficult to decide that one approach is superior to the others; these spatial discretization methods have different advantages and disadvantages.

Today the Finite Element Method (FEM) has been widely employed in solving field problems arising in modern industrial practices (Zienkiewicz & Taylor, 2000; Hinton & Owen, 1979; Huang et al., 1999; Huebner, 1975). The text in this section is short introduction to the basis of the FEM to the analysis of acoustic problems which are a very common phenomenon in many processes of engineering.

Some physical problems can be stated directly in the frame of variational principle which consists of determining the function which makes a certain integral statement called functional stationary. However the form of the variational principle is not always obvious and such a principle does not exist for many continuum problems.

As an alternative to solve such differential equations we may use a variety of weighted residual methods. Weighted residual methods are numerical techniques which can be used to solve a single or set of partial differential equations. Consider such a set in domain Ω with boundary $\delta\Omega=\Gamma$, where u is the exact solution and may represent a single variable or a column vector of variables.

Applying the method of weighted residuals involves basically two steps. The first step is to assume the general functional behaviour of the dependent field variable in some way so as to approximately satisfy the given differential equation and boundary conditions. Substitution of this approximation into the original differential equation and boundary conditions then results in some error called a residual. This residual is required to vanish in some average sense over the solution domain. The second step is to solve the equation (or equations) resulting from the first step and thereby specialize the general functional form to a particular function, which then becomes the approximate solution sought. According to Galerkin's method, the weighting functions are chosen to be the same as the approximating functions.

Let us consider time independent differential equation

$$L(u)=f \quad (32)$$

defined within a domain Ω and with boundary conditions specified at the boundary of Γ . $L(u)$ typically includes derivatives of u up to second order.

In the weighted-residual method the solution u is approximated by expressions \bar{u} of the form

$$\bar{u} = S_0 + \sum_{j=1}^N u_j S_j \quad (33)$$

where S_j are trial functions, and S_0 must satisfy all the specified boundary conditions ($S_0 = 0$ if all the specified boundary conditions are homogeneous) of the problem, and S_i must satisfy the following conditions:

- S_j should be such that $L(S_j)$ is well defined and nonzero, i.e. sufficiently differentiable;
- S_j must satisfy at least the homogeneous form of the essential boundary conditions of the problem;
- for any N , the set $\{S_j, j=1,2,\dots,N\}$ is linearly independent.

We begin by introducing the error, or residual, R_Ω in the approximation (by substitution of the approximation \bar{u} into the operator equation) which is defined by

$$R_\Omega = L(\bar{u}) - f \quad (34)$$

where \bar{u} contains trial functions and satisfies the Dirichlet boundary conditions of $\bar{u} = u_0$ at $\Gamma_1 \subseteq \Gamma$. If the residual is smaller the approximation is better. It should be noted that R_Ω is a function of position in Ω . Now we attempt to reduce this residual as close to zero as possible. If we have

$$\int_{\Omega} T_i R_\Omega d\Omega = 0 \quad (35)$$

where $T_i, i=1,2,\dots,M$ is a set of arbitrary functions and $M \rightarrow \infty$, then it can be said that the residual R_Ω vanishes. Here T_i are called weighting functions which, in general, are not the same as the approximation (trial) functions S_i . Expanding above equation we have

$$\int_{\Omega} T_i (L(\bar{u}) - f) d\Omega = 0. \quad (36)$$

A function \bar{u} that satisfies above equation for every function T_i in Ω is a weak solution of the differential equation, whereas the strong solution \bar{u} satisfies the differential equation at every point of Ω .

When the operator L is linear above equation can be simplified to the form

$$\sum_{j=1}^N \left(\int_{\Omega} T_i L(S_j) d\Omega \right) u_j = \int_{\Omega} T_i (f - L(S_0)) d\Omega \quad (37)$$

or

$$\sum_{j=1}^N A_{ij} u_j = f_i \quad (38)$$

where

$$A_{ij} = \int_{\Omega} T_i L(S_j) d\Omega \quad (39)$$

and

$$f_i = \int_{\Omega} T_i (f - L(S_0)) d\Omega. \quad (40)$$

Note that the coefficients of matrix \mathbf{A} is not symmetric $A_{ij} \neq A_{ji}$.

The weighted-residual method (when $T_i \neq S_i$) is also sometimes referred to as the Petrov-Galerkin method. For different choices of T_i the method is known by different names. We outline below the most frequently used methods.

The Galerkin method. For $T_i = S_i$ the weighted-residual method is known as the Galerkin method. When the operator is a linear differential operator of even order, the Galerkin method reduces to the Ritz method. In this case the resulting matrix will be symmetric because half of the differentiation can be transformed to the weight functions.

The least-squares method. The least-squares methods determines the constants u_j by minimizing the integral of the square of the residual

$$\frac{\partial}{\partial u_i} \int_{\Omega} R_{\Omega}^2 d\Omega = 0 \quad (41)$$

or

$$\int_{\Omega} \frac{\partial R_{\Omega}}{\partial u_i} R_{\Omega} d\Omega = 0 \quad (42)$$

A comparison of Eq. (42) with Eq. (35) shows that $T_i = \frac{\partial R_{\Omega}}{\partial u_i}$. If L is a linear operator Eq. (37)

becomes

$$\sum_{j=1}^N \left(\int_{\Omega} L(S_i) L(S_j) d\Omega \right) u_j = \int_{\Omega} L(S_i) (f - L(S_0)) d\Omega \quad (43)$$

which yields a symmetric matrix but requires the same order of differentiation as the operator equation.

The collocation method. The collocation method seeks approximate solution \bar{u} by requiring the residual $R_\Omega = R_\Omega(\mathbf{x}, \mathbf{u})$ in the equation to be identically to zero at N selected points \mathbf{x}_i , $i = 1, 2, \dots, N$ in the domain Ω

$$R_\Omega(\mathbf{x}_i, u_j) = 0. \quad (44)$$

The selection of the points \mathbf{x}_i is crucial in obtaining a well conditioned system of equations and ultimately in obtaining an accurate solution. The collocation points can be shown to be a special case of Eq. (34) $\int_\Omega T_i R_\Omega d\Omega = 0$ for $T_i = \delta(\mathbf{x} - \mathbf{x}_i)$, where $\delta(\mathbf{x})$ is the Dirac delta function

$$\int_\Omega f(\mathbf{x}) \delta(\mathbf{x} - \xi) d\Omega = f(\xi). \quad (45)$$

The Courant method. To so-called Courant method combines the basic concepts of the Ritz method and the least-squares method (for linear operator). The method seeks approximate solution \bar{u} by minimizing the modified quadratic functional

$$I_p(\bar{u}) = I(\bar{u}) + \frac{\alpha}{2} \|L(\bar{u}) - f\|^2 \quad (46)$$

where $I(u)$ is the quadratic functional associated with $L(u) = f$, when L is linear, and α is the penalty parameter. Obviously the statement make sense only for operator equation that admit functional formulation.

4. Numerical results

In this section the time-harmonic wave analysis of idealized reflective pipe with rigid walls and outlet to free acoustic space is made. The frequency dependent sound transmission loss of hard-walled pipe is calculated as well. The pipe length is L and the radius is r . Free acoustic space beyond pipe outlet (no reflected wave is coming back to pipe outlet) is modelled by larger pipe with impedance boundary condition.

One can use Perfectly Matched Layers (PMLs) to model space outside the outlet pipe (Comsol, 2007). The PMLs serve to absorb the outgoing waves so that the non-physical reflections at their exterior boundaries have a minimal influence on the pressure field inside the pipe.

Fig.2A shows example of non-dimensional pipe with length $L = 1$, radius $r = 0.2$ and free acoustic space modelled by cylinder with $L = 2$ and radius $r = 2$. The axial symmetry of the geometry and the physics makes it natural to set the model up in a 2D axisymmetric application mode (see Fig.2B).

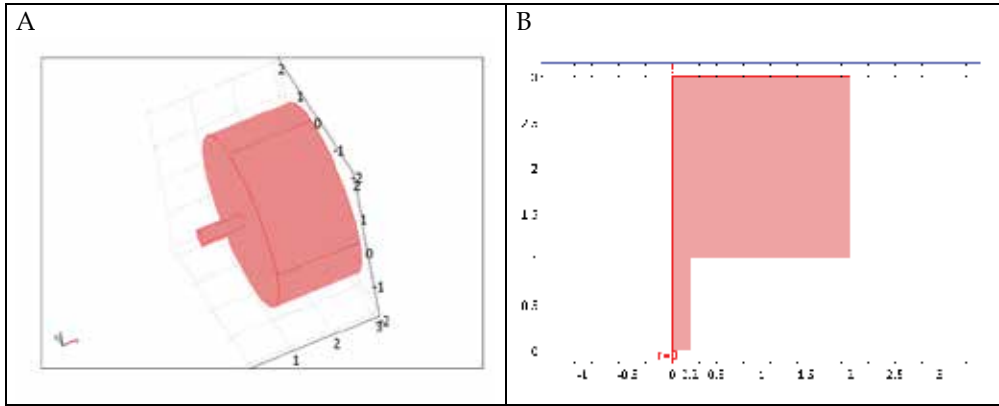


Fig. 2. Geometry: (A) three-dimensional domain and (B) 2D axisymmetric mode

All acoustic problems considered in this work are governed by dimensionless equations with appropriate boundary and initial conditions. Numerical results are obtained using standard computational code COMSOL Multiphysics (Comsol, 2007). Finite element calculations are made with second-order triangular Lagrange elements. First acoustic pressure is calculated using dimensionless time-harmonic wave equation (31) in considered axisymmetric domain.

Following boundary conditions for dimensionless model are assumed:

- pipe inlet: spherical wave is assumed;
- pipe wall: sound hard boundary $\mathbf{n} \cdot (-\nabla p) = 0$;
- cylinder's wall: impedance boundary condition $\mathbf{n} \cdot (-\nabla p) = \frac{i\omega p}{Z}$ where Z is impedance and $Z = 1$ in dimensionless form;
- axial symmetry $r = 0$.

Next using

$$W_i = \int_A p_0^2 dA, \quad W_t = \int_A |p|^2 dA \quad (47)$$

the incoming and transmitted dimensionless sound powers are calculated, respectively. Using calculated sound powers transmission loss (TL) is derived using Eq. (10).

To receive dimension frequency of time-harmonic wave, one can multiply dimensionless frequency by $\frac{c}{L}$ using formula $f = \frac{c}{L} f'$, where f' is dimensionless frequency, L is length

of pipe and c is sound velocity in medium. All figures in this chapter with frequency dependent sound transmission loss diagram present TL on vertical axis and dimensionless frequency on horizontal axis. Firstly, the influence of boundary condition at inlet on sound transmission loss is shown at Fig. 3. Two boundary conditions are assumed for pipe with ratio $r/L = 0.1$: plane wave with the applied pressure source amplitude p_0 and pressure

$p = p_0$ with $p_0 = 1$. One can observe a big difference between frequency dependent sound transmission loss values calculated with both conditions.

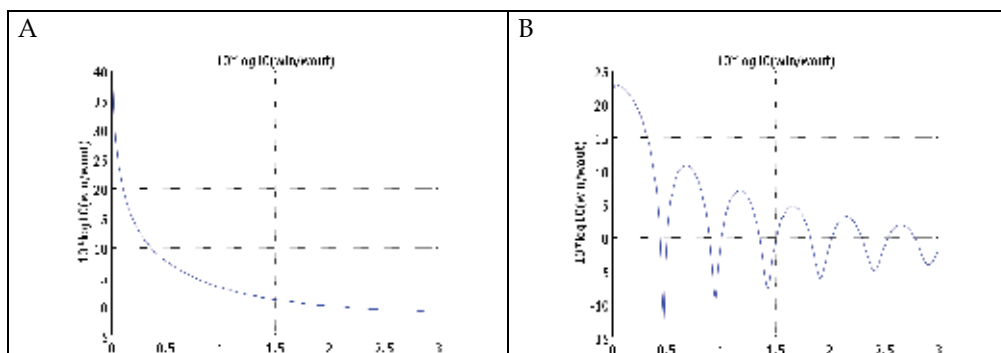


Fig. 3. Influence of inlet boundary condition on TL of pipe with ratio $r/L = 0.1$ versus dimensionless frequency: (A) plane wave, (B) pressure

Secondly, the influence of mesh density in considered domain on sound transmission loss is shown at Fig. 4, as well. The frequency dependent sound transmission loss is calculated for pipe with ratio radius to length equals 0.2 with mesh statistic collected in Table 1. However, FEM has some deficiencies at higher frequencies because of the shorter wavelengths and the higher modal density (see Fig.5). Due to the shorter wavelengths, the number of elements must be increased with frequency.

Number of:	Mesh A	Mesh B
degrees of freedom	951	14601
mesh points	251	3701
elements	450	7200

Table 1. Mesh statistic for pipe with ratio $r/L = 0.2$

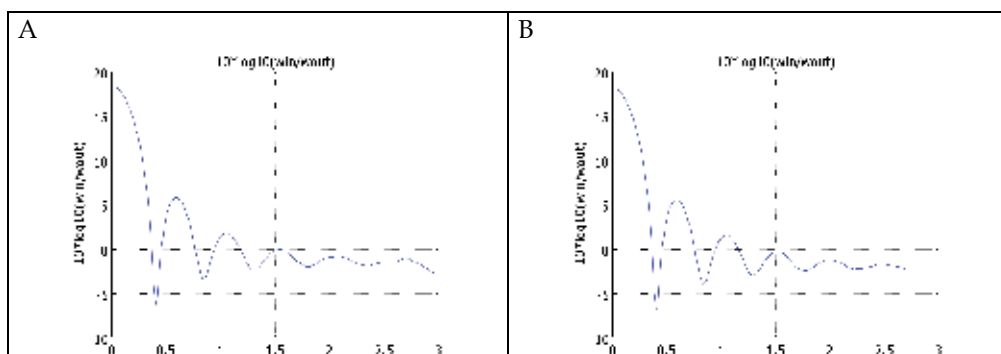


Fig. 4. Influence of mesh density on TL of pipe (ratio $r/L = 0.2$): (A) mesh A, (B) mesh B

Next computational results at Fig. 6. show very important influence of outlet boundary condition on the frequency dependent sound transmission loss. In proposed method free acoustic space outside pipe outlet is modelled by larger cylinder with impedance boundary

condition on all outer boundaries. The radius of this cylinder is at least ten times greater than pipe radius. When cylinder's radius is smaller calculated numerical results are not correct (see Fig.6). In many papers transmitted sound power is calculated on the end of device where impedance boundary condition or plane wave with zero pressure source amplitude is assumed. This assumption does not guarantee proper model of environment outside of pipe outlet.

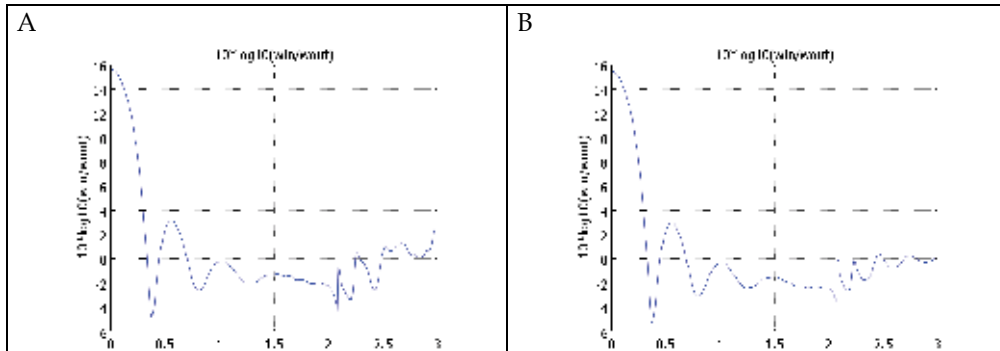
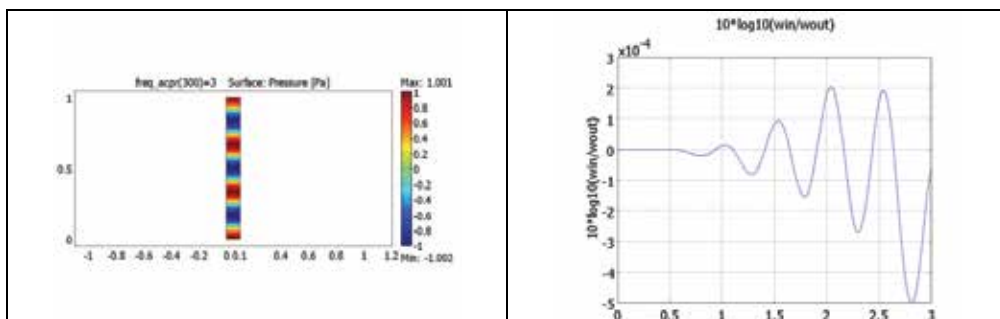


Fig. 5. Influence of mesh density on TL of pipe (ratio $r/L = 0.3$): (A) coarse, (B) fine mesh

Figures 7-9 show comparison of numerical results of calculated frequency dependent sound transmission loss, transmitted sound power and pressure at frequency $f = 3$ for pipes with different radius to length ratio equals 0.015625, 0.0625, 0.1 and 0.2. It is easy to observe that pipe with smaller ratio has greater possibilities to attenuate sound. Absolute value of transmission loss is greater.

Fig. 10 presents numerical results for pipe with ratio $r/L = 0.15$. Left figure presents frequency dependent pressure in point at inlet and outlet of pipe. Both points are placed on axial symmetry line. Right figure presents frequency dependent sound transmission loss for this pipe.

Fig.11 presents comparison of measured transmission loss of 37.5 mm diameter pipe of different lengths with numerical results calculated using 2D axisymmetric non-dimensional model with proper r/L ratio. Measured transmission loss of pipe was found from the Fourier Transforms of the acoustic pressure time histories using Eq. (12). The comparison diagram shows that results of numerical calculations (solid line) are quite good with comparison to experimental data (dashed line) at frequencies smaller than 3.



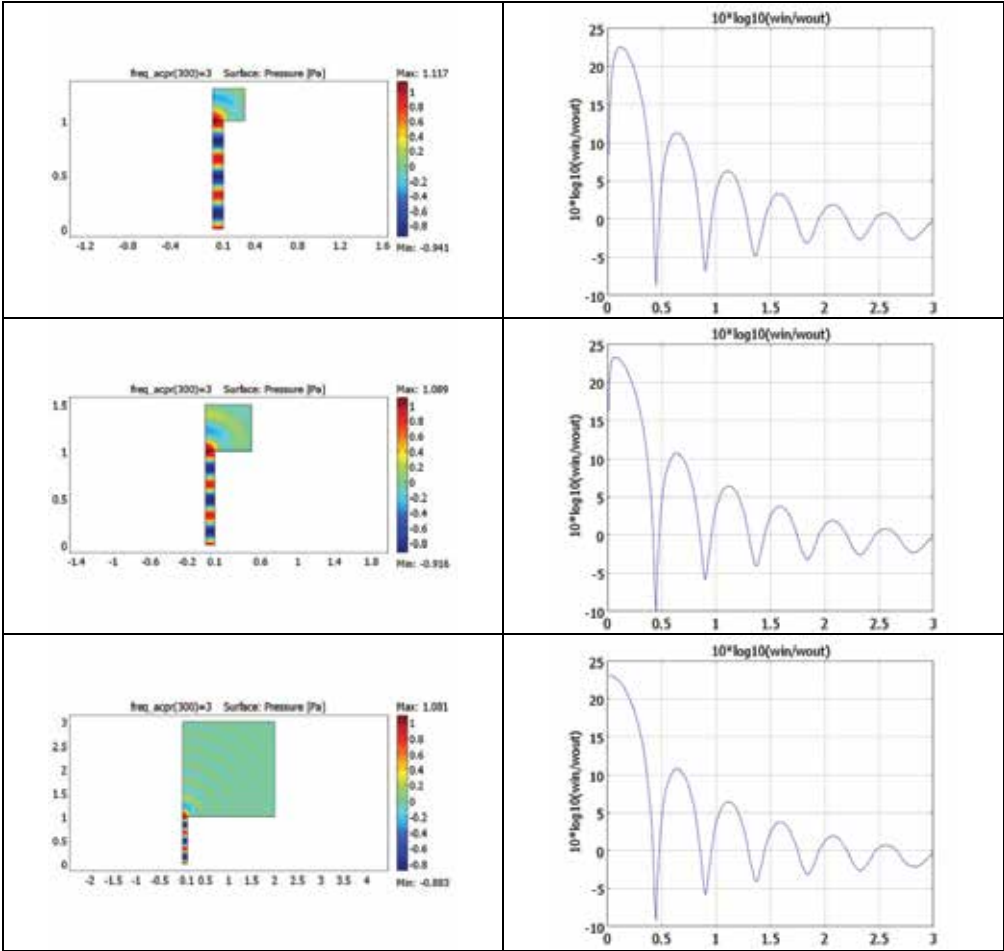
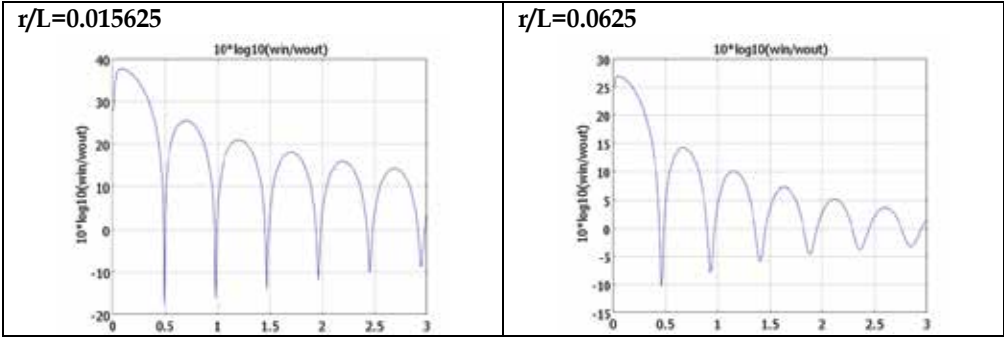


Fig. 6. Results of numerical calculation of transmission loss of pipe (length $L = 1$, radius $r = 0.1$) with free acoustic space modelled by different size cylinder.



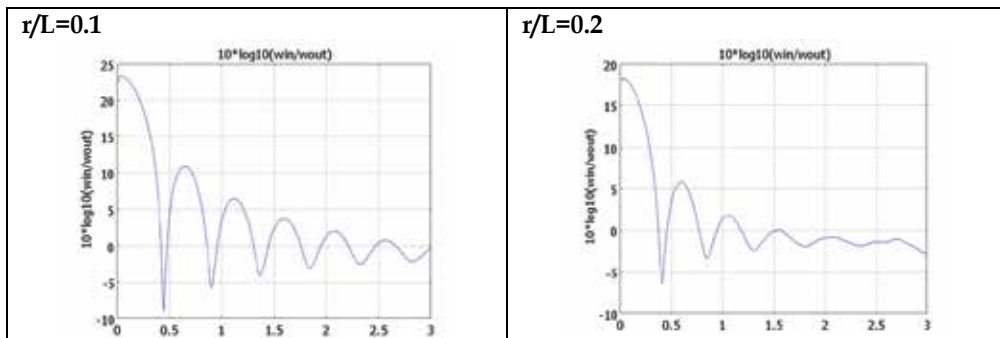


Fig. 7. Comparison of results of numerical calculation of transmission loss with different r/L ratios

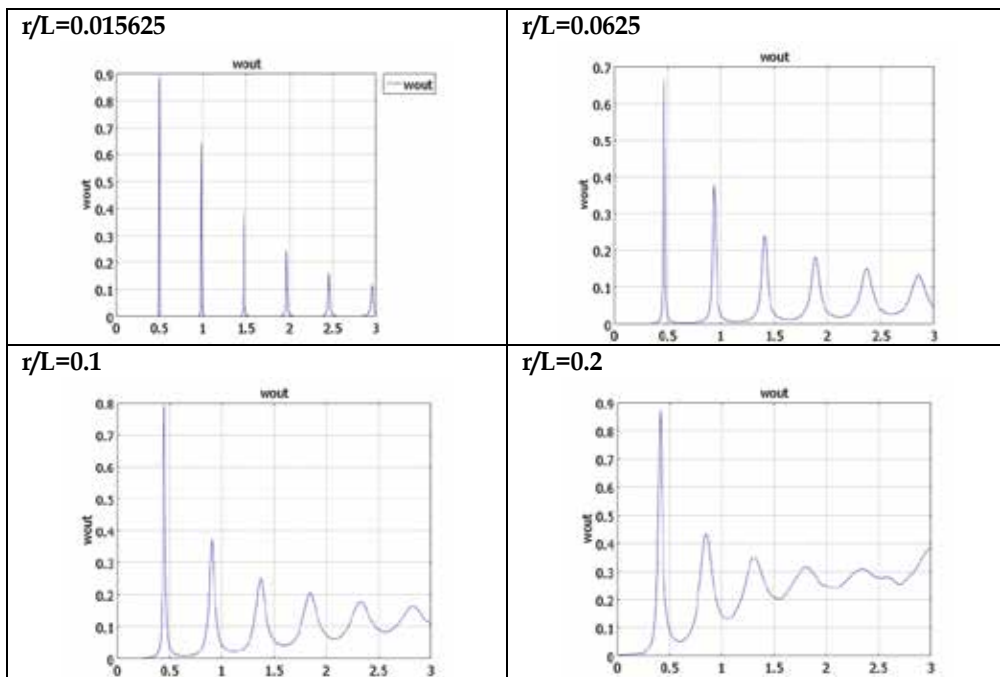


Fig. 8. Comparison of results of numerical calculation of transmitted sound power for pipes with different r/L ratios

Calculated numerical results of the frequency dependent sound transmission loss of reflective pipe for various ratios (radius to length) are compared with measured sound transmission loss for pipes with radius and length collected in Table 2.

Maximum dimension frequency presented on horizontal diagram axis is calculated using dimensionless frequency and formula $f = \frac{c}{L} f'$, where $f' = 3$ is dimensionless frequency, L is length of pipe and $c = 343 \text{ m/s}$ is sound velocity in medium.

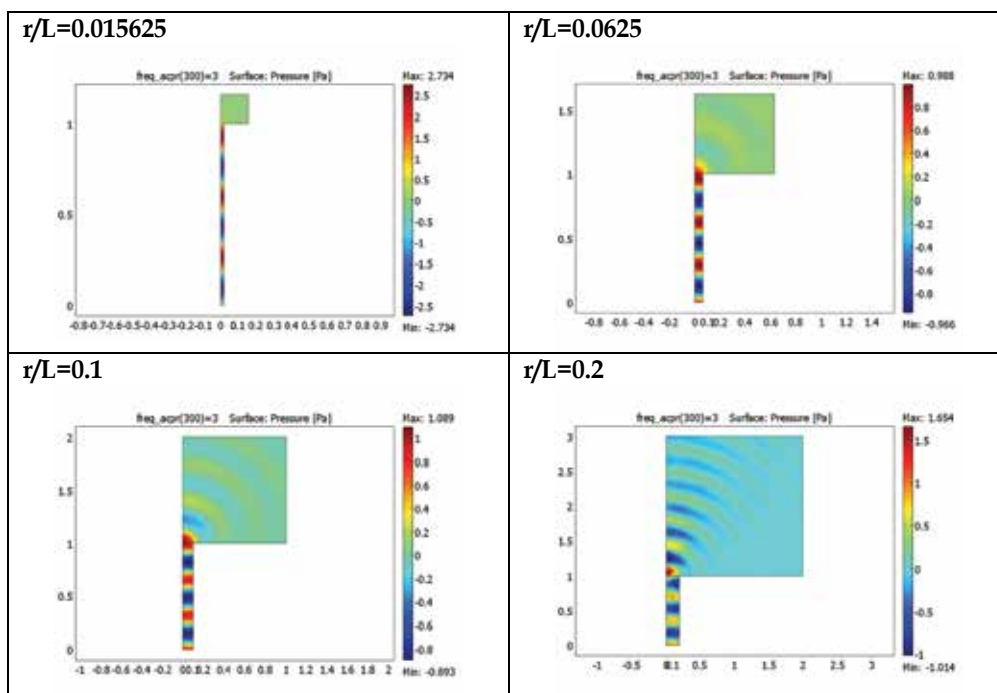


Fig. 9. Comparison of results of numerical calculation of pressure at dimensionless frequency $f = 3$ for pipes with different r/L ratios

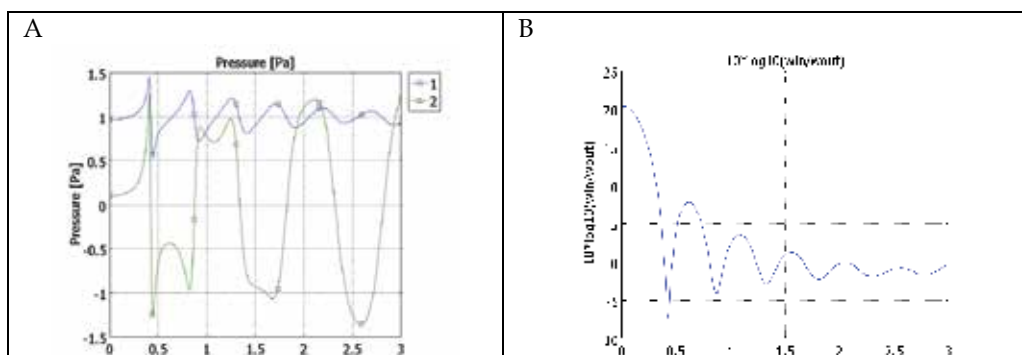


Fig. 10. Numerical results for pipe with $r/L = 0.15$: (A) frequency dependent pressure in point at pipe inlet (line - 1) and outlet (line - 2), (B) transmission loss

There exists for each pipe critical frequency after which calculated and measured frequency dependent sound transmission loss of reflective pipe begin change with different way (see Fig. 12). For example, for pipe with ratio $r/L = 0.015625$ this critical frequency is about 17 and 2 for pipe with ratio $r/L = 0.3$. If frequencies grow then wavelengths diminish and become smaller than size of finite element. Moreover, for high frequencies wave inside hard-walled pipe occurs superposition of longitudinal and transverse waves. These conclusions suggest that computational method is valid for only a limited range of frequencies or wavelengths.

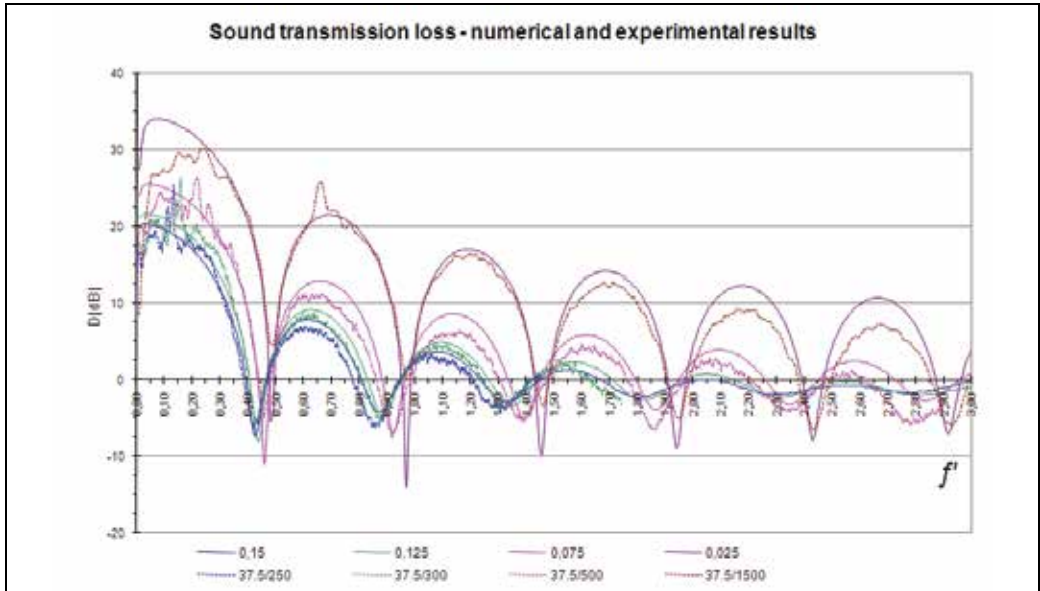


Fig. 11. Comparison of results of numerical calculation of transmission loss with measured transmission loss

Ratio, r/L	Radius, $r [mm]$	Length, $L [mm]$	Max frequency, $f_{max} [1/s]$
0.025	37.5	1500	686
0.075	37.5	500	2058
0.125	37.5	300	3430
0.15	37.5	250	4116

Table 2. Dimensionless ratio and dimension quantities for pipes

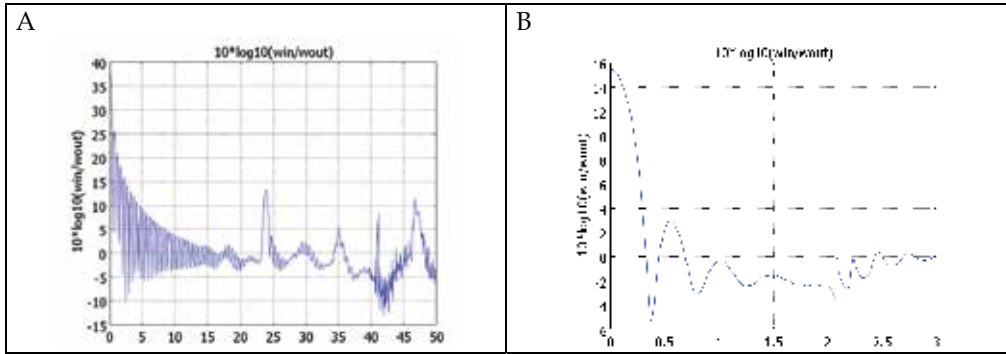


Fig. 12. Frequency dependent sound transmission loss of reflective pipe with ratio: (A) $r/L = 0.015625$, (B) $r/L = 0.3$

5. Conclusions

A lot of tools are available for the study of the time-harmonic wave analysis of acoustic devices. The classical Finite Element Method (FEM) and Boundary Element Method (BEM) are the most conventional tools and are widely used in the low frequency range.

However theoretically correct, these methods have some deficiencies at higher frequencies because of the shorter wavelengths and the higher modal density. In FEM, because of the shorter wavelengths, the number of elements must be increased with frequency. This makes the method costly at high frequencies.

In this chapter numerical method for calculation of the frequency dependent sound transmission loss within reflective pipe with outlet to free acoustic space is presented. The frequency dependent sound transmission loss of hard-walled pipe with given ratio radius to length is calculated as well.

All acoustic problems considered in this work are governed by dimensionless equations with appropriate boundary and initial conditions. Numerical results are obtained using standard computational code COMSOL Multiphysics using FEM.

The results of this study (for pipes with ratio r/L less than 0.3) show that results of numerical calculations are quite good with comparison to experimental data at dimensionless frequencies less than 3, only if size of cylinder at the pipe outlet is big enough. The FEM results are verified by a calculation with a very fine mesh. The FEM results calculated with a more coarse mesh, show a good agreement with the experiment case.

6. References

- Byrne K., Skeen M., Kessissoglou N. (2006). Measurement of the sound transmission of a small expansion chamber muffler to consider the effects of mean flow and wall compliance, *Proceedings of ACOUSTICS*, 20-22 November 2006, Christchurch, New Zealand, pp. 257-263.
- Chen K.T., Chen Y.H., Lin K.Y., Weng C.C. (1998). The Improvement on the Transmission Loss of a Duct by Adding Helmholtz Resonators, *Applied Acoustic*, Vol. 54, No. 1, pp.71-82.
- Comsol Multiphysics User's Guide (2007). *Modeling Guide and Model Library*, Documentation Set, Comsol AB.
- Comsol Multiphysics User's Guide (2007). *Acoustic Module Model*, Documentation Set, Comsol AB.
- Cummings A. (1974). Sound transmission in curved duct bends, *Journal of Sound and Vibration* 35, pp.451-477.
- Cummings A., Kirby R. (1999). Low-Frequency Sound Transmission In Ducts With Permeable Walls, *Journal of Sound and Vibration*, 226(2), pp. 237-251.
- Hinton E., Owen D.R.J. (1979). *An Introduction to Finite Element Computations*, Pineridge, Swansea.
- Hou-Cheng Huang, Zheng-Hua Li and Asif S. Usmani (1999). *Finite Element Analysis of Non-Newtonian Flow*, Springer-Verlang, London.
- Huebner K.H. (1975). *The Finite Element Method for Engineers*, Wiley, Toronto.

- Hyeon-Don Ju, Shi-Bok Lee (2005). Transmission Loss Estimation of Three Dimensional Silencers with Perforated Internal Structures Using Multi-domain BEM, *Journal of Mechanical Science and Technology*, Vol. 19, No. 8, pp. 1568-1575.
- Ko S.H., Ho L. T. (1977). Sound attenuation in acoustically lined curved ducts in the absence of fluid flow, *Journal of Sound and Vibration* 53, pp.189-201.
- Ko S.H. (1979). Three-dimensional acoustic waves propagating in acoustically lined cylindrically curved ducts without fluid flow, *Journal of Sound and Vibration* 66, pp.165-179.
- Rossing Thomas D. (Ed.) (2007). *Handbook of Acoustic*, Springer, LLC New York, 2007.
- Rostafinski W. (1974), Analysis of propagation of waves of acoustic frequencies in curves ducts, *Journal of the Acoustical Society of America* 56, pp.11-15.
- Sarigug A.S. (1999). Sound Attenuation Characteristics of Right-Angle Pipe Bends, *Journal of Sound and Vibration*, 228(4), pp.837-844.
- Thompson L.L. (2006). A review of finite-element methods for time-harmonic acoustics, *Journal of Acoustic Society of America*, 119 (3), pp. 1315-1330.
- Utsumi M. (2001). Sound Transmission in Circular Ducts of Continuously Varying Cross-Sectional Area, *Journal of Sound and Vibration*, 242(2), pp.369-376.
- Wilson G.P., Soroka W.W. (1965). Approximation to the Diffraction of Sound by a Circular Aperture in a Rigid Wall of Finite Thickness, *The Journal of The Acoustical Society of America*, Volume 37, Number 2, pp.286-297.
- Zienkiewicz O.C., Taylor R.L. (2000). *The Finite Element Method*, Vol. 1-3: The Basis, Solid Mechanics, Fluid Dynamics (5th ed.), Butterworth-Heinemann, Oxford.



Edited by David Moratal

Finite element analysis is an engineering method for the numerical analysis of complex structures. This book provides a bird's eye view on this very broad matter through 27 original and innovative research studies exhibiting various investigation directions.

Through its chapters the reader will have access to works related to Biomedical Engineering, Materials Engineering, Process Analysis and Civil Engineering. The text is addressed not only to researchers, but also to professional engineers, engineering lecturers and students seeking to gain a better understanding of where Finite Element Analysis stands today.

Photo by archerix / iStock

IntechOpen

