Journal of
Mechanics of
Materials and Structures

JOURNAL OF MECHANICS OF MATERIALS AND STRUCTURES

jomms.org

Founded by Charles R. Steele and Marie-Louise Steele

Cover design: Alex Scorpan   Cover photo: Ev Shafrir

∎
∎∎msp

DEDICATION

This issue of the *Journal of Mechanics of Materials and Structures* is dedicated to Professor Charles R. Steele, the journal's Founding Editor, and his late wife, Marie-Louise Steele, Associate Editor of *JoMMS* until her untimely death in 2009. It includes papers written by colleagues and friends who thereby wanted to honor the two. We, who have succeeded Charles and Marie-Louise at the helm of the journal, consider it a privilege to have acted as editors of this issue.

The issue has a twofold purpose: to celebrate the Steeles' accomplishments and to commemorate the passing of Marie-Louise, who was an indispensable partner in running the journal and is to be credited to no small degree with supporting Charles in his professional endeavors. They both made an important contribution to the engineering community by founding the nonprofit *Journal of Mechanics of Materials and Structures*, and prior to that by their editorship of the *International Journal of Solid and Structures*. Charles's own research work has had a lasting impact on theoretical and applied mechanics and biomechanics, addressing problems in shell theory, elasticity theory, cochlear mechanics, and bone mechanics, among others. He has been recognized for these contributions by being elected to National Academy of Engineering in 1995 and by the ASME Koiter medal in 1999.

The original papers presented in this issue are authored by researchers whose scientific interests fall within the broad spectrum of *JoMMS*. They include theoretical, numerical and experimental reports spanning the fields of mechanics, mathematics and biology, covering fundamental problems in continuum and discrete mechanics, deformation, flow and failure of materials (beams, shells, composites, fractal media), wave propagation of linear and nonlinear waves, cellular mechanics, adaptation of biological tissues, and a range of other topics.

We extend sincere thanks to all the authors who contributed to this issue, for which the manuscripts were reviewed according to the rules of *JoMMS*. We thank them for their cooperation throughout the process.

DAVIDE BIGONI: bigoni@ing.unitn.it
Dipartimento di Ingegneria Meccanica e Strutturale, Facoltà di Ingegneria - University of Trento, Via Mesiano 77, I-38123 Povo (TN), Italy

IWONA JASIUK: ijasiuk@illinois.edu
Department of Mechanical Science and Engineering, University of Illinois at Urbana-Champaign, 1206 West Green Street, Urbana, IL 61801, United States

YASUHIDE SHINDO: shindo@material.tohoku.ac.jp
Department of Materials Processing, Graduate School of Engineering, Tohoku University, Aoba-yama 6-6-02, Aoba-ku, Sendai 980-8579, Japan

FROM THE EDITORAL BOARD

The Editorial Board of *JoMMS* and the Publisher join the Chief Editors today in remembering Marie-Louise Steele with fondness, expressing their warmest appreciation and good wishes to Charles, and renewing their commitment to *JoMMS*, so that it may evermore be a tribute to the monumental work of its founders.

▪msp

MARIE-LOUISE (BÜHLER) STEELE (1943–2009)

On February 17, 2009, Marie-Louise (MarieLu) was diagnosed with metastatic breast cancer. This was a complete shock, since she had a clear mammogram just 18 months before, there is little cancer in the family, her father lived to 92, and her mother lived to 99. Her mother always told MarieLu that she would not live as long, because she used too much of her energy in each day. As many know, MarieLu did not spare effort in her projects and in helping friends and family. Below is a photograph taken at the beginning of radiotherapy; the one on the next page was taken just before the first chemotherapy session on April 15. Despite her good physical condition and very positive outlook, she reacted very poorly to the chemotherapy, entered the hospital intensive care on April 30, and died on May 14. Although she grew weaker by the day, she always had a smile to greet visitors.

When MarieLu received the health report on February 17, she said there should be no tears and quoted Edith Piaf: *Je ne regrette rien.* She and Charles had such good fortune in meeting each other and sharing life in journal work, morning workout followed by coffee latte in the hot tub, so much opportunity to travel and keep friends around the world, and healthy children and grandchildren.

Personal background

MarieLu was born in Heilbronn (Germany) in 1943 to Karl and Elise Bühler, and remembered well growing up in the aftermath of World War II, an experience that helped to form her lifelong dedication to conservation and hard work. Her father was a prisoner of war in Siberia, and was one of the few to return. After recuperation he became Burgermeister of Brackenheim, a small town about 30 miles from Stuttgart. She always remembered as a nine-year-old, looking up at the stars and vowing that she would

MarieLu with grandchildren (and granddog) on March 14, 2009.

MarieLu on a hike with the family at the Golden Gate Bridge, San Francisco on April 5, 2009.

see the world one day. She loved poring over maps and picked two places she had to visit. One was San Francisco and the other was Rio de Janeiro. As it transpired, she has lived at Stanford near San Francisco, and has spent four New Year's Eves on Copacabana, related to attending PACAM meetings with Charles.

The acceptable means for a young lady to escape to the large world was through language. So, after finishing the arbitur, MarieLu attended interpreter school in Stuttgart and then in Geneva. The international flavor was greatly to her liking, and she became fluent in French. However, her focus was English, and she decided that she must spend time in an English-speaking country to achieve proficiency. So she immigrated to the United States, which was relatively easy then for northern Europeans. In Geneva she worked part-time in an office of Hewlett-Packard. Apparently her skills were already developed at the age of 21, since a position was offered to her if she would come to Palo Alto, California. She first spent three months working for Siemens in the Empire State Building. The chaos of Manhattan, however, did not appeal to her, but she decided to give the US one more chance. In Geneva she had purchased a bus ticket: 100 days for 100 dollars. With her direct practical thinking, she took the bus nonstop from Manhattan to Palo Alto. After a few years in the international office of Hewlett-Packard, she decided to go to school full time, but then walked into the life of Charles Steele, a single father of three boys — Eric, Brett and Jay — and changed both of their lives forever. Thus began a 40-year love affair, enhanced by the arrival in 1971 of their son Ryan.

When Ryan was only one month old, the family packed up and traveled to Switzerland, where Charles was invited to teach for a year. MarieLu was not daunted by caring for a newborn and setting up a new household in a foreign country. She thrived on challenges and later completed the BS and MS degrees, made possible by organizing an equal distribution of household chores and cooking among all family members. Over the years, the family went on to spend sabbatical time in Taiwan, Sweden, Germany and South Africa.

MarieLu will be remembered by all for her bubbly personality, her passion for life and culture. After the boys were grown up and on their own, MarieLu and Charles continued to travel, and spent months at

a time in Italy, Spain, Germany, Thailand, and India. In recent years, they traveled with three computers, in order to combine work and pleasure. Around the globe, MarieLu won friends with her vivaciousness and zest for life.

MarieLu further will be remembered for her generous assistance to others less fortunate. She provided a home away from home for generations of foreign visitors and relatives, both distant and close. She gave to all she met, thrilled to be able to help in anyway she could. In 1993, she helped a Ukranian friend start a dress shop in the Ukraine with second-hand wedding dresses collected in the United States. She sponsored two of her German nieces to come to school in the States and put two step-grandchildren through University. She helped each of her sons secure a home for their families, and often handled the home improvements herself. The examples of generosity and selflessness go on and on. She was tireless, and even while the cancer attacked her body, she worked on the journal, painted, cleaned, worked out at the YMCA, and played with her grandchildren.

MarieLu loved all creatures of the world, and would even rescue spiders from the shower to liberate in the garden. She cherished her furry friends and the little birdies she fattened up with birdseed in the deluxe feeders and birdhouses she constructed in her garden. She often cared for her grandkitties and granddogs, spoiling them as only a grandmother can. Some recent fond memories are of her butterfly walks this past Easter with her young granddaughters at her beloved Sea Ranch in Sonoma County, California, where the family maintained a vacation home.

Journal work

In 1965, George Herrmann founded the *International Journal of Solids and Structures* (*IJSS*), one of a number of journals launched by Pergamon Press around that time. In 1984, he retired, and Charles Steele agreed to succeed him as editor. Because of a previous change in staff, the editorial office was not in good order, with piles of manuscripts and cabinets full of unanswered correspondence, some more than a year old. MarieLu offered to help for a few days. However, she quickly proved to be indispensible and was appointed as Associate Editor. The few days turned into 20 years. She was well suited for this, since her great skill was bringing order to a chaotic situation, with her desire to help people and her interest in the international community. She grew to know thousands of authors and reviewers around the world. Correspondents were appreciative of her cheerfulness, efficiency and warmth — rarities when dealing with a technical journal! As the figure on the next page shows, *IJSS* was stable in the first 20 years of existence, with around 150 submissions per year. After MarieLu began work, the number of submissions increased exponentially to over 900 in 2004. This is despite the launching of a number of competing journals of mechanics during this time. Standards were not relaxed, with only about 60% of the submitted papers published all that time.

Unfortunately, in the hands of Elsevier, the price per page also increased substantially. The Cornell University Library includes *IJSS* among the four most outrageous examples of pricing for engineering journals:

http://engineering.library.cornell.edu/about/stickershock_4

Generally, the price per article (or per page) of technical journals from commercial publishers are many times those published by nonprofit organizations. Consequently in response to the international library crisis, Charles and MarieLu, with 21 of the 23 members of the Board of Editors, resigned from
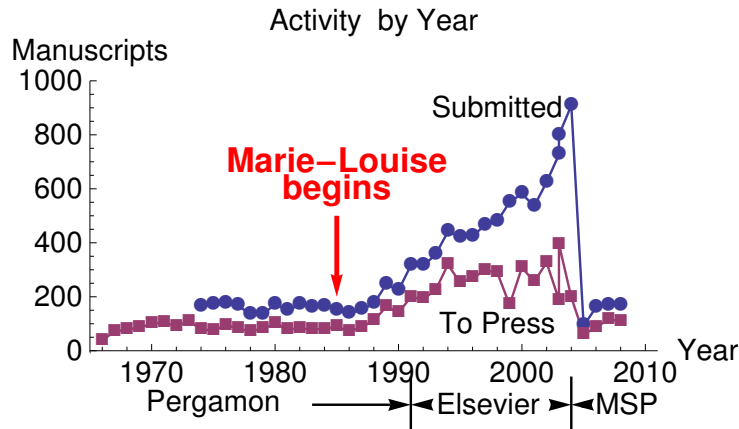
Manuscripts submitted for publication and those accepted and sent to the publisher for publication, by year. Included are *IJSS*, from 1965 through 2004 under Pergamon and Elsevier, and *JoMMS* from 2005 through 2008 under MSP. After Marie-Louise Steele began with *IJSS* in 1985, the increase in activity is substantial. A new phase starts with the low-cost *JoMMS*.

IJSS to establish the *Journal of Mechanics of Materials and Structures* (*JoMMS*). The publisher is the nonprofit Mathematical Sciences Publications (MSP). *JoMMS* has made a very good beginning, now with about 160 submissions per year. MarieLu's ambition was to see *JoMMS* catch up to *IJSS* in activity, but with the cost to libraries remaining low. The advantages of *JoMMS* for authors and libraries make this a possibility. In any case, the mechanics community will miss the contribution of Marie-Louise Steele. In the hospital, she was pleased and quite relieved to hear that Davide Bigoni, Iwona Jasiuk, and Yasuhide Shindo agreed to take over the responsibility for *JoMMS*, and joint Chief Editors. In their hands, her ambition for *JoMMS* will be fulfilled.

ELIZABETH WILLES, friend, and CHARLES STEELE, husband
6 October 2009

msp

FINITE STRAIN MICROMECHANICAL MODELING OF THERMOVISCOELASTIC MATRIX COMPOSITES

Jacob Aboudi

A finite strain micromechanical analysis is generalized for the modeling of thermoviscoelastic matrix composites. The thermoviscoelastic matrix of the composite is represented by a finite thermoviscoelasticity theory that permits (in contrast to finite linear thermoviscoelasticity theories) large deviations away from thermodynamic equilibrium. As a result, it is possible to subject the composite to large thermomechanical loadings. In addition, the possibility of evolving damage in the matrix is included. The derived micromechanical model is applied to investigate the behavior of a thermoviscoelastic rubber-like matrix reinforced by steel fibers in various circumstances. By subjecting the composite to mechanical loading under isentropic conditions, the micromechanical model is employed for the prediction of thermoelastic inversion point at which the Gough–Joule phenomenon at the rubber-like phase occurs. Results are given that show the effect of damage, elevated temperature and viscoelasticity of the matrix on the global response of the composite including its creep and relaxation behavior.

1. Introduction

The viscoelastic effects of polymers that are undergoing large deformations can be modeled by finite linear viscoelasticity. In the framework of this theory the strains are finite, but the deviations away from thermodynamic equilibrium are assumed to be small. This implies that the equations that control the evolution of the internal variables are linear; see [Lockett 1972; Christensen 1982; Holzapfel 2000; Simo 1987], for example. For large deviations from equilibrium the finite linear viscoelasticity is not applicable any more and nonlinear evolution laws must be introduced to allow more accurate modeling. To this end a finite viscoelasticity theory was formulated in [Reese and Govindjee 1998b] in which the equations of evolution are nonlinear thus allowing very large strains to take place. In [Govindjee and Reese 1997] comparisons have been made between the finite linear viscoelasticity of [Simo 1987] and their developed finite viscoelasticity theory.

The development of thermoviscoelasticity theories at finite strain that include viscous and nonisothermal effects is very important since the behavior of polymeric materials is strongly influenced by temperature changes. A thermoviscoelasticity theory at finite strains was presented in [Holzapfel and Simo 1996]. In this formulation, however, the evolution equations of the internal variables are linear and, therefore, this theory can be considered as finite linear thermoviscoelasticity. Finite thermoviscoelasticity theories that allow finite perturbations away from thermodynamic equilibrium were presented by [Lion 1996; Reese

and Govindjee 1998a]. The formulation in the latter reference is based on entropic elasticity and involves nonlinear evolution equations that allow the modeling of significant thermomechanical deformations of the material and permits large deformation rates.

The derived constitutive equations of [Reese and Govindjee 1998a] are based on the multiplicative decomposition of the deformation gradient into elastic and viscous parts. In addition, the free energy is decomposed into equilibrium (which corresponds to time-independent thermoelastic deformation) and nonequilibrium (which corresponds to the time-dependent deformation) parts. Damage considerations are not included in the finite thermoviscoelasticity of that paper. In the present investigation, evolving damage in finite thermoviscoelastic materials is included by adopting, in the framework of continuum damage mechanics, the derivation of [Lin and Schomburg 2003; Miehe and Keck 2000], according to which the rate of damage depends upon the kinematic arc-length. Isothermal finite viscoelasticity with evolving damage is obtained as a special case, and by neglecting the thermal and viscous effects, the special case of a hyperelastic material with evolving damage is obtained.

In [Aboudi 2004], a micromechanical analysis was proposed for the prediction of the behavior of composites undergoing large deformations in which one of the phases is modeled by the finite linear thermoviscoelasticity theory of [Holzapfel and Simo 1996]. In the present investigation, this finite strain micromechanical analysis, referred to as *high-fidelity generalized method of cells* (HFGMC), is extended to incorporate polymeric phases that can be modeled by the finite thermoviscoelasticity of [Reese and Govindjee 1998a] in which damage can evolve, the rate of which depends upon the kinematic arc-length. As a result of the present generalization, finite strain constitutive equations that govern the macroscopic behavior of the anisotropic thermoviscoelasticity composites undergoing large deformations with evolving damage in the polymeric phase are established. These equations involve the damaged instantaneous mechanical and thermal tangent tensors as well as a global tensor that includes the current viscoelasticity and damage effects. Every one of these three tensors is given by closed-form expressions that involve the instantaneous properties of the phases and the corresponding current mechanical, thermal and viscoelastic-damage tensors which have been established from the micromechanical procedure. The special case of composites that consist of finite viscoelastic phases has been recently investigated in [Aboudi 2010].

Results are given for a thermoviscoelastic rubber-like matrix reinforced by thermoelastic steel fibers. The thermoviscoelastic rubber-like matrix consists of a system of a thermoelastic element together with a single Maxwell element which is represented by the equilibrium and the nonequilibrium parts of the free energy. The thermoelastic inversion point which is characteristic for rubber-like materials at which the Gough–Joule effect occurs is determined by analyzing an isentropic process that provides the induced temperature by stretching the composite in the transverse direction (perpendicular to the fibers). In addition, the effects of elevated temperature, viscoelasticity and damage on the steel/rubber-like composite response to mechanical and thermal loading-unloading conditions are examined, as well as its creep and relaxation behavior.

This paper is organized as follows. After a brief summary in Section 2 of the thermoviscoelastic model of [Reese and Govindjee 1998a] for monolithic materials and its coupling with evolving damage, the HFGMC analysis is described in Section 3. Section 4 includes the application of the finite strain thermoviscoelastic composite model in various circumstances, followed by conclusions and suggestions for future research.

2. Finite strain thermoviscoelasticity coupled with damage model of monolithic materials

In the present section we briefly present the constitutive behavior of finite strain thermoviscoelastic polymeric materials that exhibit evolving damage. The presentation follows the papers of [Reese and Govindjee 1998a] where no damage is accounted to and [Lin and Schomburg 2003] where evolving damage is included. The present thermoviscoelastic modeling allows finite strain and large deviations from the thermodynamic equilibrium state.

Let X and x denote the location of a point in the material with respect to the initial (Lagrangian) and current systems of coordinates, respectively, and t is the time. In terms of the local deformation gradient tensor $F(X, t)$, $dx = F(X, t)dX$. The deformation gradient F is expressed by the multiplicative decomposition

$$F(X, t) = F^e(X, t)F^v(X, t), \tag{1}$$

where F^e and F^v are the elastic and viscous parts. The Jacobians that correspond to F and F^e are: $J = \det F$ and $J^e = \det F^e$, respectively.

The modeling that is presented herein is based on a single Maxwell and elastic elements, but it can be extended to include several Maxwell elements. The total free energy per unit reference volume is decomposed into equilibrium (EQ) which represents the strain energy of the elastic element and a nonequilibrium (NEQ) part that accounts for the Maxwell element:

$$\psi = \psi^{\text{EQ}} + \psi^{\text{NEQ}}. \tag{2}$$

The equilibrium part is given by

$$\psi^{\text{EQ}} = (1 - D) \left[f_{\text{EQ}} \psi_0^{\text{EQ}} + (e_0)_{\text{EQ}} \left(1 - \frac{\theta}{\theta_0} \right) + c_0 \left(\theta - \theta_0 - \theta \log \frac{\theta}{\theta_0} \right) \right], \tag{3}$$

where D denotes the amount of damage such that $0 \leq D \leq 1$, θ and θ_0 are the current and reference temperatures, respectively, and c_0 is the heat capacity. In this relation f_{EQ} and $(e_0)_{\text{EQ}}$ are of the form

$$f_{\text{EQ}} = \frac{\theta}{\theta_0}, \qquad (e_0)_{\text{EQ}} = K^e \alpha^e \log J \theta_0, \tag{4}$$

where K^e and α^e are the bulk modulus and thermal expansion coefficient, respectively, of the elastic element. It follows from (3) that ψ_0^{EQ} is the equilibrium part of the free energy at the reference temperature θ_0 in the presence of damage.

The nonequilibrium part is given by

$$\psi^{\text{NEQ}} = (1 - D) \left[f_{\text{NEQ}} \psi_0^{\text{NEQ}} + (e_0)_{\text{NEQ}} \left(1 - \frac{\theta}{\theta_0} \right) \right], \tag{5}$$

where

$$f_{\text{NEQ}} = \frac{\theta}{\theta_0} \quad \text{and} \quad (e_0)_{\text{NEQ}} = K^v \alpha^v \log J_e \theta_0, \tag{6}$$

K^v and α^v being the bulk modulus and thermal expansion coefficient, respectively, of the viscous part of the material. Here too, $\psi^{\text{NEQ}} = \psi_0^{\text{NEQ}}$ for $\theta = \theta_0$ and $D \neq 0$.

The Kirchhoff stresses can be derived from the free-energy expressions above:

$$\tau^{\mathrm{EQ}} = 2\boldsymbol{F}\frac{\partial \psi^{\mathrm{EQ}}}{\partial \boldsymbol{C}}\boldsymbol{F}^T \equiv (1-D)\tau_0^{\mathrm{EQ}}, \tag{7}$$

$\boldsymbol{C} = \boldsymbol{F}^T\boldsymbol{F}$ being the right Cauchy–Green deformation tensor, and

$$\tau^{\mathrm{NEQ}} = 2\boldsymbol{F}\frac{\partial \psi^{\mathrm{NEQ}}}{\partial \boldsymbol{C}}\boldsymbol{F}^T = 2\boldsymbol{F}^e\frac{\partial \psi^{\mathrm{NEQ}}}{\partial \boldsymbol{C}^e}\boldsymbol{F}^{eT} \equiv (1-D)\tau_0^{\mathrm{NEQ}}, \tag{8}$$

where $\boldsymbol{C}^e = \boldsymbol{F}^{eT}\boldsymbol{F}^e$ and τ_0^{EQ}, τ_0^{NEQ} correspond to the Kirchhoff stresses of the undamaged material.

Let the left Cauchy–Green tensor $\boldsymbol{B} = \boldsymbol{F}\boldsymbol{F}^T$ be represented in terms of its eigenvalues b_p and unit principal directions $\boldsymbol{e}_p$, $p = 1, 2, 3$:

$$\boldsymbol{B} = \mathrm{diag}\,[b_1, b_2, b_3], \quad \text{i.e.,} \quad \boldsymbol{B} = \sum_{p=1}^{3} b_p\, \boldsymbol{e}_p \otimes \boldsymbol{e}_p. \tag{9}$$

With $J = \det \boldsymbol{F} = \sqrt{b_1 b_2 b_3}$, the volume preserving tensor $\bar{\boldsymbol{B}} = J^{-2/3}\boldsymbol{B}$ can be accordingly represented in the form

$$\bar{\boldsymbol{B}} = \mathrm{diag}\,[\bar{b}_1, \bar{b}_2, \bar{b}_3] = (b_1 b_2 b_3)^{-1/3}\,\mathrm{diag}\,[b_1, b_2, b_3]. \tag{10}$$

The finite strain isothermal contribution ψ_0^{EQ} can be modeled by the Ogden's compressible material representation [Ogden 1984; Holzapfel 2000] as follows

$$\psi_0^{\mathrm{EQ}} = \sum_{p=1}^{3} \frac{\mu_p^e}{\alpha_p^e}\left((\bar{b}_1)^{\alpha_p^e/2} + (\bar{b}_2)^{\alpha_p^e/2} + (\bar{b}_3)^{\alpha_p^e/2} - 3\right) + \frac{K^e}{4}\left(J^2 - 2\log J - 1\right), \tag{11}$$

where μ_p^e and α_p^e are material parameters of the elastic element.

For Maxwell's element, the isothermal free energy ψ_0^{NEQ} is represented by [Reese and Govindjee 1998a]:

$$\psi_0^{\mathrm{NEQ}} = \sum_{p=1}^{3} \frac{\mu_p^v}{\alpha_p^v}\left((\bar{b}_1^e)^{\alpha_p^v/2} + (\bar{b}_2^e)^{\alpha_p^v/2} + (\bar{b}_3^e)^{\alpha_p^v/2} - 3\right) + \frac{K^v}{4}\left((J^e)^2 - 2\log J^e - 1\right), \tag{12}$$

where

$$\boldsymbol{B}^e = \boldsymbol{F}^e[\boldsymbol{F}^e]^T = \mathrm{diag}\,[b_1^e, b_2^e, b_3^e] \tag{13}$$

and $J^e = \sqrt{b_1^e b_2^e b_3^e}$, $\bar{b}_A^e = (J^e)^{-2/3}b_A^e$, and μ_p^v, α_p^v, are material parameters.

The entropy of the system can be determined from

$$\begin{aligned}
\eta &= -\frac{\partial \psi}{\partial \theta} = -\frac{\partial \psi^{\mathrm{EQ}}}{\partial \theta} - \frac{\partial \psi^{\mathrm{NEQ}}}{\partial \theta} \\
&= -(1-D)\left[\frac{1}{\theta_0}\psi_0^{\mathrm{EQ}} - \frac{1}{\theta_0}(e_0)_{\mathrm{EQ}} - c_0\log\frac{\theta}{\theta_0}\right] - (1-D)\left[\frac{1}{\theta_0}\psi_0^{\mathrm{NEQ}} - \frac{1}{\theta_0}(e_0)_{\mathrm{NEQ}}\right] \\
&\equiv (1-D)\eta_0.
\end{aligned} \tag{14}$$

The evolution equation for the internal variables is given in [Reese and Govindjee 1998a]:

$$-\frac{1}{2}L_v[\boldsymbol{B}^e][\boldsymbol{B}^e]^{-1} = \frac{1}{2\eta_D}\mathrm{dev}\,[\tau^{\mathrm{NEQ}}] + \frac{1}{9\eta_V}\mathrm{trace}\,[\tau^{\mathrm{NEQ}}], \tag{15}$$

where η_D and η_V are the deviatoric and volumetric viscosities, respectively, and $L_v[\boldsymbol{B}^e]$, the Lie derivative of $\boldsymbol{B}^e$, can be expressed as

$$L_v[\boldsymbol{B}^e] = \boldsymbol{F}\dot{\boldsymbol{C}}^{v-1}\boldsymbol{F}^T, \tag{16}$$

with $\boldsymbol{C}^v = \boldsymbol{F}^{vT}\boldsymbol{F}^v$. For elastic bulk behavior, $1/\eta_V = 0$ and the relaxation time is given by $\xi = \eta_D/\mu$ where μ is the small strain shear modulus of the Maxwell element (the nonequilibrium part).

The integration of the evolution equation (15) is performed by means of the return mapping algorithm in conjunction with the logarithmic strain and the backward exponential approximation which were developed in the framework of elastoplasticity, see [Weber and Anand 1990; Eterovic and Bathe 1990; Cuitino and Ortiz 1992; Simo 1992]. Thus, by employing the exponential mapping algorithm, (15) is reduced to

$$\epsilon_{n+1,A}^e = \epsilon_{n+1,A}^{e \text{ trial}} - \Delta t \left[\frac{1}{2\eta_D}\text{dev}\,[\tau_A^{\text{NEQ}}] + \frac{1}{9\eta_V}\text{trace}\,[\tau^{\text{NEQ}}]\right]_{n+1}, \tag{17}$$

with $A = 1, 2, 3$, where the principal values of the elastic logarithmic strain ϵ_A^e are given by $\epsilon_A^e = \frac{1}{2}\log b_A^e$ and Δt is the time increment between the current and previous step. In (17), the trial values of $\epsilon_{n+1,A}^{e \text{ trial}}$ can be expressed in terms of the eigenvalues $b_{n+1,A}^{e \text{ trial}}$ of $\boldsymbol{B}_{n+1}^{e \text{ trial}}$ namely, $\epsilon_{n+1,A}^{e \text{ trial}} = \frac{1}{2}\log b_{A,n+1}^{e \text{ trial}}$, where

$$\boldsymbol{B}_{n+1}^{e \text{ trial}} = \boldsymbol{f}_{n+1}\boldsymbol{B}_n^e\boldsymbol{f}_{n+1}^T, \tag{18}$$

with

$$\boldsymbol{f}_{n+1} = \boldsymbol{F}_{n+1}\boldsymbol{F}_n^{-1}, \tag{19}$$

Equation (17) forms a system of coupled nonlinear equations in the unknowns $\epsilon_{n+1,A}^e$, $A = 1, 2, 3$. It can be rewritten in terms of the elastic logarithmic strain increments in the form

$$\Delta\epsilon_A^e = \epsilon_{n+1,A}^{e \text{ trial}} - \epsilon_{n,A}^e - \Delta t \left[\frac{1}{2\eta_D}\text{dev}\,[\tau_A^{\text{NEQ}}] + \frac{1}{9\eta_V}\text{trace}\,[\tau^{\text{NEQ}}]\right]_{n+1}. \tag{20}$$

In this equation, the principal values of τ^{NEQ} are given according to (8) by

$$\tau_A^{\text{NEQ}} = 2\frac{\partial\psi^{\text{NEQ}}}{\partial b_A}b_A = 2\frac{\partial\psi^{\text{NEQ}}}{\partial b_A^e}b_A^e. \tag{21}$$

The rate of damage evolution is given by (see [Lin and Schomburg 2003; Miehe and Keck 2000])

$$\dot{D} = \frac{\dot{z}}{\eta_{\text{dam}}}(D^\infty - D), \tag{22}$$

where the rate of kinematic arc length is defined by

$$\dot{z} = \sqrt{\frac{2}{3}}\|\dot{\boldsymbol{H}}\|, \qquad \boldsymbol{H} = \frac{1}{2}\log\boldsymbol{C}, \tag{23}$$

with the saturation value

$$D^\infty = \frac{1}{1 + D_0^\infty\exp\left(-\beta_{\text{dam}}/\alpha_{\text{dam}}\right)} \tag{24}$$

and

$$\beta_{\text{dam}} = \max_{0 \le \xi \le t}\sqrt{\frac{2}{3}}\|\boldsymbol{H}(\xi)\|. \tag{25}$$

In these relations, η_{dam}, D_0^∞ and α_{dam} are material parameters.

The incremental form of the constitutive equations of the finite thermoviscoelastic material and the corresponding instantaneous tangent tensor that are needed in the following micromechanical analysis, are determined as follows. From (8) the following expression can be established

$$\Delta \tau_A^{\mathrm{NEQ}} = (1 - D)\Delta \tau_{0A}^{\mathrm{NEQ}} - \tau_{0A}^{\mathrm{NEQ}} \Delta D. \tag{26}$$

Let the second-order tensor $\boldsymbol{M}^{\mathrm{NEQ}}$ be defined by

$$\boldsymbol{M}^{\mathrm{NEQ}} = [M_{AB}^{\mathrm{NEQ}}] \equiv \left[\frac{\partial \tau_{0A}^{\mathrm{NEQ}}}{\partial \epsilon_B^e} \right], \qquad A, B = 1, 2, 3, \tag{27}$$

In addition, the components of the thermal stress vector $\boldsymbol{\Gamma}^{\mathrm{NEQ}}$ are determined from

$$\Gamma_A^{\mathrm{NEQ}} = -\frac{\partial \tau_{0A}^{\mathrm{NEQ}}}{\partial \theta}, \qquad A = 1, 2, 3. \tag{28}$$

In conjunction with (20), we obtain from (26) that

$$\Delta \tau_A^{\mathrm{NEQ}} = (1 - D)M_{AB}^{\mathrm{NEQ}} \left\{ \epsilon_{n+1,B}^{e \ \mathrm{trial}} - \epsilon_{n,B}^e - \Delta t \left[\frac{1}{2\eta_D} \mathrm{dev} \, [\tau_B^{\mathrm{NEQ}}] + \frac{1}{9\eta_V} \mathrm{trace} \, [\tau^{\mathrm{NEQ}}] \right]_{n+1} \right\}$$
$$- (1 - D)\Gamma_A^{\mathrm{NEQ}} \Delta \theta - \tau_{0A}^{\mathrm{NEQ}} \Delta D. \tag{29}$$

Let $\Delta \epsilon_A^e$ and $\Delta \epsilon_A^{\mathrm{ved}}$ be defined by

$$\Delta \epsilon_A \equiv \epsilon_{n+1,A}^{e \ \mathrm{trial}} - \epsilon_{n,A}^e,$$
$$\Delta \epsilon_A^{\mathrm{ved}} \equiv \Delta t \left[\frac{1}{2\eta_D} \mathrm{dev} \, [\tau_A^{\mathrm{NEQ}}] + \frac{1}{9\eta_V} \mathrm{trace} \, [\tau^{\mathrm{NEQ}}] \right]_{n+1} + [M^{\mathrm{NEQ}}]_{AB}^{-1} \tau_{0B}^{\mathrm{NEQ}} \frac{\Delta D}{1 - D} \cdot \tag{30}$$

Therefore (29) can be represented by

$$\Delta \tau_A^{\mathrm{NEQ}} = (1 - D)\left\{ M_{AB}^{\mathrm{NEQ}}[\Delta \epsilon_B - \Delta \epsilon_B^{\mathrm{ved}}] - \Gamma_A^{\mathrm{NEQ}} \Delta \theta \right\}$$
$$\equiv (1 - D)[M_{AB}^{\mathrm{NEQ}} \Delta \epsilon_B - \Gamma_A^{\mathrm{NEQ}} \Delta \theta] - \Delta W_A^{\mathrm{NEQ}}, \tag{31}$$

where the components $\Delta W_A^{\mathrm{NEQ}}$ involve the thermoviscoelastic and damage effects.

The fourth-order tangent tensor $\boldsymbol{d}^{\mathrm{NEQ}}$ is defined by

$$\boldsymbol{d}^{\mathrm{NEQ}} = 2\frac{\partial \boldsymbol{S}^{\mathrm{NEQ}}}{\partial \boldsymbol{C}} = 4\frac{\partial^2 \psi^{\mathrm{NEQ}}}{\partial \boldsymbol{C} \partial \boldsymbol{C}}, \tag{32}$$

where $\boldsymbol{S}^{\mathrm{NEQ}}$ is the second Piola–Kirchhoff stress tensor:

$$\boldsymbol{S}^{\mathrm{NEQ}} = 2\frac{\partial \psi^{\mathrm{NEQ}}}{\partial \boldsymbol{C}} = 2[\boldsymbol{F}^v]^{-1}\frac{\partial \psi^{\mathrm{NEQ}}}{\partial \boldsymbol{C}^e}[\boldsymbol{F}^v]^{-T}, \tag{33}$$

whose the principal values are given by

$$S_A^{\mathrm{NEQ}} = 2\frac{\partial \psi^{\mathrm{NEQ}}}{\partial b_A} = \frac{2}{b_A^v}\frac{\partial \psi^{\mathrm{NEQ}}}{\partial b_A^e}, \qquad A = 1, 2, 3, \tag{34}$$

with $b_A^v = b_A/b_A^e$ being the principal values of $\boldsymbol{B}^v = \boldsymbol{F}^v[\boldsymbol{F}^v]^T = \text{diag } [b_1^v, b_2^v, b_3^v]$. The principal values of $\boldsymbol{d}^{\text{NEQ}}$ can be determined from the following expression [Holzapfel 2000]:

$$\boldsymbol{d}^{\text{NEQ}} = \sum_{A=1}^{3} \sum_{B=1}^{3} \frac{1}{\lambda_B} \frac{\partial S_A^{\text{NEQ}}}{\partial \lambda_B} \boldsymbol{N}_A \otimes \boldsymbol{N}_A \otimes \boldsymbol{N}_B \otimes \boldsymbol{N}_B$$

$$+ \sum_{A=1}^{3} \sum_{B \neq A=1}^{3} \frac{S_B^{\text{NEQ}} - S_A^{\text{NEQ}}}{(\lambda_B)^2 - (\lambda_A)^2} (\boldsymbol{N}_A \otimes \boldsymbol{N}_B \otimes \boldsymbol{N}_A \otimes \boldsymbol{N}_B + \boldsymbol{N}_A \otimes \boldsymbol{N}_B \otimes \boldsymbol{N}_B \otimes \boldsymbol{N}_A), \quad (35)$$

where

$$\frac{1}{\lambda_B} \frac{\partial S_A^{\text{NEQ}}}{\partial \lambda_B} = 2 \frac{\partial S_A^{\text{NEQ}}}{\partial b_B} = \frac{4}{b_A^v b_B^v} \frac{\partial^2 \psi^{\text{NEQ}}}{\partial b_A^e b_B^e}, \tag{36}$$

with $\lambda_A = \sqrt{b_A}$ and $\boldsymbol{N}_A$ denotes the principal referential orthonormal directions. It should be noted that for $\lambda_A = \lambda_B$, a Taylor expansion shows that

$$\lim_{\lambda_B \to \lambda_A} \frac{S_B^{\text{NEQ}} - S_A^{\text{NEQ}}}{(\lambda_B)^2 - (\lambda_A)^2} = \frac{1}{2\lambda_B} \left[\frac{\partial S_B^{\text{NEQ}}}{\partial \lambda_B} - \frac{\partial S_A^{\text{NEQ}}}{\partial \lambda_B} \right]. \tag{37}$$

The second-order thermal stress tensor $\boldsymbol{\gamma}^{\text{NEQ}}$ is determined from

$$\boldsymbol{\gamma}^{\text{NEQ}} = -\frac{\partial \boldsymbol{S}^{\text{NEQ}}}{\partial \theta} = -2 \frac{\partial^2 \psi^{\text{NEQ}}}{\partial \boldsymbol{C} \partial \theta} \tag{38}$$

and its principal values are given by

$$\gamma_A^{\text{NEQ}} = -2 \frac{\partial^2 \psi^{\text{NEQ}}}{\partial b_A \partial \theta} = -\frac{2}{b_A^v} \frac{\partial^2 \psi^{\text{NEQ}}}{\partial b_A^e \partial \theta}. \tag{39}$$

The values of γ_A^{NEQ} can be readily related to Γ_A^{NEQ} in (28) by employing the relation: $\tau = \boldsymbol{F} \boldsymbol{S} \boldsymbol{F}^T$ from which the equality $\Gamma_A^{\text{NEQ}} = b_A \gamma_A^{\text{NEQ}}$ is obtained.

The fourth-order first tangent tensor $\boldsymbol{R}^{\text{NEQ}}$ which is defined by

$$\boldsymbol{R}^{\text{NEQ}} = \frac{\partial \boldsymbol{T}^{\text{NEQ}}}{\partial \boldsymbol{F}}, \tag{40}$$

where $\boldsymbol{T}^{\text{NEQ}}$ is the first Piola–Kirchhoff stress tensor, can be determined from

$$\boldsymbol{R}^{\text{NEQ}} = \boldsymbol{F} \, \boldsymbol{d}^{\text{NEQ}} \, \boldsymbol{F}^T + \boldsymbol{S}^{\text{NEQ}} \otimes \boldsymbol{I}, \tag{41}$$

with $\boldsymbol{I}$ denoting the unit second-order tensor. Thus, the rate form of the nonequilibrium portion of the constitutive equations of the finite thermoviscoelastic material is given by

$$\dot{\boldsymbol{T}}^{\text{NEQ}} = \boldsymbol{R}^{\text{NEQ}} : \dot{\boldsymbol{F}} - \boldsymbol{H}^{\text{NEQ}} \dot{\theta} - \dot{\boldsymbol{V}}^{\text{NEQ}}, \tag{42}$$

where by taking into account the relation between the Kirchhoff τ and the first Piola–Kirchhoff $\boldsymbol{T}$ stress tensors: $\tau = \boldsymbol{F}\boldsymbol{T}$, the following expressions for the thermal stress $\boldsymbol{H}^{\text{NEQ}}$ and viscous-damage $\dot{\boldsymbol{V}}^{\text{NEQ}}$

terms can be established:

$$H^{\mathrm{NEQ}} = F^{-1} \Gamma^{\mathrm{NEQ}} = \gamma^{\mathrm{NEQ}} F^T, \tag{43}$$

$$\dot{V}^{\mathrm{NEQ}} = F^{-1} \dot{W}^{\mathrm{NEQ}}. \tag{44}$$

The same procedure can be followed for the establishment of the first tangent tensor R^{EQ} of equilibrium elastic element where $\dot{W}^{\mathrm{EQ}}$ involves this time the damage effects only and $F^v = I$. It yields

$$\dot{T}^{\mathrm{EQ}} = R^{\mathrm{EQ}} : \dot{F} - H^{\mathrm{EQ}}\dot{\theta} - \dot{V}^{\mathrm{EQ}}. \tag{45}$$

The final total rate form of the finite thermoviscoelastic material is as follows

$$\dot{T} = R : \dot{F} - H\dot{\theta} - \dot{V}, \tag{46}$$

where $\dot{T} = \dot{T}^{\mathrm{NEQ}} + \dot{T}^{\mathrm{EQ}}$, $R = R^{\mathrm{NEQ}} + R^{\mathrm{EQ}}$, $H = H^{\mathrm{NEQ}} + H^{\mathrm{EQ}}$, and $\dot{V} = \dot{V}^{\mathrm{NEQ}} + \dot{V}^{\mathrm{EQ}}$. Constitutive equations can be obtained from (46) as specials cases in the presence/absence of damage and viscous effects.

3. Finite-strain micromechanical analysis

Finite strain HFGMC micromechanical analyses for the establishment of the macroscopic constitutive equations of various types of composites with doubly periodic microstructure undergoing large deformations have been previously reviewed by [Aboudi 2008]. These micromechanical analyses are based on the homogenization technique in which a repeating unit cell of the periodic composite can identified. This repeating unit cell represents the periodic composite in which the double periodicity is taken in the transverse $2 - 3$ plane, so that the axial 1-direction corresponds to the continuous direction; see Figure 1. (For a unidirectional fiber-reinforced material, for example, the 1-direction coincides with the
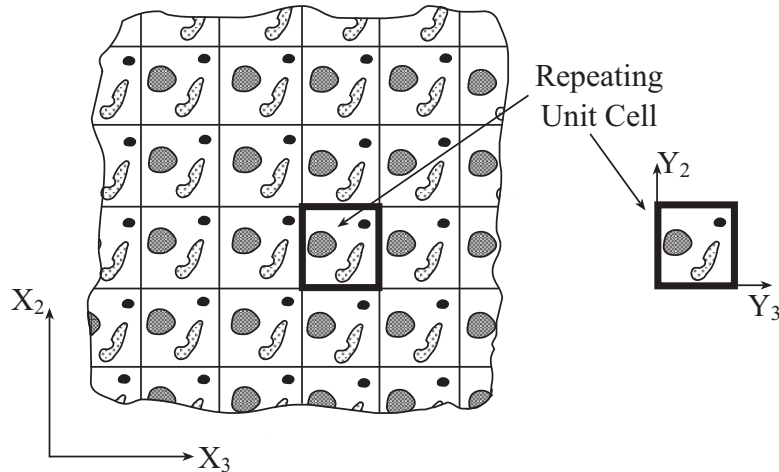

Figure 1. A multiphase composite with doubly periodic microstructures defined with respect to global initial coordinates of the plane $X_2 - X_3$. The repeating unit cell is defined with respect to local initial coordinates of the plane $Y_2 - Y_3$.

fiber orientation.) In the framework of these HFGMC micromechanical models, the displacements are asymptotically expanded and the repeating unit cell is discretized. This is followed by imposing the equilibrium equations, the displacement and traction interfacial conditions as well as the conditions that ensure that the displacements and tractions are periodic across the repeating unit cell. In particular, the imposition of the equilibrium equations provide the strong form of the Lagrangian equilibrium conditions of the homogenization theory that must be satisfied. In addition, since the solution of the repeating unit cell is determined with a constant vector, the corners of this cell are rigidly clamped to prevent this indeterminacy. The resulting homogenization derivation establishes the deformation concentration tensor $A(Y, \theta)$, where Y are the local Lagrangian system of coordinates with respect to which field variables in the repeating unit cell are characterized. This tensor relates the rate of the local deformation gradient $\dot{F}(Y, \theta)$ at a material point Y within the repeating unit cell to the externally applied deformation gradient rate $\dot{\bar{F}}$ on the composite in the form

$$\dot{F}(Y, \theta) = A(Y, \theta) : \dot{\bar{F}} + A^{\text{th}}(Y, \theta)\dot{\theta} + \dot{A}^{\text{vd}}(Y, \theta), \tag{47}$$

where A^{th} and $\dot{A}^{\text{vd}}$ are the thermal concentration tensor and the viscous-damage contribution. The mechanical concentration tensor A is determined at every increment of loading when the thermal and viscous-damage effects are absent. Similarly, the thermal concentration tensor A^{th} is determined at every increment in the absence of mechanical and viscous-damage effects. Finally, the current $\dot{A}^{\text{vd}}$ can be determined when no mechanical or thermal effects are present. It follows from (47) and in conjunction with (46) that the local stress rate at this point is given by

$$\dot{T}(Y, \theta) = R(Y, \theta) : \left[A(Y, \theta) : \dot{\bar{F}} + A^{\text{th}}(Y, \theta)\dot{\theta} + \dot{A}^{\text{vd}}(Y, \theta) \right] - H(Y, \theta)\dot{\theta} - \dot{V}(Y, \theta). \tag{48}$$

Hence the resulting rate form of the macroscopic constitutive equation for the multiphase thermoviscoelastic composite undergoing large deformation is given by

$$\dot{\bar{T}} = R^* : \dot{\bar{F}} - H^*\dot{\theta} - \dot{\bar{V}}, \tag{49}$$

where $\dot{\bar{T}}$ is the rate of the overall (global) first Piola–Kirchhoff stress tensor and R^* is the instantaneous effective stiffness (first tangent) tensor of the multiphase composite. It can be expressed in terms of the first tangent tensors of the constituents $R(Y, \theta)$ and the established deformation concentration tensor $A(Y, \theta)$ in the form

$$R^* = \frac{1}{S_Y} \iint_{S_Y} R(Y, \theta) \, A(Y, \theta) \, dS_Y, \tag{50}$$

where S_Y is the area of the repeating unit cell. The effective instantaneous thermal stress tensor H^* in (49) is established from

$$H^* = -\frac{1}{S_Y} \iint_{S_Y} \left[R(Y, \theta) \, A^{\text{th}}(Y, \theta) - H(Y, \theta) \right] \, dS_Y. \tag{51}$$

Finally, the current viscous-damage contribution to the macroscopic constitutive equations (49) is given by

$$\dot{\bar{V}} = -\frac{1}{S_Y} \iint_{S_Y} \left[R(Y, \theta) : \dot{A}^{\text{vd}}(Y, \theta) - \dot{V}(Y, \theta) \right] dS_Y. \tag{52}$$

More details can be found in [Aboudi 2008]. It should be noted that the current values of $\boldsymbol{R}^*$, $\boldsymbol{H}^*$ and $\dot{\boldsymbol{V}}$ of the composite are affected by the current value of damage variable through the instantaneous value of the tensors $\boldsymbol{R}(Y, \theta)$, $\boldsymbol{H}(Y, \theta)$ and $\dot{\boldsymbol{V}}(Y, \theta)$ of the finite strain constituents.

The finite strain HFGMC micromechanical model predictions were assessed and verified by comparison with analytical and numerical large deformation solutions in [Aboudi and Pindera 2004; Aboudi 2009] for composites with hyperelastic and hyperelastic coupled with damage constituents, respectively.

4. Applications

In the present section, applications are given that exhibit under various circumstances the response of a composite undergoing large deformations, which consists of a thermoviscoelastic rubber-like material reinforced by continuous thermoelastic fibers. The thermoviscoelastic matrix is characterized by the free energy functions (3) and (5) that represent elastic and maxwell elements, respectively in conjunction with the corresponding isothermal free-energy functions (11) and (12). The parameters in these functions [Reese and Govindjee 1998a] are given in Tables 1 and 2, together with η_D and $1/\eta_V = 0$ (assuming elastic bulk deformations). The damage mechanism affects the thermoviscoelastic matrix only and its parameters which appear in (22)–(24) are: $\eta_{\text{dam}} = 0.1$, $D_0^\infty = 1$ and $\alpha_{\text{dam}} = 1$. The effect of damage can be totally neglected by choosing $1/\eta_{\text{dam}} = 0$.

The continuous thermoelastic steel fibers are oriented in the 1-direction and they are characterized by the free energy function [Reese and Govindjee 1998a]

$$\psi^{\text{steel}} = \frac{(\mu_1)_{st}}{(\alpha_1)_{st}} \left[(\bar{b}_1)^{(\alpha_1)_{st}/2} + (\bar{b}_2)^{(\alpha_1)_{st}/2} + (\bar{b}_3)^{(\alpha_1)_{st}/2} - 3 \right] + \frac{K_{st}}{4} \left[J^2 - 2 \log J - 1 \right], \qquad (53)$$

where $(\mu_1)_{st}$, $(\alpha_1)_{st}$ and K_{st} are material parameters of the steel fibers which are given in Table 3. The volume fraction of the fibers is $v_f = 0.05$ which is characteristic for a rubber-like material reinforced by steel fibers.

μ_1^e (MPa)	μ_2^e (MPa)	μ_3^e (MPa)	α_1^e	α_2^e	α_3^e	K^e (MPa)	$\alpha^e (K^{-1})$	c_0^e (MPa K^{-1})
0.13790	−0.04827	0.01034	1.8	−2	7	50	223.33×10^{-6}	1.7385

Table 1. Material parameters in the function ψ_0^{EQ} of (11). The parameters μ_p^e and α_p^e, $p = 1, 2, 3$ are the Ogden's material constants, K^e is the bulk modulus, α^e is its thermal expansion coefficient and c_0^e is its heat capacity. In the small strain domain, the shear modulus of this material is 0.208 MPa.

μ_1^v (MPa)	μ_2^v (MPa)	μ_3^v (MPa)	α_1^v	α_2^v	α_3^v	K^v (MPa)	η_D (MPa s)	α^v (K^{-1})
0.3544	−0.1240	0.0266	1.8	−2	7	50	9.38105	223.33×10^{-6}

Table 2. Material parameters in the function ψ_0^{NEQ} of (12). The parameters μ_p^v and α_p^v, $p = 1, 2, 3$ are the Ogden's material constants, K^v is the bulk modulus and α^v is its thermal expansion coefficient. η_D and η_V are the viscoelastic constants with $\eta_V \to \infty$ implying elastic bulk behavior. In the small strain domain, the shear modulus of this material is 0.536 MPa.

$(\mu_1)_{st}$ (MPa)	$(\alpha_1)_{st}$	K_{st} (MPa)	$(\alpha)_{st}(K^{-1})$	$(c_0)_{st}$ (MPa K^{-1})
80769.231	2	121153.85	12×10^{-6}	3.768

Table 3. Material parameters in the function ψ^{steel} of (53). The parameters $(\mu)_{st}$ and $(\alpha)_{st}$ are the Ogden's material constants of the steel fibers, K_{st} is the bulk modulus, $(\alpha)_{st}$ its coefficient of thermal expansion and $(c_0)_{st}$ is its heat capacity. In the small strain domain, the shear modulus of the material is 80769.231 MPa.

4.1. *The thermoelastic inversion effect.* Consider a uniaxially stretched specimen of rubber in the 1-direction that is subjected to thermal loading (isomeric behavior). For low values of constant stretch, the graph of the gradient of the stress against temperature is negative (as in glass and metals), but becomes positive at certain critical stretching. This change of gradient sign characterizes the thermoelastic inversion effect also referred to as the Gough–Joule effect. Similarly, in ordinary materials with a prescribed stress and subjected to a thermal loading (isotonic behavior), the gradient of deformation with respect to temperature is always positive, but in rubbers this gradient becomes negative at loadings beyond a critical value. As was shown by [Ogden 1992], the derivative of the stress T_{11} with respect to the temperature θ at constant stretch λ_1: $(\partial T_{11}/\partial \theta)_{\lambda_1}$, and the derivative of the temperature θ with respect to the stretch λ_1 at constant entropy η: $(\partial \theta/\partial \lambda_1)_\eta$ (isentropic behavior) vanish simultaneously. Thus, it is possible to detect the critical value of stretch at which the thermoelastic inversion takes place by considering a strip of rubber subjected to a uniaxial stress loading under isentropic conditions, namely $\dot{\eta} = 0$. The minimum value of the generated temperature against applied stretch graph corresponds to the critical inversion point.

For a homogeneous material the (undamaged) entropy η_0 is given by (14). Under uniaxial stress loading in the 1-direction, the principal values b_2 and b_3 of the left Cauchy–Green deformation tensor **B** can be expressed in terms of the stretch $\lambda_1 = \sqrt{b_1}$ and temperature θ. Hence $\eta_0 = \eta_0(\lambda_1, \theta)$. For isentropic stretching

$$d\eta_0(\lambda_1, \theta) = \left(\frac{\partial \eta_0}{\partial \lambda_1}\right)_\theta d\lambda_1 + \left(\frac{\partial \eta_0}{\partial \theta}\right)_{\lambda_1} d\theta = 0, \tag{54}$$

where

$$\left(\frac{\partial \eta_0}{\partial \lambda_A}\right)_\theta = 2\sqrt{\frac{b_A^e}{b_A^v}} \frac{\partial \eta_0}{\partial b_A^e}, \qquad A = 1, 2, 3. \tag{55}$$

This relation provides

$$\frac{d\theta}{d\lambda_1} = -\frac{(\partial \eta_0/\partial \lambda_1)_\theta}{(\partial \eta_0/\partial \theta)_{\lambda_1}}. \tag{56}$$

The integration of this differential equation provides the graph of temperature deviation $\Delta\theta = \theta - \theta_0$ against the stretch λ_1 which shows initially a falling and then rising temperature, thus exhibiting the thermoelastic inversion effect. The critical value of stretch where thermoelastic inversion effect takes place correspond to the minimum of this curve.

For a rubber specimen that is subjected to prescribed stress, the critical value of the stress at which the slope of $(\partial \lambda_1 / \partial \theta)_{T_{11}}$ changes sign can be determined as follows. The Maxwell relations

$$\left(\frac{\partial T_{11}}{\partial \theta}\right)_{\lambda_1} = -\left(\frac{\partial \eta_0}{\partial \lambda_1}\right)_\theta, \qquad \left(\frac{\partial \lambda_1}{\partial \theta}\right)_{T_{11}} = \left(\frac{\partial \eta_0}{\partial \lambda_1}\right)_\theta \tag{57}$$

show that $(\partial T_{11} / \partial \theta)_{\lambda_1} = 0$ yields $(\partial \eta_0 / \partial \lambda_1)_\theta = 0$; namely, $(\partial \lambda_1 / \partial \theta)_{T_{11}} = 0$. But from

$$dT_{11} = \left(\frac{\partial T_{11}}{\partial \lambda_1}\right)_\theta d\lambda_1 + \left(\frac{\partial T_{11}}{\partial \theta}\right)_{\lambda_1} d\theta \tag{58}$$

and (54) we obtain that $(\partial \eta_0 / \partial \lambda_1)_\theta = 0$ implies that $(\partial \theta / \partial T_{11})_{\eta_0} = 0$. Hence the critical value of stress T_{11} can be determined from the minimum of the curve θ vs. T_{11} generated during an isentropic procedure.

In order to determine the thermoelastic inversion effect in thermoviscoelastic composites, the following tensor is defined in terms of its principal values $\partial \eta_0 / \partial \lambda_p$ and unit principal directions $\boldsymbol{k}_p$, $p = 1, 2, 3$:

$$\boldsymbol{K} = \frac{\partial \eta_0}{\partial \boldsymbol{F}} = \text{diag}\left[\frac{\partial \eta_0}{\partial \lambda_1}, \frac{\partial \eta_0}{\partial \lambda_2}, \frac{\partial \eta_0}{\partial \lambda_3}\right], \quad \text{i.e.,} \quad \boldsymbol{K} = \sum_{p=1}^{3} \frac{\partial \eta_0}{\partial \lambda_p} \boldsymbol{k}_p \otimes \boldsymbol{k}_p. \tag{59}$$

It follows that the increment of the local value of the entropy is given by

$$\Delta \eta_0(\boldsymbol{Y}, \theta) = \boldsymbol{K}(\boldsymbol{Y}, \theta) : \Delta \boldsymbol{F}(\boldsymbol{Y}, \theta) + P(\boldsymbol{Y}, \theta) \Delta \theta, \tag{60}$$

where $P = c_0 / \theta$ of the constituent. By substituting (47) in this relation the following expression is established

$$\Delta \eta_0(\boldsymbol{Y}, \theta) = \boldsymbol{K}(\boldsymbol{Y}, \theta) : \left[\boldsymbol{A}(\boldsymbol{Y}, \theta) : \Delta \bar{\boldsymbol{F}} + \boldsymbol{A}^{\text{th}}(\boldsymbol{Y}, \theta) \Delta \theta + \Delta \boldsymbol{A}^{\text{vd}}(\boldsymbol{Y}, \theta) \right] + P(\boldsymbol{Y}, \theta) \Delta \theta. \tag{61}$$

The increment of the global entropy of the composite is given by

$$\Delta \bar{\eta}_0 = \frac{1}{S_Y} \iint_{S_Y} \Delta \eta_0(\boldsymbol{Y}, \theta) \, dS_Y. \tag{62}$$

Consider a thermoviscoelastic composite that subjected to a uniaxial stress loading in the transverse 2-direction such that all components of $\bar{\boldsymbol{T}}$ are zero except $\bar{T}_{22}$. Here, there are eight unknown deformation gradients $\Delta \bar{\boldsymbol{F}}$ (except $\Delta \bar{F}_{22}$) and $\Delta \theta$. There are, on the other hand eight equations $\Delta \bar{\boldsymbol{T}} = \boldsymbol{0}$ (except $\Delta \bar{T}_{22}$) and for isentropic procedure there is the additional relation: $\Delta \bar{\eta}_0 = 0$. Thus, the above relations enable the computation of the temperature θ that is generated at applying a transverse stretch $\bar{F}_{22}$ in a stepwise manner. Figure 2(a) shows the generated temperature deviation $\Delta \theta = \theta - \theta_0$ under isentropic and uniaxial stress loading conditions against the applied average transverse deformation gradient $\bar{F}_{22}$. This figure shows the resulting behavior of the homogeneous (unreinforced) and the unidirectional steel/rubber-like composite where the matrix is considered as thermoviscoelastic (TVE) as well as thermoelastic (TE) in which the viscous effects have been neglected. The minima of the curves correspond to the critical stretch $\bar{F}_{22}$ at which the thermoelastic inversions take place. It can be readily observed that the viscous effects on the critical points are negligible. The critical stretch of the TVE homogeneous matrix occurs at a stretch of $F_{22} = 1.023$, but due to the presence of reinforcement it moves to $\bar{F}_{22} = 1.032$. The corresponding graph which exhibits the resulting temperature deviation $\Delta \theta$ against transverse stress $\bar{T}_{22}$ is shown in Figure 2(b). This latter figure also shows the locations of the inversion thermoelastic points when the
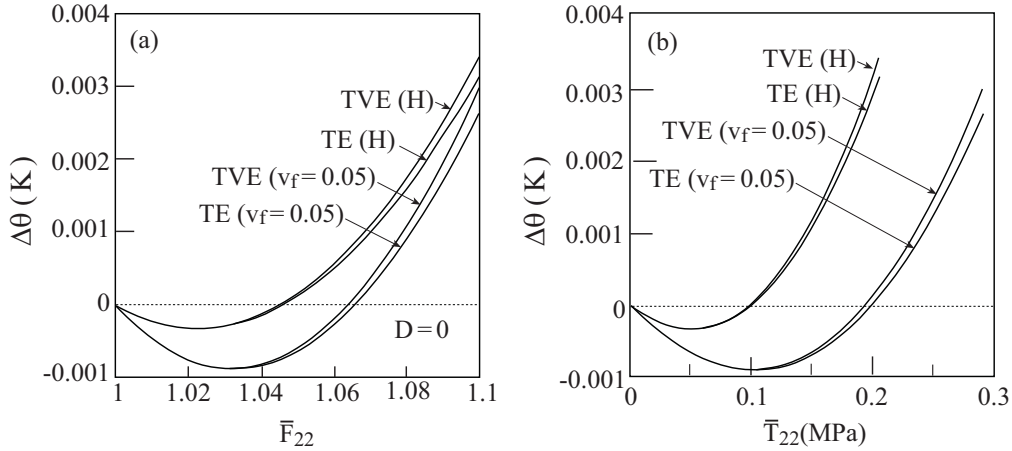
Figure 2. Temperature deviations generated by applying, under isentropic conditions, a uniaxial transverse stress loading on thermoviscoelastic (TVE) and thermoelastic (TE) composites. Also shown are the corresponding temperature deviations generated in the homogeneous (H) unreinforced matrix. The minima correspond to the locations of the thermoelastic inversions. Temperature deviations are plotted against average transverse deformation gradient (a) and against average transverse stress (b).

composite and its homogeneous unreinforced matrix are subjected to a thermal loading in conjunction with a prescribed transverse uniaxial stress. These critical stresses occur at about $\bar{T}_{22} = 0.1\,\text{MPa}$ and $T_{22} = 0.05\,\text{MPa}$ for the composite and its homogeneous matrix, respectively.

4.2. *The composite and its unreinforced matrix responses.* We now study the behaviors of the unidirectional thermoviscoelastic composite and its matrix under various circumstances. Figures 3 and 4

Figure 3. Response of the homogeneous thermoviscoelastic (TVE) and thermoelastic (TE) rubber-like material to uniaxial stress loading in the 1-direction, applied at elevated temperature $\Delta\theta = 100\,\text{K}$ at a rate of $r = \dot{F}_{11} = 0.01/\text{s}$: stress-deformation gradient response (a), damage evolution (b), and stress-deformation gradient response in the absence of damage (c).

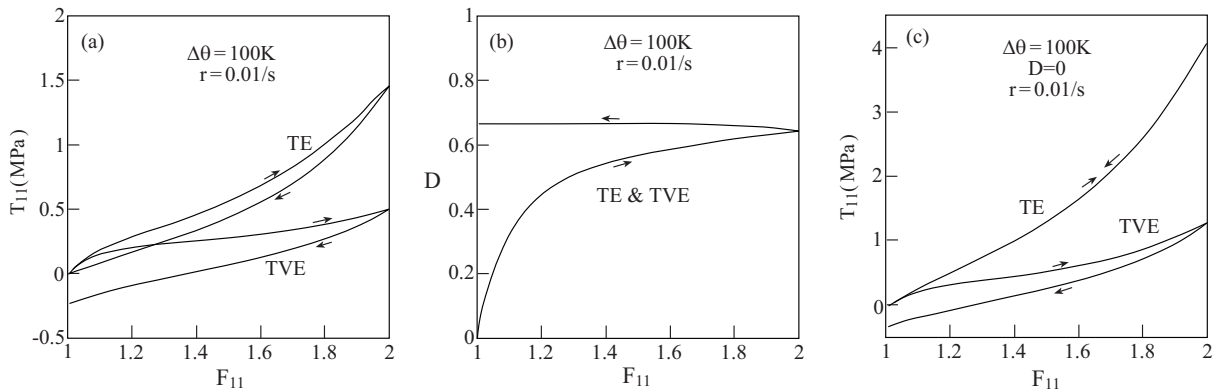show the behavior under uniaxial stress loading in the 1-direction (i.e., when all components T are zero except T_{11}) of the homogeneous (unreinforced) thermoviscoelastic (TVE) rubber-like matrix at elevated temperature $\Delta\theta = 100\,\text{K}$ as well as the thermoelastic (TE) matrix when the viscous effects have been neglected. Figure 3(a) compares the response of the TE and TVE matrix, whereas Figure 3(b) shows the evolution of damage in the matrix as the loading proceeds. Here and in the following, the value of the damage refers to its maximum amount that evolves in all locations of the rubber-like phase. Figure 3(c) shows the counterpart behavior of the homogeneous matrix in the absence of any damage effects (note that the scale of the plot in the latter case is twice that of the damaged case). In all cases the rate of loading is $r = \dot{F}_{11} = 0.01/s$. This figure exhibits very well the significant viscous and damage effects on the rubber-like material behavior.

Figure 4 provides comparisons of the rubber-like material response at the reference temperature (i.e., $\Delta\theta = 0$) such that the material behavior is viscoelastic (VE), and at elevated temperature $\Delta\theta = 100\,\text{K}$
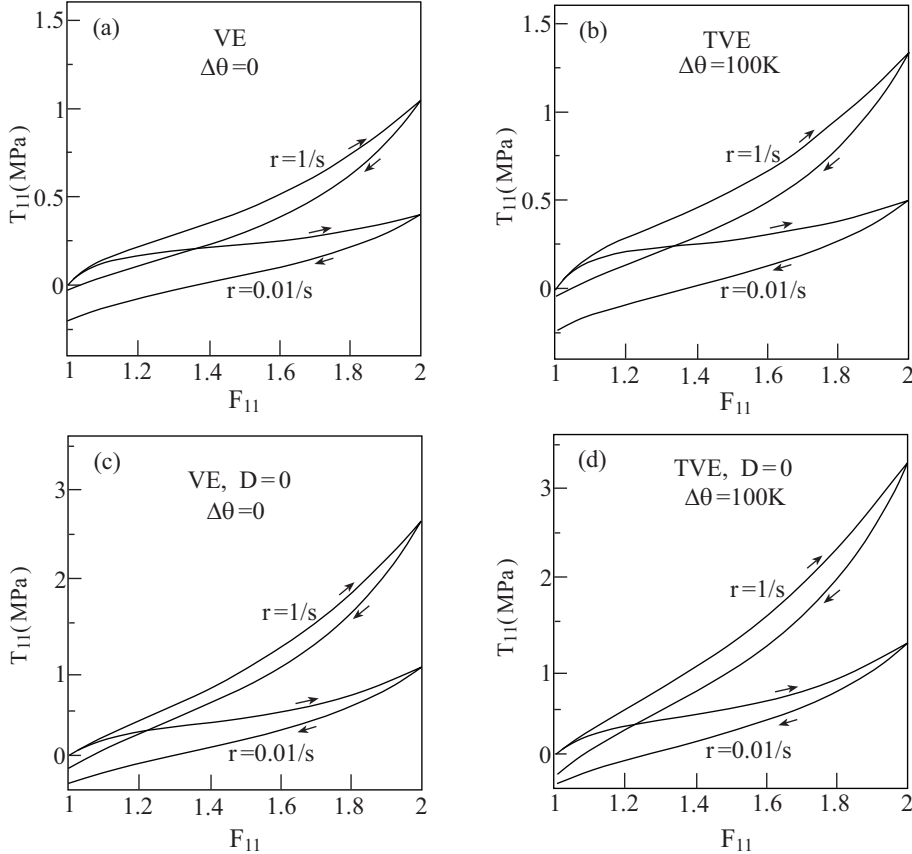
Figure 4. Top row: stress-deformation gradient response of the homogeneous viscoelastic (a) and thermoviscoelastic (b) rubber-like material to uniaxial stress loading in the 1-direction, applied at two rates: $r = \dot{F}_{11} = 1/s$ and $0.01/s$. In (a) the VE material is kept at $\theta = \theta_0$, in (b) the TVE material undergoes a temperature change of $\Delta\theta = 100\,\text{K}$. Bottom row: corresponding results in the absence of damage effects.
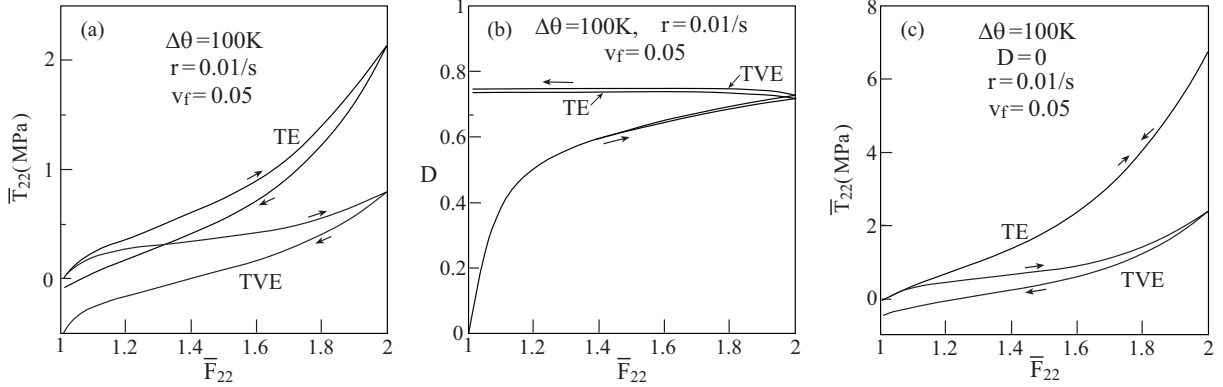
Figure 5. Response of the steel/rubber-like thermoviscoelastic (TVE) and thermoelastic (TE) composite to uniaxial stress loading in the transverse 2-direction, applied at elevated temperature $\Delta\theta = 100$ K at a rate of $r = \dot{\bar{F}}_{22} = 0.01$/s: global stress-deformation gradient response (a), damage evolution (b), and global stress-deformation gradient response in the absence of damage (c).

resulting in a thermoviscoelastic behavior (TVE). In this figure, comparisons are also shown when the uniaxial stress loading is applied at two rates $r = 1$/s and 0.01/s in the presence and absence of damage (in the latter case the scale of the graph is twice the former). Significant differences between the various cases can be clearly observed.

Similar studies can be carried out in order to show the behavior of the unidirectional steel/rubber-like composite. In this case, the application of a transverse uniaxial stress loading perpendicular to the fiber direction is the most interesting loading, since for loading in the fiber direction (1-direction) the much stiffer elastic steel will dominate the response of the composite. In this type of loading all components of the average stress $\bar{T}$ are equal to zero except $\bar{T}_{22}$. The transverse loading is performed by applying the average transverse deformation gradient $\bar{F}_{22}$ at a rate of $r = \dot{\bar{F}}_{22}$. Figure 5 exhibits the viscous and damage effects of the rubber-like phase on the behavior of the composite loaded at a rate of $r = 0.01$/s (note that the scale of the plot in the undamaged case is three times the damaged one). Figure 6 shows the effect of elevated temperature, rate of loading and damage on the macroscopic transverse response of the composite.

Let the unidirectional steel/rubber-like thermoviscoelastic composite be subjected to a cyclic thermal loading at a rate $\dot{\theta} = 1$ K/s while keeping the composite traction-free ($\bar{T} = 0$). In this thermal loading case, the temperature deviation increases/decreases linearly such that -100 K $\leq \Delta\theta \leq 100$ K. The resulting average of the transverse deformation gradient $\bar{F}_{22}$ caused by this thermal loading is shown in Figure 7 during 7 cycles followed by 1/4 cycle after which the applied temperature deviation reaches $\Delta\theta = 100$ K, where the damage reaches the value of $D = 0.5$. This figure shows the transverse deformation gradient in both the presence and absence of evolving damage in the rubber-like matrix. In presence of damage, its evolution with applied thermal loading is shown in Figure 7(b). It can be readily observed that the effect of damage on the resulting transverse deformation gradient is negligible, and that the induced strains are quite small and can be regarded to belong to the infinitesimal domain. In addition, it turns out that the rate of applied thermal loading has no appreciable effect on these strains.
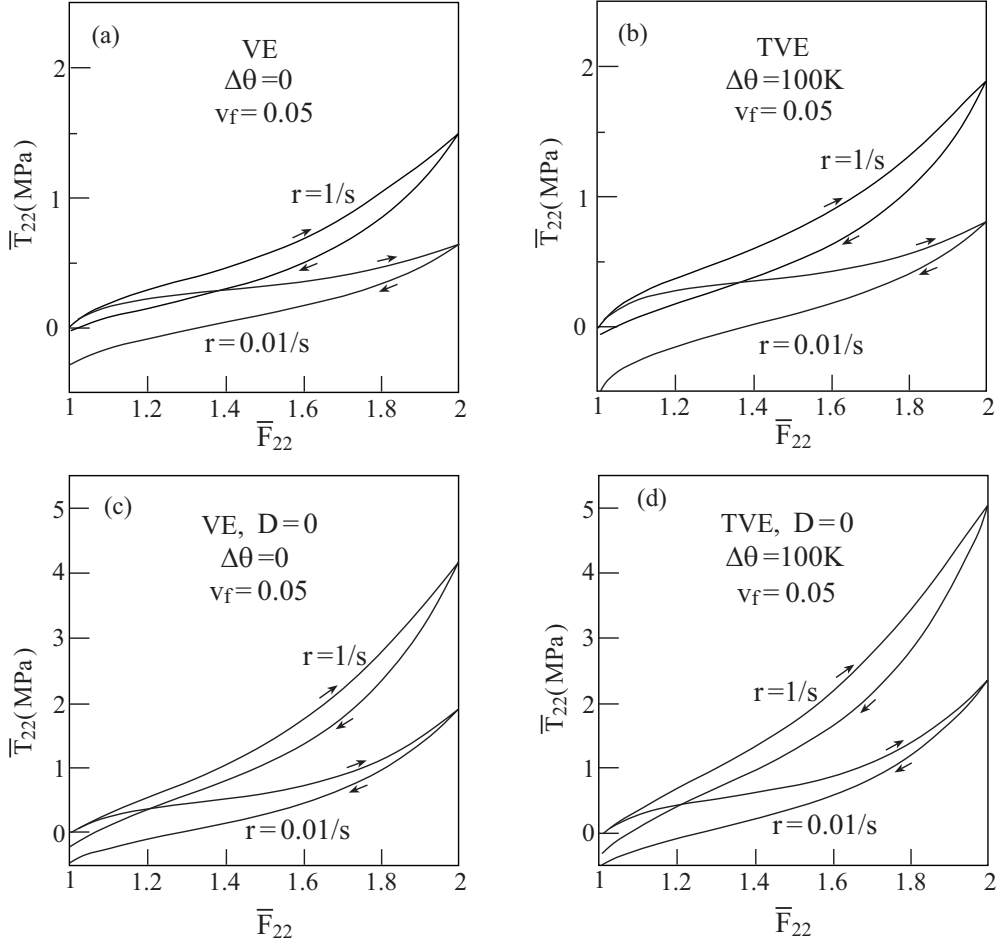
Figure 6. Top row: global stress-deformation gradient response of the steel/rubber-like viscoelastic and thermoviscoelastic composite to uniaxial stress loading in the transverse 2-direction, applied at two rates: $r = \dot{\bar{F}}_{22} = 1/s$ and $0.01/s$. In (a) the VE material is kept at $\theta = \theta_0$, in (b) the TVE material undergoes a temperature change of $\Delta\theta = 100\,K$. Bottom row: corresponding results in the absence of damage effects.

Although the previous thermal loading case that was shown in Figure 7 indicates that the effect of evolving damage on the free-thermal expansion response of the thermoviscoelastic composite appears to be minor, it should be interesting to investigate the response of the composite by applying a uniaxial transverse stress of loading and unloading at a rate of $r = \dot{\bar{F}}_{22} = 1/s$, which immediately follows the previously applied 7.25 cycles of thermal loading, $-100\,K \leq \Delta\theta \leq 100\,K$, which will be referred to as case 1. The response of the composite in case 1 is compared to the response which results by subjecting the steel/rubber-like composite to a uniaxial transverse stress of loading and unloading applied at a rate of $r = \dot{\bar{F}}_{22} = 1/s$ at elevated temperature $\Delta\theta = 100\,K$, referred to as case 2. The resulting comparison is shown in Figure 8 together with the evolving damage in both cases. It is readily observed that although the cyclic thermal loading ends with a damage of $D = 0.5$ in the rubber-like matrix, its effect on the
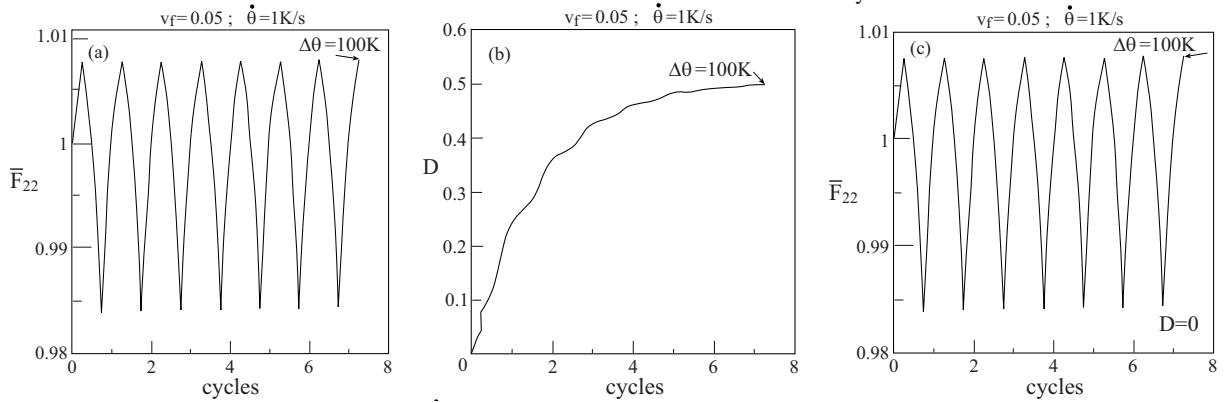
Figure 7. Response of the thermoviscoelastic steel/rubber-like composite that is subjected to cyclic thermal loading $-100\,\text{K} \leq \Delta\theta \leq 100\,\text{K}$ at a rate of $\dot{\theta} = 1$ K/s, while keeping it traction-free: transverse deformation gradient variation with cycles (a), damage evolution with cycles (b), and transverse deformation gradient variation with cycles in the absence of damage (c).
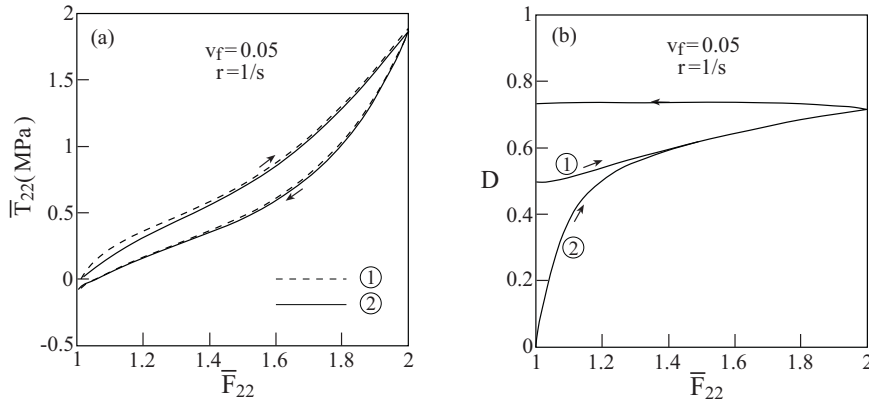
Figure 8. Response of thermoviscoelastic steel/rubber-like composite subjected to (1) cyclic thermal loading $-100\,\text{K} \leq \Delta\theta \leq 100\,\text{K}$ at a rate $\dot{\theta} = 1$ K /s, while staying traction-free, then to a uniaxial transverse stress loading-unloading, applied at a rate of $r = \dot{\bar{F}}_{22} = 1/\text{s}$, at elevated temperature $\Delta\theta = 100\,\text{K}$; and (2) stress loading-unloading alone, as in (1). Plot (a) shows the average transverse stress deformation gradient, plot (b) the damage evolution.

subsequent mechanical transverse loading is quite small. Applying the uniaxial transverse stress loading at the lower rate of $r = \dot{\bar{F}}_{22} = 0.01/\text{s}$ gave the same closeness between the two cases.

The following four figures exhibit the creep behavior of the composite under various circumstances. Figure 9(a) compares the creep behavior of the thermoviscoelastic steel/rubber-like composite with the corresponding behavior of the homogeneous (H) matrix. In both cases a uniaxial transverse stress loading is applied at elevated temperature $\Delta\theta = 100\,\text{K}$ such that all components of the stress $\bar{T}$ are zero except $\bar{T}_{22} = 1\,\text{MPa}$. Figure 9(b) shows the evolving damage to saturation in the thermoviscoelastic matrix of
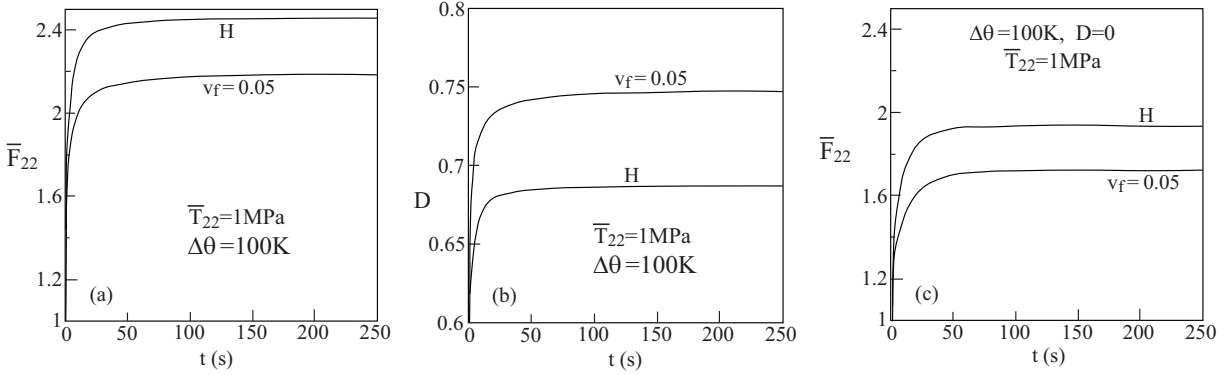
Figure 9. Creep behavior of thermoviscoelastic steel/rubber-like composite subjected to a uniaxial transverse stress loading $\overline{T}_{22} = 1\,\text{MPa}$ at elevated temperature $\Delta\theta = 100\,\text{K}$. Also shown is the corresponding creep behavior of the homogeneous (H) unreinforced thermoviscoelastic matrix. As functions of time, the plots show the global transverse deformation gradient (a), damage evolution (b), and global transverse deformation gradient in the absence of damage effects.

the composite as well is in the unreinforced material. Finally, Figure 9(c) is the counterpart of Figure 9(a) in the absence of any damage effects.

It is interesting to observe that whereas the existence of the steel fibers decreases, as expected, the resulting macroscopic deformation gradient of the composite, the damage induced in the reinforced matrix is higher that the one that evolves in the homogeneous material. Figure 10 compares the creep behavior of thermoviscoelastic steel/rubber-like composite when it is subjected to uniaxial transverse stress loadings of $\overline{T}_{22} = 1\,\text{MPa}$ and $2\,\text{MPa}$ at elevated temperature $\Delta\theta = 100\,\text{K}$. This figure shows that by doubling the value of the applied stress, the global transverse displacement gradient $\overline{F}_{22} - 1$ of the

Figure 10. Comparison between the creep behaviors of the thermoviscoelastic steel/rubber-like composite which is subjected to uniaxial transverse stress loadings $\overline{T}_{22} = 1\,\text{MPa}$ and $2\,\text{MPa}$ at elevated temperature $\Delta\theta = 100\,\text{K}$. As functions of time, the plots show the global transverse deformation gradient (a), damage evolution (b), and global transverse deformation gradient in the absence of damage effects.

Figure 11. Comparison between the creep behaviors of thermoviscoelastic (TVE) and thermoelastic (TE) steel/rubber-like composites where in the latter the viscous effects in the matrix phase have been neglected. Both composites are subjected to uniaxial transverse stress loading $\bar{T}_{22} = 1\,\text{MPa}$ at elevated temperature $\Delta\theta = 100\,\text{K}$. As functions of time, the plots show the global transverse deformation gradient (a), damage evolution (b), and global transverse deformation gradient in the absence of damage effects (c).

composite increases due to the nonlinearity by about 1.25 times only. On the other hand, the amount of saturation value of damage increase is just about 1.03 times. Next, Figure 11 shows the viscous effects in the matrix phase of the composite. This figure compares the creep response to uniaxial transverse stress loading $\bar{T}_{22} = 1\,\text{MPa}$ of the thermoviscoelastic (TVE) steel/rubber-like composite at elevated temperature $\Delta\theta = 100\,\text{K}$ with the corresponding thermoelastic (TE) one in which the viscoelasticity of the matrix is neglected and, therefore, there is no creep effect. The graphs show that the viscous effects are significant. It increases the saturated transverse macroscopic displacement gradient of the composite and the saturation value of damage by about 1.8 and 1.14 times, respectively. Finally, in Figure 12 the
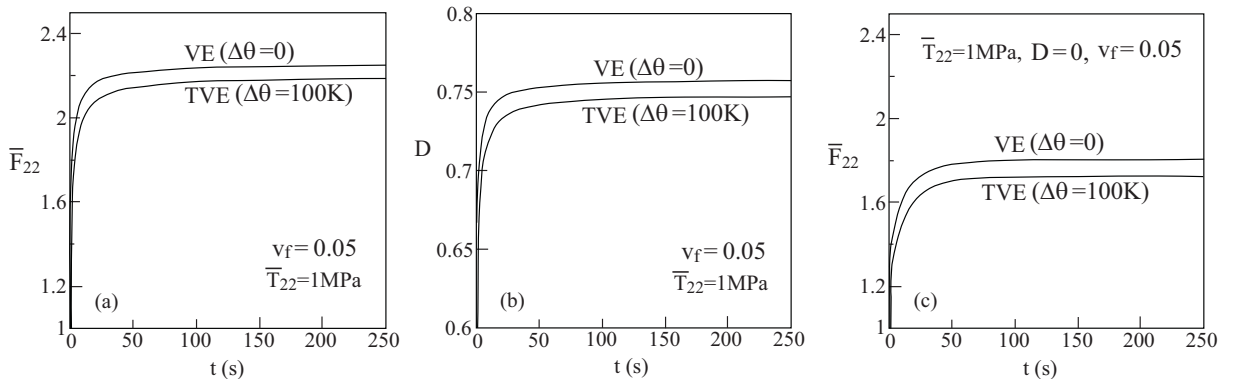
Figure 12. Comparison between the creep behaviors of thermoviscoelastic (TVE) at elevated temperature $\Delta\theta = 100\,\text{K}$, and viscoelastic (VE) at the reference temperature ($\Delta\theta = 0$) of steel/rubber-like composites. Both are subjected to uniaxial transverse stress loading $\bar{T}_{22} = 1\,\text{MPa}$. Global transverse deformation gradient (a), damage evolution (b), and global transverse deformation gradient in the absence of damage effects (c).

Figure 13. Relaxation behavior of the thermoviscoelastic steel/rubber-like composite which is subjected to a uniaxial transverse stress loading such that $\bar{F}_{22} = 2$ at elevated temperature $\Delta\theta = 100\,\text{K}$. Also shown is the corresponding relaxation behavior of the homogeneous (H) unreinforced thermoviscoelastic matrix. Global transverse stress (a), damage evolution (b), and global transverse stress in the absence of damage effects (c).

effect of elevated temperature on the creep of the thermoviscoelastic (TVE) steel/rubber-like composite loaded by a transverse stress $\bar{T}_{22} = 1\,\text{MPa}$ at $\Delta\theta = 100\,\text{K}$ is compared with that of a viscoelastic (VE) composite that is subjected to the same loading but by keeping the composite at the reference temperature $\theta = \theta_0$. This figure shows that the effect of elevated temperature on the creep of the composite and the evolving damage in its matrix are not appreciable.

Corresponding to these four figures that describe the creep behavior of the composite and its matrix in various circumstances, the following four figures show the relaxation behavior of the composite and its matrix. Figure 13(a) and (c) display the relaxation at elevated temperature $\Delta\theta = 100\,\text{K}$ of the thermoviscoelastic steel/rubber-like composite and its homogeneous (H) unreinforced matrix when they are subjected to a transverse deformation gradient of $\bar{F}_{22} = 2$ under uniaxial stress loading conditions (all components of $\bar{T}$ are equal to zero except $\bar{T}_{22}$) in the presence and absence of damage effects in the matrix. (Note that the scale of the graph when $D = 0$ is three times the scale of the damaged case.) Figure 13(b) shows is the damage evolution with time. The effect of the steel fibers is well exhibited in this figure.

The effect of doubling the applied transverse displacement gradient $\bar{F}_{22} - 1$ on the relaxation behavior of the composite is displayed in Figure 14 in the presence and absence of damage (it should be noted that the scale of the plot in the latter case is four time the scale of the former one). Doubling the applied deformation gradient generates a stress at saturation of about 5.3 and 7 times in the damaged and undamaged case, respectively. The damage increases by about 1.1 times, however.

The viscous effects in the rubber-like matrix on the relaxation behavior of the composite are shown in Figure 15 at elevated temperature $\Delta\theta = 100\,\text{K}$ in the presence and absence of damage. The relaxation stresses of the thermoviscoelastic (TVE) composite are observed to be much lower than the thermoelastic (TE) case in which no relaxation effects exist. In both cases, however, the viscoelasticity has a very small effect on the damage evolution in the matrix.

The final illustration is given in Figure 16, where a comparison between the relaxation behaviors of the thermoviscoelastic (TVE) composite at elevated temperature $\Delta\theta = 100\,\text{K}$ and a viscoelastic (VE)
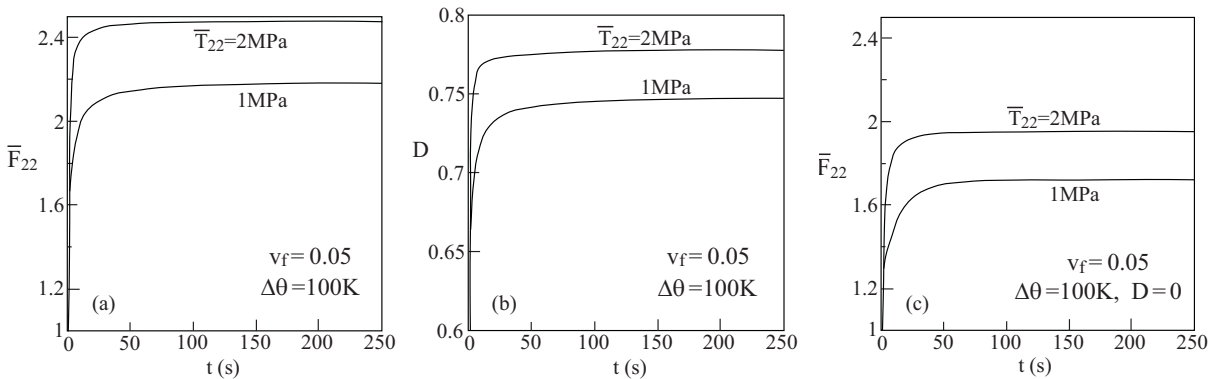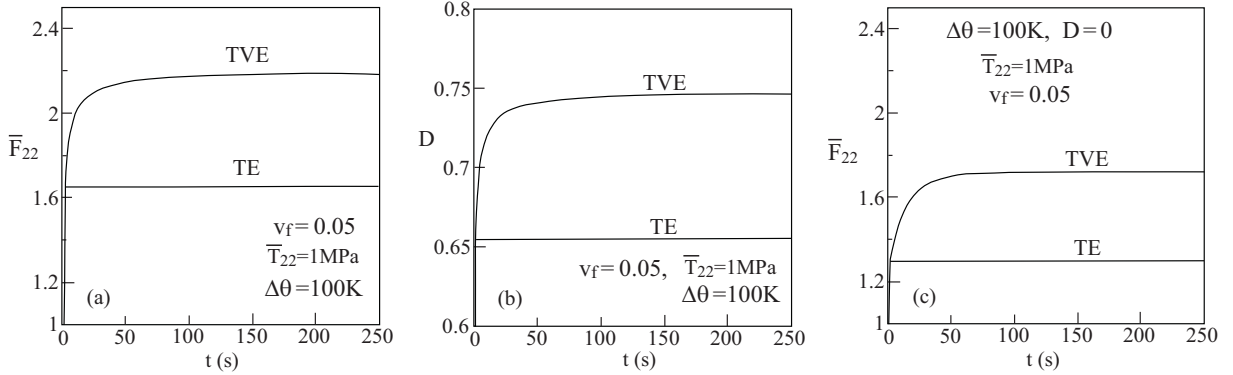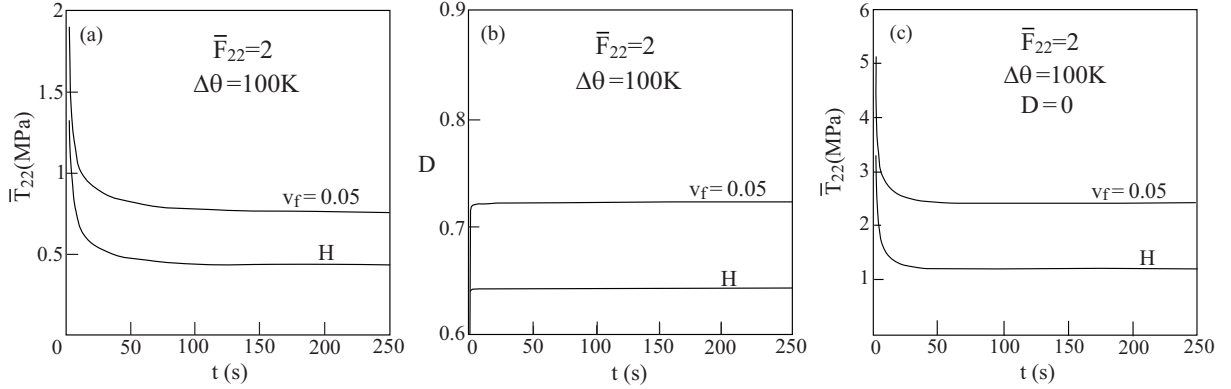
Figure 14. Comparison between the relaxation behaviors of the thermoviscoelastic steel/rubber-like composite which is subjected to uniaxial transverse stress loadings such that $\bar{F}_{22} = 2$ and 3 at elevated temperature $\Delta\theta = 100$ K. Global transverse stress (a), damage evolution (b), and global transverse stress in the absence of damage effects (c).
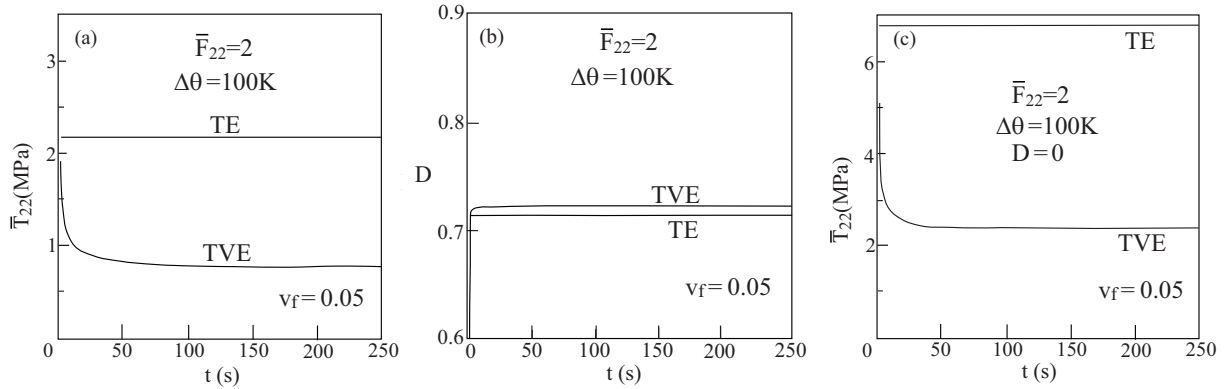
Figure 15. Comparison between the relaxation behaviors of thermoviscoelastic (TVE) and thermoelastic (TE) steel/rubber-like composites where in the latter the viscous effects of in the matrix phase have been neglected. Both composites are subjected to uniaxial transverse stress loading such that $\bar{F}_{22} = 2$ at elevated temperature $\Delta\theta = 100$ K. Global transverse stress (a), damage evolution (b), and global transverse stress in the absence of damage effects (c).

composite which is kept at the reference temperature $\theta = \theta_0$. As in the creep case, the effect of elevated temperature appears to be moderate.

Conclusions

The finite strain HFGMC micromechanical model which is capable of predicting the global behavior of thermoviscoelastic rubber-like matrix composites that are subjected to arbitrarily large thermomechanical loading has been presented. The rubber-like matrix is modeled by finite thermoviscoelasticity, which in contrast to finite linear thermoviscoelasticity where the deformations are large but the deviations from
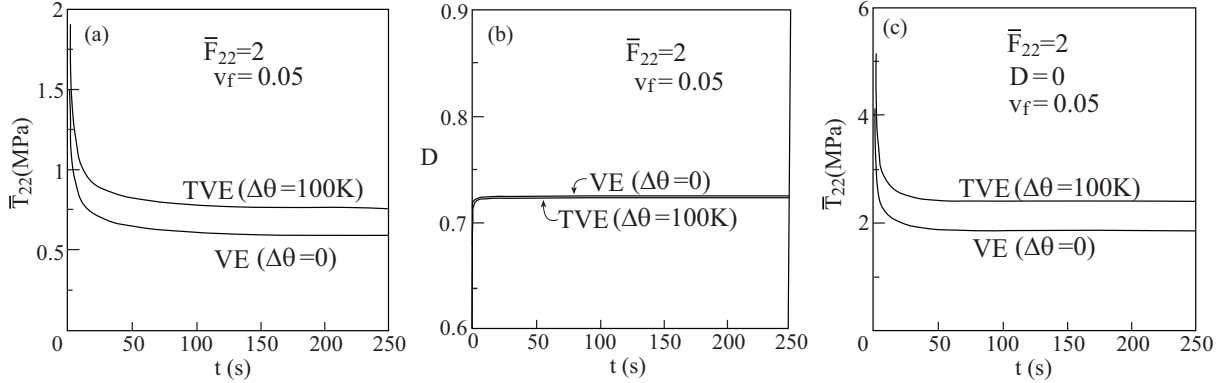
Figure 16. Comparison between the relaxation behaviors of thermoviscoelastic (TVE) at elevated temperature $\Delta\theta = 100\,\mathrm{K}$, and viscoelastic (VE) at t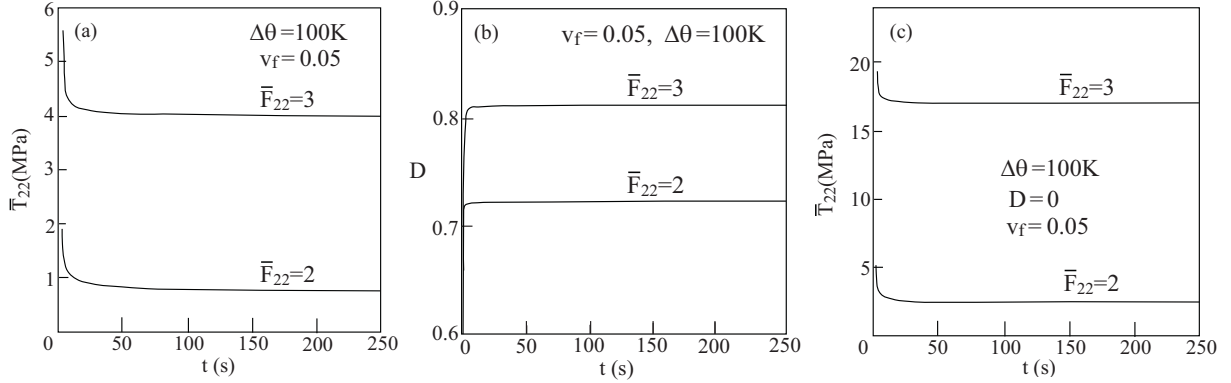he reference temperature ($\Delta\theta = 0$) of steel/rubber-like composites. Both composites are subjected to uniaxial transverse stress loading such that $\bar{F}_{22} = 2$. Global transverse stress (a), damage evolution (b), and global transverse stress in the absence of damage effects (c).

equilibrium are small (i.e., $\boldsymbol{B}^e \approx \boldsymbol{I}$, implying that the dependence on the strain is nonlinear but the dependence on the strain rate is linear), permits large deviation from equilibrium. In addition, the effect of evolving damage in the finite thermoviscoelastic matrix is incorporated. The finite strain HFGMC analysis establishes the rate form of the macroscopic constitutive equations that govern the composite's global response. The results exhibit the response of the composite and its unreinforced thermoviscoelastic matrix under various circumstances including their creep and relaxation behaviors.

The present derivation is confined to one-way thermomechanical coupling according to which the mechanical effects do not affect the temperature. It is possible, however, to generalize the micromechanical analysis by including a full (two-way) thermomechanical coupling.

The established global finite strain constitutive equations of the unidirectional composite can be employed to investigate the behavior of thermoviscoelastic laminates. They can be also employed to investigate the response of thermoviscoelastic structures such as laminated plates and shells undergoing large deformations. This can be performed by coupling the present micromechanical model to a finite element software such that the nonlinear composite structure response at each integration point is governed by the established macroscopic constitutive equations at each increment. This multiscale approach has been recently implemented by [Kim 2009] who coupled the hyperelastic HFGMC model to the finite element ABAQUS software in order to investigate the behavior of various types of tissue materials including the human arterial wall layers and porcine aortic valves leaflets.

References

[Aboudi 2004] J. Aboudi, "Micromechanics-based thermoviscoelastic constitutive equations for rubber-like matrix composites at finite strains", *Int. J. Solids Struct.* **41** (2004), 5611–5629.

[Aboudi 2008] J. Aboudi, "Finite strain micromechanical modeling of multiphase composites", *Int. J. Multiscale Comput. Engrg.* **6** (2008), 411–434.

[Aboudi 2009] J. Aboudi, "Finite strain micromechanical analysis of rubber-like matrix composites incorporating the Mullins damage effect", *Int. J. Damage Mech* **18** (2009), 5–29.

[Aboudi 2010] J. Aboudi, "Micromechanical modeling of viscoelastic behavior of polymer matrix composites undergoing large deformations", in *Creep and fatigue in polymer matrix composites*, edited by R. M. Guedes, Woodhead Publishing, Cambridge, UK, 2010.

[Aboudi and Pindera 2004] J. Aboudi and M.-J. Pindera, "High-fidelity micromechanical modeling of continuously reinforced elastic multiphase materials undergoing finite deformations", *Math. Mech. Solids* **9**:6 (2004), 599–628.

[Christensen 1982] R. M. Christensen, *Theory of Voscoelasticy*, Academic Press, New York, 1982.

[Cuitino and Ortiz 1992] A. Cuitino and M. Ortiz, "A material-independent method for extending stress update algorithms from small-strain plasticity with multiplicative kinematics", *Engrg. Comp.* **9** (1992), 437–451.

[Eterovic and Bathe 1990] A. L. Eterovic and K.-J. Bathe, "A hyperelastic based large strain elasto-plastic constitutive formulation with combined isotropic-kinematic hardening using the logarithmic stress and strain measures", *Int. J. Num. Meth. Engrg.* **30** (1990), 1099–1114.

[Govindjee and Reese 1997] S. Govindjee and S. Reese, "A presentation and comparison of two large deformation viscoelasticity models", *J. Engrg. Mater. Tech.* **119** (1997), 251–255.

[Holzapfel 2000] G. A. Holzapfel, *Nonlinear solid mechanics*, John Wiley & Sons Ltd., Chichester, 2000. A continuum approach for engineering.

[Holzapfel and Simo 1996] G. A. Holzapfel and J. C. Simo, "A new viscoelastic constitutive model for continuous media at finite thermomechanical changes", *Int. J. Solids Struct.* **33** (1996), 3019–3034.

[Kim 2009] H. S. Kim, *Nonlinear multiscale anisotropic material and structural models for prosthetic and native aortic heart valves*, Ph.D. dissertation, Georgia Institute of Technology, Atlanta, 2009, available at http://smartech.gatech.edu/handle/1853/29671.

[Lin and Schomburg 2003] R. C. Lin and U. Schomburg, "A finite elastic-viscoelastic-elastoplastic material law with damage: theoretical and numerical aspects", *Comput. Meth. Appl. Mech. Engrg.* **192** (2003), 1591–1627.

[Lion 1996] A. Lion, "A physically based method to represent the thermo-mechanical behaviour of elastomers", *Acta Mech.* **123** (1996), 1–25.

[Lockett 1972] F. J. Lockett, *Nonlinear viscoelastic solids*, Academic Press, New York, 1972.

[Miehe and Keck 2000] C. Miehe and J. Keck, "Superimposed finite elastic-viscoelastic stress response with damage in filled rubbery polymers: Experiments, modelling and algorithmic implementation", *J. Mech. Phys. Solids* **48** (2000), 323–365.

[Ogden 1984] R. W. Ogden, *Non-linear elastic deformations*, Ellis Horwood, Chichester, 1984.

[Ogden 1992] R. W. Ogden, "On the thermoelastic modeling of rubberlike solids", *J. Thermal Stresses* **15**:4 (1992), 533–557.

[Reese and Govindjee 1998a] S. Reese and S. Govindjee, "Theoretical and numerical aspects in the thermo-viscoelastic material behaviour of rubber-like polymers", *Mech. Time-Dependent Mater.* **1** (1998), 357–396.

[Reese and Govindjee 1998b] S. Reese and S. Govindjee, "A theory of finite viscoelasticity and numerical aspects", *Int. J. Solids Struct.* **35** (1998), 3455–3482.

[Simo 1987] J. C. Simo, "On a fully three-dimensional finite-strain viscoelastic damage model: formulation and computational aspects", *Comp. Meth. Appl. Mech. Engrg.* **60** (1987), 153–173.

[Simo 1992] J. C. Simo, "Algorithms for static and dynamic multiplicative plasticity that preserve the classical return mapping schemes of the infinitesimal theory", *Comput. Methods Appl. Mech. Engrg.* **99**:1 (1992), 61–112.

[Weber and Anand 1990] G. Weber and L. Anand, "Finite deformation constitutive equations and a time integration procedure for isotropic, hyperelastic-viscoplastic solids", *Comp. Meth. Appl. Mech. Engrg.* **79** (1990), 173–202.

JACOB ABOUDI: aboudi@eng.tau.ac.il
Faculty of Engineering, Tel Aviv University, 69978 Ramat-Aviv, Israel
http://www.eng.tau.ac.il/~aboudi/

GENERALIZED THERMOELASTIC WAVES IN CYLINDERS
DUE TO LOCALIZED HEATING

HAO BAI, RAVI CHITIKIREDDY, ARVIND H. SHAH AND SUBHENDU K. DATTA

This paper presents a theoretical study of propagation of thermoelastic waves generated by concentrated heating of the outer surface of circular cylindrical shells. The generalized thermoelastic theory proposed by Lord and Shulman is used to model the response of a circular cylindrical shell to a pulsed laser focused on the surface of the cylinder. Guided wave modes in the cylinder are obtained by a semianalytical finite element method. Dynamic response is constructed numerically by superposition of guided wave modes. In this method, the cylinder is discretized in the radial direction into several coaxial circular cylinders (subcylinders) and the radial dependence of the displacement and temperature in each subcylinder is approximated by quadratic interpolation polynomials. Numerical results for the variation through the thickness of various physical quantities of interest at a location away from the source are presented for a silicon nitride (Si_3N_4) tube for illustration purposes. The frequency dependence of the response quantities is discussed and attention is focused on convergence and accuracy of the computed results.

1. Introduction

Laser-based ultrasonic techniques have been used in several recent studies to generate elastic waves in solids. These techniques provide a number of advantages over conventional ultrasonic methods, such as higher spatial resolution, noncontact generation and detection of waves, and the ability to operate on curved and rough surfaces [Scruby and Drain 1990]. The use of generalized thermoelasticity theories to analyze thermoelastic waves generated by a pulsed laser beam has received some attention in recent years. The classical theory of heat conduction in solids treats the flux of heat as proportional to the gradient of temperature in the media. Thus, the heat conduction equation is a parabolic partial differential equation, which predicts an infinite thermal wave speed. This assumption of infinite speed of heat propagation is contrary to physical reality. To rectify this paradox, several generalizations to the classical heat conduction equation and the thermoelastic wave equations have been proposed. These generalizations take into account the finite wave speed of the thermal pulse travelling through the body.

Lord and Shulman [1967] presented a generalized theory of thermoelasticity for an isotropic body that was a modification of the Fourier law of heat conduction by the inclusion of a relaxation time. This then predicts a finite thermal wave speed. Other generalizations were proposed by Green and Lindsay [1972], who developed a temperature-rate dependent theory that included two relaxation times, and by Green and Naghdi [1993], who introduced a thermoelasticity theory without energy dissipation. See [Ignaczak and Ostoja-Starzewski 2010] for a theoretical development of thermoelasticity with finite wave speeds.

Guided surface (Rayleigh) waves and Rayleigh–Lamb waves in plates have been studied by several investigators using generalized thermoelasticity theories. Rayleigh–Lamb waves in plates using Lord–Shulman theory have been discussed in [Datta and Shah 2009]. References to many other works can be found in [Sharma et al. 2000; Verma and Hasebe 2001; Verma 2002; Al-Qahtani and Datta 2004; Al-Qahtani et al. 2005]. Compared to many studies that have dealt with plates, there is a very limited amount of reported work on waves in cylinders using generalized thermoelastic theories.

Dispersion of longitudinal thermoelastic wave propagation in a circular isotropic cylinder was studied by Erbay and Şuhubi [1986], who considered the cylinder surface to be stress-free and at a constant temperature. Elnagar and Abd-Alla [1987] studied the influence of the initial stress on Rayleigh wave propagation in a generalized thermoelastic cylinder. Three-dimensional vibration of a homogeneous transversely isotropic thermoelastic cylindrical panel was investigated in [Sharma 2001; Sharma and Sharma 2002]. Recently, wave propagation in a thermoelastic cylinder of an arbitrary cross section was reported by Ponnusamy [2007], who used a collocation method to study the dispersive waves.

Circumferential isothermal elastic waves in an isotropic cylinder generated by a laser pulse and their scattering by a surface defect was studied experimentally in [Clorennec and Royer 2003]. Pan et al. [2004; 2006] investigated theoretically and experimentally the isothermal acoustic waves generated by a laser point pulse in an isotropic and a transversely isotropic cylinder. Three-dimensional Fourier transforms were used to find the dynamic displacements at the cylinder surface.

In this work, we present an analysis of generalized thermoelastic waves in cylinders due to a pulsed laser beam focused on the surface of the cylinders. The generalized heat conduction theory of [Lord and Shulman 1967] has been adopted here. A semianalytical finite element (SAFE) method was employed earlier in [Chitikireddy et al. 2010] to obtain the dispersion relations for guided thermoelastic waves in free cylinders. These modes are used here to represent the time-harmonic solutions for the field quantities due to the laser excitation. Based on this solution, the steady-state Green's function for the thermoelastic cylinder can also be constructed. Zhuang et al. [1999] used the SAFE method to study the steady-state Green's function for isothermal composite cylinders. The response in the time domain can be obtained by applying an inverse Fourier transform. Here, the distributions of the stresses, displacements, and heat flux through the thickness of the cylinder are presented at certain chosen frequencies to show the convergence of the results.

2. Formulation

2.1. *Governing equations.* Consider an infinite thermoelastic cylinder of inner radius r_i and outer radius r_o in the cylindrical coordinate system (r, θ, z) with the origin at the centre of the cross section of the cylinder as shown in Figure 1. The generalized Lord–Shulman governing equations of thermoelasticity, in the presence of a body force and heat source, are given by [Al-Qahtani et al. 2005]

$$\sigma_{ij,j} + f_i = \rho \ddot{u}_i, \quad T_0 \rho \dot{\eta} + Q = -q_{i,i}, \quad \sigma_{ij} = C_{ijkl}\varepsilon_{kl} - \beta_{ij}T,$$
$$\rho \dot{\eta} = \beta_{ij}\dot{\varepsilon}_{ij} + \frac{\rho c_E}{T_0}\dot{T}, \quad q_i + \tau_0 \dot{q}_i = -k_{ij}T_{,j}. \tag{1}$$

The physical quantities and material constants appearing in these equations are: σ_{ij}, the components of the stress tensor; u_i, the components of the displacement; ε_{ij}, the components of the strain tensor; C_{ijkl}, the elastic constants; q_i, the components of the heat flux vector; ρ, the mass density; T_0, the

reference temperature; η, the entropy density; T, the temperature change; c_E, the specific heat at constant deformation; τ_0, the thermal relaxation time; β_{ij}, the thermal coefficients; k_{ij}, the coefficients of thermal conductivity; f_i, the body force per unit volume; and Q, the heat source. In the above equations, a superposed dot indicates the derivative with respect to time.

We define nondimensional quantities

$$r_i^* = \frac{r_i}{H}, \quad u_i^* = \frac{u_i}{\delta H}, \quad T^* = \frac{T}{\bar{T}}, \quad t^* = \frac{\bar{v}}{H}t, \quad \sigma_{ij}^* = \frac{\sigma_{ij}}{\delta \bar{c}}, \quad q_i^* = \frac{q_i}{\bar{q}}, \quad \varepsilon_{ij}^* = \frac{\varepsilon_{ij}}{\delta},$$

$$\rho^* = \frac{\rho}{\bar{\rho}}, \quad c_E^* = \frac{c_E}{\bar{c}_E}, \quad \beta_{ij}^* = \frac{\beta_{ij}}{\bar{\beta}}, \quad k_{ij}^* = \frac{k_{ij}}{\bar{k}}, \quad c_{ijkl}^* = \frac{c_{ijkl}}{\bar{c}}, \quad \tau_0^* = \frac{\bar{v}}{H}\tau_0, \quad T_0^* = \frac{T_0}{\bar{T}},$$

(2)

where

$$\bar{v} = \sqrt{\frac{\bar{c}}{\bar{\rho}}}, \quad \delta = \sqrt{\frac{\bar{k}\bar{T}}{H\bar{v}\bar{c}}}, \quad \bar{q} = \frac{\bar{k}\bar{T}}{H}, \quad \bar{\beta} = \frac{\delta \bar{c}}{\bar{T}}, \quad \bar{c}_E = \frac{\bar{k}}{\bar{\rho}\bar{v}H}.$$

(3)

In the sequel, two more nondimensional quantities are defined: the wave number $k^* = kH$ and frequency $\omega^* = \omega H/\bar{v}$. Here $\bar{\rho}, \bar{k}, \bar{c}, H$, and $\bar{T}$ are the basic normalization quantities, and the rest can be derived from them as shown above. Note that $\bar{\rho}, \bar{k}, \bar{c}$, and $\bar{T}$ can be taken suitably depending upon the material properties of the cylinder. This nondimensionalization scheme yields all dimensionless equations in the same form as their dimensional counterparts. Therefore, this normalization could be used to solve multilayer structures.

Using (1), the governing coupled thermoelastic equations can be written as

$$c_{ijkl}u_{k,lj} - \beta_{ij}T_{,j} + f_i = \rho \ddot{u}_i, \qquad \left(1 + \tau_0 \frac{\partial}{\partial t}\right)[T_0\beta_{ij}\dot{\varepsilon}_{ij} + \rho c_E \dot{T} + Q] = k_{ij}T_{,ij}.$$

(4)

Writing these equations in terms of the nondimensional quantities defined in (2) and introducing

$$f_i^* = \frac{H}{\delta \bar{c}}f_i, \quad Q^* = \frac{H^2}{\bar{k}\bar{T}}\left(1 + \tau_0^* \frac{\partial}{\partial t^*}\right)Q,$$
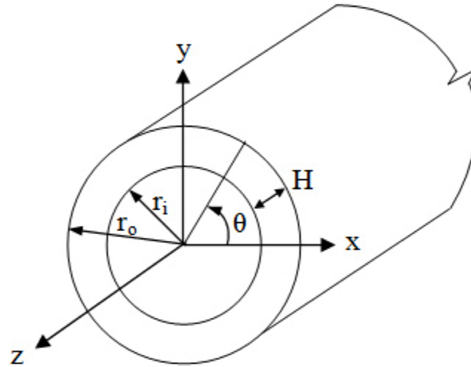
Figure 1. Geometry of the cylinder.

we obtain

$$c_{ijkl}^* u_{k,lj}^* - \beta_{ij}^* T_{,j}^* + f_i^* = \rho^* \ddot{u}_i^*,$$

$$\beta_{ij}^* T_0^* \left(1 + \tau_0^* \frac{\partial}{\partial t^*}\right) \dot{u}_{i,j}^* - k_{ij}^* T_{,ij}^* + \rho^* c_E^* \left(1 + \tau_0^* \frac{\partial}{\partial t^*}\right) \dot{T}^* = -Q^*. \tag{5}$$

These two equations have to be solved by numerical techniques for a generally anisotropic medium. If the material of the cylinder is transversely isotropic with the symmetry axis coinciding with the z-axis, the equations can be solved analytically in terms of Bessel functions. This will be discussed in a future publication. In the following, we will use the SAFE method to solve the problem. For convenience, the superscript * will be dropped.

2.2. Semianalytical finite element formulation. In this method, the radial dependence of the displacement $\boldsymbol{u}$ and the temperature T are approximated by one-dimensional isoparametric finite elements. The total thickness of the cylinder H is composed of cylindrical layers and each layer can have distinct thermoelastic properties and thickness. By using the SAFE method, the thickness of the cylinder is discretized into N laminas. Quadratic interpolation polynomials are used to approximate the displacement and temperature field over each lamina in the radial direction.

The displacement and temperature of the k-th lamina are expressed as

$$\boldsymbol{u}(r, \theta, z, t) = \boldsymbol{N}_1(\xi) \boldsymbol{u}^e(\theta, z, t),$$

$$T(r, \theta, z, t) = \boldsymbol{N}_2(\xi) \boldsymbol{T}^e(\theta, z, t), \tag{6}$$

with

$$r = \boldsymbol{N}_2(\xi) \boldsymbol{R}^e. \tag{7}$$

Here

$$\boldsymbol{N}_1(\xi) = \begin{bmatrix} n_1 & 0 & 0 & n_2 & 0 & 0 & n_3 & 0 & 0 \\ 0 & n_1 & 0 & 0 & n_2 & 0 & 0 & n_3 & 0 \\ 0 & 0 & n_1 & 0 & 0 & n_2 & 0 & 0 & n_3 \end{bmatrix}, \quad \boldsymbol{N}_2(\xi) = \begin{bmatrix} n_1 & n_2 & n_3 \end{bmatrix}, \tag{8}$$

$$\boldsymbol{u}^e = (u_{r1}, u_{\theta 1}, u_{z1}, u_{r2}, u_{\theta 2}, u_{z2}, u_{r3}, u_{\theta 3}, u_{z3})^T, \quad \boldsymbol{T}^e = (T_1, T_2, T_3)^T, \quad \boldsymbol{R}^e = (R_b, R_m, R_f)^T. \tag{9}$$

In (8), the shape functions are

$$n_1 = \tfrac{1}{2}\xi(\xi - 1), \quad n_2 = 1 - \xi^2, \quad n_3 = \tfrac{1}{2}\xi(\xi + 1), \quad -1 \le \xi \le 1. \tag{10}$$

In (9) the nodal displacements u_{rj}, $u_{\theta j}$, and u_{zj} and temperature T_j, where $j = 1, 2, 3$, are taken at the inner surface ($r = R_b$), middle layer ($r = R_m = (R_b + R_f)/2$), and outer surface ($r = R_f$) of the k-th lamina.

The strain tensor and temperature gradient are expressed as

$$\boldsymbol{\varepsilon} = \boldsymbol{D}_1 \boldsymbol{u}^e + \boldsymbol{D}_2 \boldsymbol{u}_{,\theta}^e + \boldsymbol{D}_3 \boldsymbol{u}_{,z}^e, \quad \boldsymbol{T}' = \boldsymbol{B}_1 \boldsymbol{T}^e + \boldsymbol{B}_2 \boldsymbol{T}_{,\theta}^e + \boldsymbol{B}_3 \boldsymbol{T}_{,z}^e. \tag{11}$$

The matrices $\boldsymbol{B}_1$, $\boldsymbol{B}_2$, $\boldsymbol{B}_3$, $\boldsymbol{D}_1$, $\boldsymbol{D}_2$, and $\boldsymbol{D}_3$ are defined in the Appendix. The stress vector is given by

$$\boldsymbol{\sigma} = \boldsymbol{C}(\boldsymbol{D}_1 \boldsymbol{u}^e + \boldsymbol{D}_2 \boldsymbol{u}_{,\theta}^e + \boldsymbol{D}_3 \boldsymbol{u}_{,z}^e) - \boldsymbol{\beta} \boldsymbol{N}_2 \boldsymbol{T}^e. \tag{12}$$

The variational principle of thermoelasticity [Al-Qahtani et al. 2005] is

$$\int_{t_0}^{t_1}\int_V \left(\delta\boldsymbol{\varepsilon}^T\boldsymbol{\sigma} - \delta\boldsymbol{T'}^T\boldsymbol{K}\boldsymbol{T'} - \delta\boldsymbol{T'}^T(\boldsymbol{q}+\tau_0\dot{\boldsymbol{q}})\right)dV\,dt = \int_{t_0}^{t_1}\int_V (\boldsymbol{f}-\delta\boldsymbol{u}^T\rho\ddot{\boldsymbol{u}})\,dV\,dt. \qquad (13)$$

The first term on the left-hand side is

$$\int_{t_0}^{t_1}\int_V \delta\boldsymbol{\varepsilon}^T\boldsymbol{\sigma}\,dV\,dt = \int_{t_0}^{t_1}\int_V (\boldsymbol{D}_1\delta\boldsymbol{u}^e + \boldsymbol{D}_2\delta\boldsymbol{u}^e_{,\theta} + \boldsymbol{D}_3\delta\boldsymbol{u}^e_{,z})^T[\boldsymbol{C}(\boldsymbol{D}_1\boldsymbol{u}^e + \boldsymbol{D}_2\boldsymbol{u}^e_{,\theta} + \boldsymbol{D}_3\boldsymbol{u}^e_{,z}) - \boldsymbol{\beta}\boldsymbol{N}_2\boldsymbol{T}^e]\,dV\,dt$$

$$= \int_{t_0}^{t_1}\int_z\int_\theta \delta u^{eT}\begin{bmatrix} \boldsymbol{K}_{11}\boldsymbol{u}^e + (\boldsymbol{K}_{12}-\boldsymbol{K}_{21})\boldsymbol{u}^e_{,\theta} + (\boldsymbol{K}_{13}-\boldsymbol{K}_{31})\boldsymbol{u}^e_{,z} - \\ \boldsymbol{K}_{22}\boldsymbol{u}^e_{,\theta\theta} - (\boldsymbol{K}_{23}+\boldsymbol{K}_{32})\boldsymbol{u}^e_{,\theta z} - \boldsymbol{K}_{33}\boldsymbol{u}^e_{,zz} - \\ \boldsymbol{K}^e_{01}\boldsymbol{T}^e + \boldsymbol{K}^e_{02}\boldsymbol{T}^e_{,\theta} + \boldsymbol{K}^e_{03}\boldsymbol{T}^e_{,z} \end{bmatrix}d\theta\,dz\,dt. \quad (14)$$

The second term in (13) is

$$\int_{t_0}^{t_1}\int_V \delta\boldsymbol{T'}^T\boldsymbol{K}\boldsymbol{T'}\,dV\,dt = \int_{t_0}^{t_1}\int_V (\delta\boldsymbol{T}^{eT}\boldsymbol{B}_1^T + \delta\boldsymbol{T}^{eT}_{,\theta}\boldsymbol{B}_2^T + \delta\boldsymbol{T}^{eT}_{,z}\boldsymbol{B}_3^T)\boldsymbol{K}(\boldsymbol{B}_1\boldsymbol{T}^e + \boldsymbol{B}_2\boldsymbol{T}^e_{,\theta} + \boldsymbol{B}_3\boldsymbol{T}^e_{,z})\,dV\,dt$$

$$= \int_{t_0}^{t_1}\int_z\int_\theta \delta\boldsymbol{T}^{eT}(\boldsymbol{g}_{11}\boldsymbol{T}^e - \boldsymbol{g}_{22}\boldsymbol{T}^e_{,\theta\theta} - \boldsymbol{g}_{33}\boldsymbol{T}^e_{,zz})\,d\theta\,dz\,dt. \qquad (15)$$

The third term in (13) is

$$\int_{t_0}^{t_1}\int_V \delta\boldsymbol{T'}^T(\boldsymbol{q}+\tau_0\dot{\boldsymbol{q}})\,dV\,dt = -\int_{t_0}^{t_1}\int_V \delta\boldsymbol{T}^T(\nabla\cdot\boldsymbol{q}+\tau_0\nabla\cdot\dot{\boldsymbol{q}})\,dV\,dt$$

$$= \int_{t_0}^{t_1}\int_V \delta\boldsymbol{T}^T(T_0\rho\dot{\eta} + Q + \tau_0 T_0\rho\ddot{\eta} + \tau_0\dot{Q})\,dV\,dt$$

$$= \int_{t_0}^{t_1}\int_V \delta\boldsymbol{T}^T\left(T_0\boldsymbol{\beta}^T\dot{\boldsymbol{\varepsilon}} + \rho C_E\dot{T} + Q + \tau_0(T_0\boldsymbol{\beta}^T\ddot{\boldsymbol{\varepsilon}} + \rho C_E\ddot{T} + \dot{Q})\right)dV\,dt \qquad (16)$$

$$= \int_{t_0}^{t_1}\int_z\int_\theta \delta\boldsymbol{T}^{eT}\begin{bmatrix} (\boldsymbol{f}_1\dot{\boldsymbol{u}}^e + \boldsymbol{f}_2\dot{\boldsymbol{u}}^e_{,\theta} + \boldsymbol{f}_3\dot{\boldsymbol{u}}^e_{,z} + \boldsymbol{m}_0\dot{\boldsymbol{T}}^e) + \\ \tau_0(\boldsymbol{f}_1\ddot{\boldsymbol{u}}^e + \boldsymbol{f}_2\ddot{\boldsymbol{u}}^e_{,\theta} + \boldsymbol{f}_3\ddot{\boldsymbol{u}}^e_{,z} + \boldsymbol{m}_0\ddot{\boldsymbol{T}}^e) + \boldsymbol{Q}^e \end{bmatrix}d\theta\,dz\,dt.$$

The right-hand side of (13) has the form

$$\int_{t_0}^{t_1}\int_V \delta\boldsymbol{u}^T(\boldsymbol{f}-\rho\ddot{\boldsymbol{u}})\,dV\,dt = \int_{t_0}^{t_1}\int_z\int_\theta \delta\boldsymbol{u}^{eT}(\boldsymbol{f}^e - \boldsymbol{M}\ddot{\boldsymbol{u}}^e)\,d\theta\,dz\,dt. \qquad (17)$$

The element matrices appearing in equations (14)–(17) are defined in the Appendix. Equating the coefficients of $\delta\boldsymbol{u}^e$ in (13) to zero gives the equation

$$\boldsymbol{M}\ddot{\boldsymbol{u}}^e + \boldsymbol{K}_{11}\boldsymbol{u}^e - \boldsymbol{K}_{01}\boldsymbol{T}^e + (\boldsymbol{K}_{12}-\boldsymbol{K}_{21})\boldsymbol{u}^e_{,\theta} + \boldsymbol{K}_{02}\boldsymbol{T}^e_{,\theta}$$
$$+ (\boldsymbol{K}_{13}-\boldsymbol{K}_{31})\boldsymbol{u}^e_{,z} + \boldsymbol{K}_{03}\boldsymbol{T}^e_{,z} - \boldsymbol{K}_{22}\boldsymbol{u}^e_{,\theta\theta} - (\boldsymbol{K}_{23}+\boldsymbol{K}_{32})\boldsymbol{u}^e_{,\theta z} - \boldsymbol{K}_{33}\boldsymbol{u}^e_{,zz} = \boldsymbol{f}^e. \quad (18)$$

Similarly, equating the coefficients of $\delta\boldsymbol{T}^e$ in (13) yields

$$\tau_0\boldsymbol{f}_1\ddot{\boldsymbol{u}}^e + \tau_0\boldsymbol{m}_0\ddot{\boldsymbol{T}}^e + \tau_0\boldsymbol{f}_2\ddot{\boldsymbol{u}}^e_{,\theta} + \tau_0\boldsymbol{f}_3\ddot{\boldsymbol{u}}^e_{,z} + \boldsymbol{m}_0\dot{\boldsymbol{T}}^e + \boldsymbol{f}_1\dot{\boldsymbol{u}}^e + \boldsymbol{f}_2\dot{\boldsymbol{u}}^e_{,\theta} + \boldsymbol{f}_3\dot{\boldsymbol{u}}^e_{,z} + \boldsymbol{g}_{11}\boldsymbol{T}^e - \boldsymbol{g}_{22}\boldsymbol{T}^e_{,\theta\theta} - \boldsymbol{g}_{33}\boldsymbol{T}^e_{,zz} = \boldsymbol{Q}^e.$$
$$(19)$$

Combining (18) and (19) and assembling the element matrices into global matrices leads to the governing equations of motion

$$H_1\ddot{V}+H_2\ddot{V}_{,\theta}+H_3\ddot{V}_{,z}+H_4\dot{V}+H_5\dot{V}_{,\theta}+H_6\dot{V}_{,z}+H_7V+H_8V_{,\theta}+H_9V_{,z}+H_{10}V_{,\theta\theta}+H_{11}V_{,\theta z}+H_{12}V_{,zz}$$
$$=F, \quad (20)$$

where H_i ($i = 1, 2, \ldots, 12$) are the global matrices, given in the Appendix, V is the global nodal displacement and temperature vector, and F is a load vector defined by

$$F = \begin{bmatrix} f^e \\ Q^e \end{bmatrix}. \quad (21)$$

The traction-free boundary conditions on surfaces of the cylinder require that the stresses at the inner and outer surfaces of the cylinder are zero:

$$\sigma_{rr} = \sigma_{r\theta} = \sigma_{rz} = 0 \quad \text{at} \quad r = r_{\mathrm{i}} \text{ and } r = r_{\mathrm{o}}. \quad (22)$$

The thermal boundary conditions are considered as

$$T_{,r} = 0 \quad \text{at} \quad r = r_{\mathrm{i}} \text{ and } r = r_{\mathrm{o}}. \quad (23)$$

These thermal boundary conditions imply that heat does not flow into or out of the system via the boundaries. Note that for a free cylinder without body forces and heat sources the right-hand side of (20) will be zero. Then, the homogeneous equation (20) leads to the eigenvalue problem for the determination of the dispersion relation between the frequency ω and the z-direction wave number k for a fixed circumferential integral wave number n. This was studied in [Chitikireddy et al. 2010].

3. Solution procedure for steady-state loading

In (20), the force vector F and the response V are assumed to be time harmonic with frequency ω. The θ-dependence of the load and response can be expressed in Fourier series as:

$$F(\theta, z, t) = e^{-i\omega t} F(\theta, z) = e^{-i\omega t} \sum_{n=-\infty}^{n=+\infty} F_n(z)e^{in\theta},$$
$$V(\theta, z, t) = e^{-i\omega t} V(\theta, z) = e^{-i\omega t} \sum_{n=-\infty}^{n=+\infty} V_n(z)e^{in\theta}. \quad (24)$$

Substitution of (24) into (20) yields a system of ordinary differential equations with Fourier coefficients V_n in terms of z. For each circumferential wave number (n), we obtain

$$H_{12}V_{n,zz} + (-\omega^2 H_3 - i\omega H_6 + H_9 + inH_{11})V_{n,z}$$
$$- [\omega^2(H_1 + inH_2) + i\omega(H_4 + inH_5) - (H_7 + inH_8 - n^2 H_{10})]V_n = F_n. \quad (25)$$

The following Fourier integral transform pairs are used to treat the z-dependence in (25):

$$\bar{V}_n(k_n) = \int_{-\infty}^{+\infty} V_n(z)e^{-ik_n z}\, dz, \quad V_n(z) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \bar{V}_n(k_n)e^{ik_n z}\, dk_n. \quad (26)$$

Application of Fourier transformation to (25) yields an algebraic equation in terms of the transform parameter (k_n):

$$-k_n^2 \boldsymbol{H}_{12} \bar{\boldsymbol{V}}_n + k_n \boldsymbol{H}_B \bar{\boldsymbol{V}}_n + \boldsymbol{H}_A \bar{\boldsymbol{V}}_n = \bar{\boldsymbol{F}}_n, \tag{27}$$

where

$$\boldsymbol{H}_B = -i\omega^2 \boldsymbol{H}_3 + \omega \boldsymbol{H}_6 + i \boldsymbol{H}_9 - n \boldsymbol{H}_{11},$$

$$\boldsymbol{H}_A = -\omega^2 (\boldsymbol{H}_1 + in\boldsymbol{H}_2) - i\omega(\boldsymbol{H}_4 + in\boldsymbol{H}_5) + (\boldsymbol{H}_7 + in\boldsymbol{H}_8 - n^2 \boldsymbol{H}_{10}).$$

Equation (27) is the governing equation for the n-th circumferential harmonic in the transformed domain.

The solution of (27) will be obtained in the form of an expansion in guided wave modes in the z-direction. For this purpose, we consider the homogeneous equation, which is a three-parameter algebraic eigensystem in ω, n, and k_n. If k_n serves as an eigenvalue for given values of ω and n, (27) gives a quadratic eigenvalue problem. Equation (27) can be converted into two first-order equations in the form

$$\begin{bmatrix} 0 & I \\ \boldsymbol{H}_A & \boldsymbol{H}_B \end{bmatrix} \begin{Bmatrix} \bar{\boldsymbol{V}}_n \\ k_n \bar{\boldsymbol{V}}_n \end{Bmatrix} - k_n \begin{bmatrix} I & 0 \\ 0 & \boldsymbol{H}_{12} \end{bmatrix} \begin{Bmatrix} \bar{\boldsymbol{V}}_n \\ k_n \bar{\boldsymbol{V}}_n \end{Bmatrix} = \begin{bmatrix} 0 \\ \bar{\boldsymbol{F}}_n \end{bmatrix}, \tag{28}$$

which we rewrite more compactly as

$$[\boldsymbol{A} - k_n \boldsymbol{B}]\boldsymbol{U}_n = \boldsymbol{P}_n. \tag{29}$$

If the displacement and temperature vector $\bar{\boldsymbol{V}}_n$ has dimension M, the dimension of $\boldsymbol{U}_n$ in (29) is $2M$. A nontrivial solution to the homogeneous form of (29) in terms of k_n yields $2M$ roots, denoted by k_{nm}. They represent axial wave numbers which can be real, purely imaginary, or complex. A real wave number represents a propagating wave and purely imaginary or complex wave numbers represent nonpropagating (evanescent) waves.

Once the wavenumbers and wave functions are found from (29), the response due to the n-th circumferential mode in the Fourier series representation of the applied loads can be obtained by modal summation. Associated with each eigenvalue there are right and left eigenvectors, $\boldsymbol{\Phi}_{nm}^R$ and $\boldsymbol{\Phi}_{nm}^L$, respectively, and they satisfy the equations

$$[\boldsymbol{A} - k_n \boldsymbol{B}]\boldsymbol{\Phi}_{nm}^R = 0, \quad [\boldsymbol{A}^T - k_n \boldsymbol{B}^T]\boldsymbol{\Phi}_{nm}^L = 0. \tag{30}$$

The right and left eigenvectors also satisfy the biorthogonality relations

$$\boldsymbol{\Phi}_{nm}^{L^T} \boldsymbol{B} \boldsymbol{\Phi}_{np}^R = \operatorname{diag}(B_{nm}), \quad \boldsymbol{\Phi}_{nm}^{L^T} \boldsymbol{A} \boldsymbol{\Phi}_{np}^R = \operatorname{diag}(k_{nm} B_{nm}), \tag{31}$$

where diag( ) denotes a diagonal matrix. The eigenvectors can also be partitioned into the following upper and lower halves (represented by subscript u and l, respectively):

$$\boldsymbol{\Phi}_{nm}^R = \begin{bmatrix} \boldsymbol{\Phi}_{nm\mathrm{u}}^R \\ \boldsymbol{\Phi}_{nm\mathrm{l}}^R \end{bmatrix}, \quad \boldsymbol{\Phi}_{nm}^L = \begin{bmatrix} \boldsymbol{\Phi}_{nm\mathrm{u}}^L \\ \boldsymbol{\Phi}_{nm\mathrm{l}}^L \end{bmatrix}. \tag{32}$$

The solution of (29) is expressed by summation of right eigenvectors as

$$\boldsymbol{U}_n = \sum_{m=1}^{2M} U_{nm} \boldsymbol{\Phi}_{nm}^R. \tag{33}$$

The coefficients U_{nm} are calculated by substituting (33) into (28) and using the biorthogonality relations (31). Then, the solution vector U_n is written as

$$U_n = \sum_{m=1}^{2M} \frac{\Phi_{mn}^{L^T} P_n}{(k_{nm} - k_n) B_{nm}} \Phi_{nm}^R. \tag{34}$$

The solution vector $\bar{V}_n$ in (28), occupying the upper half of U_n, takes the form

$$\bar{V}_n = \sum_{m=1}^{2M} \frac{\Phi_{mnl}^{L^T} \bar{F}_n}{(k_{nm} - k_n) B_{nm}} \Phi_{nmu}^R. \tag{35}$$

Inverse Fourier transform of (35) gives the response, $V_n(z)$, of the n-th circumferential harmonic in the spatial domain:

$$V_n(z) = \frac{1}{2\pi} \sum_{m=1}^{2M} \int_{-\infty}^{+\infty} \frac{\Phi_{mnl}^{L^T} \bar{F}_n}{(k_{nm} - k_n) B_{nm}} \Phi_{nmu}^R e^{ik_n z} \, dk_n. \tag{36}$$

In many problems, $\bar{F}_n$, Φ_{mn}^L, Φ_{nm}^R, and B_{nm} are independent of wave number k_n, so that application of the Cauchy residue theorem yields the modal response $V_n(z)$ in a straightforward way. The eigendata can be divided into two groups for travelling or decaying modes from the origin in $\pm z$ directions, respectively. Therefore, $V_n(z)$ can be expressed as summation of motions in the positive and negative directions as

$$V_n(z, \omega) = -i \sum_{m=1}^{M} \frac{\Phi_{mnl}^{L^T} \bar{F}_n}{B_{nm}} \Phi_{nmu}^R e^{ik_{nm} z} - i \sum_{m=M+1}^{2M} \frac{\Phi_{mnl}^{L^T} \bar{F}_n}{B_{nm}} \Phi_{nmu}^R e^{-ik_{nm} z}. \tag{37}$$

The response in the time domain is now obtained by applying an inverse Fourier transformation to (37) and is calculated numerically as

$$V_n(z, t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} V_n(z, \omega) e^{-i\omega t} \, d\omega. \tag{38}$$

4. Heat source representation

The heat input due to the laser pulse is assumed to be of the form

$$Q = I_0 f(t) \delta(\theta) g_r(r) g_z(z), \tag{39}$$

where I_0 is the energy of the laser pulse. The temporal profile $f(t)$ is assumed as

$$f(t) = \frac{t}{t_0^2} e^{-\frac{t}{t_0}}, \tag{40}$$

where t_0 is the pulse rise time. The spatial profile $g_z(z)$ is assumed to have a Gaussian profile in the z-direction,

$$g_z(z) = \frac{1}{\sqrt{\pi} a} e^{-z^2/a^2}, \tag{41}$$

where a is the radius of the laser beam. The depth dependence, $g_r(r)$, of the pulse is taken as

$$g_r(r) = \frac{\gamma}{r} e^{-\gamma(r_0 - r)}, \tag{42}$$

where γ is the extinction coefficient. A schematic representation of the pulse and the frequency spectrum of $f(t)$ are shown in Figures 2 and 3, respectively. Here, the nondimensional frequency, $\omega^* = \omega H / \bar{v}$, where $\bar{v} = 2.5 \times 10^2$ m/s. Note that the rise time of the pulse, t_0, has been taken to be $1\,\mu$s.
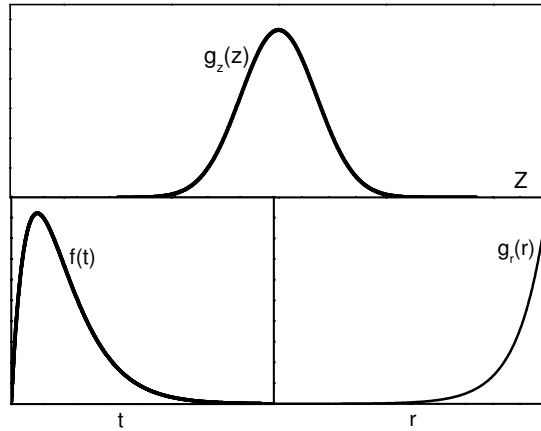
Figure 2. Temporal and spatial profiles of the pulse.
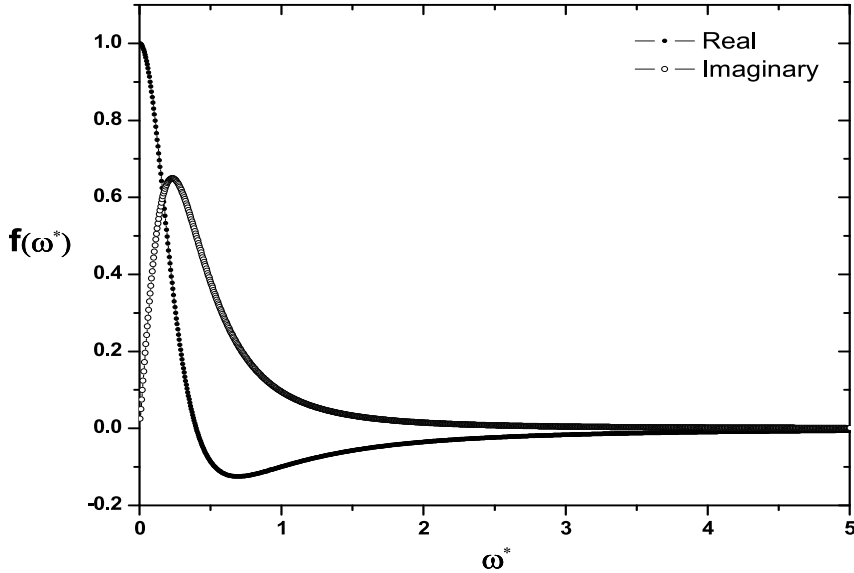
Figure 3. Frequency spectrum of $f(t)$.

Once the fundamental response functions $\boldsymbol{\Gamma}(z)$ are constructed due to $Q = I_0 f(t)\delta(\theta)g_r(r)\delta(z)$, the response due to the heat source Q, represented by (39), can be calculated as

$$\boldsymbol{v}(z_0) = \int_{-\infty}^{+\infty} \boldsymbol{\Gamma}(z) g_z(z - z_0)\, dz. \tag{43}$$

In the present work, only the thermal load is considered for constructing the response function. The spatial representation of the heat source along the radial profile in the frequency domain is

$$\boldsymbol{Q} = I_0 \bar{f}(\omega)\delta(\theta)\delta(z)\,\boldsymbol{Q}_0, \tag{44}$$

where $\boldsymbol{Q}_0 = (Q_1, Q_2, \ldots, Q_k, \ldots, Q_{2N+1})^T$ and Q_k is the value of $g_r(r)$ at the k-th node using a consistent load formulation. The solution procedure involves expansion of $\delta(\theta)$ in a Fourier series. It is well known that the Fourier series representation of a δ-function does not converge. Hence, it is necessary to replace the point source by a uniform spatial pulse of intensity q_0 over a circumferential distance $2r_o\theta_0$. For equivalence of a unit concentrated source, q_0 is given by

$$\int_{-\theta_0}^{+\theta_0} q_0 r_o d\theta = 1 \quad \text{or} \quad q_0 = \frac{1}{2r_o\theta_0}. \tag{45}$$

Therefore, the θ-dependence of $\boldsymbol{Q}$ in (44) is represented by a Fourier series expansion in the circumferential direction as

$$\boldsymbol{Q} = \sum_{n=-\infty}^{n=+\infty} e^{in\theta}\,\boldsymbol{Q}_{n\theta}, \tag{46}$$

where

$$\boldsymbol{Q}_{n\theta} = \frac{1}{2\pi r_o} \frac{\sin n\theta_0}{n\theta_0} I_0 \bar{f}(\omega)\delta(z)\,\boldsymbol{Q}_0. \tag{47}$$

The Fourier transform of $\boldsymbol{Q}_{n\theta}$ in the z-direction is

$$\bar{\boldsymbol{Q}}_{n\theta} = \frac{1}{2\pi r_o} \frac{\sin n\theta_0}{n\theta_0} I_0 \bar{f}(\omega)\,\boldsymbol{Q}_0. \tag{48}$$

Since mechanical load is not considered here, the load vector in (27) becomes

$$\bar{\boldsymbol{F}}_n = \begin{bmatrix} 0 \\ \bar{\boldsymbol{Q}}_{n\theta} \end{bmatrix} = \frac{I_0 \bar{f}(\omega)}{2\pi r_o} \frac{\sin n\theta_0}{n\theta_0} \boldsymbol{F}_0, \tag{49}$$

where $\boldsymbol{F}_0 = \begin{bmatrix} 0 & \boldsymbol{Q}_0 \end{bmatrix}^T$. Substitution of (49) into (37) and considering propagation only in the $+z$ directions yields the n-th circumferential mode displacement and temperature response functions as

$$V_n(z, \omega) = \frac{-i\, I_0 \bar{f}(\omega)}{2\pi r_o} \frac{\sin n\theta_0}{n\theta_0} \sum_{m=1}^{M} \frac{\boldsymbol{\Phi}_{mnl}^{L^T} \boldsymbol{F}_0}{B_{nm}} \boldsymbol{\Phi}_{nmu}^{R} e^{ik_{nm}z}. \tag{50}$$

Using (50) in (46), we obtain

$$V(\theta, z, \omega) = \frac{-i\, I_0 \bar{f}(\omega)}{2\pi r_o} \sum_{n=-\infty}^{n=+\infty} \frac{\sin n\theta_0}{n\theta_0} e^{in\theta} \sum_{m=1}^{M} \frac{\boldsymbol{\Phi}_{mnl}^{L^T} \boldsymbol{F}_0}{B_{nm}} \boldsymbol{\Phi}_{nmu}^{R} e^{ik_{nm}z}. \tag{51}$$

Equation (51) is used in the next section to calculate the vector V at a given location (θ, z) for different frequencies ω. Attention is focused on the convergence of the numerical results by varying the number of nodes through the thickness of the cylinder, the number of terms $|n|$ in the Fourier series, and the number of modes M.

5. Numerical results and discussion

In this section, we consider propagation of thermoelastic waves in a cylindrical shell generated by the heat source represented by (39). The material of the cylinder is taken to be silicon nitride (Si$_3$N$_4$). Amorphous and textured Si$_3$N$_4$ has been widely studied in the past for its excellent mechanical properties, such as high resistance to thermal shock, resistance to chemical attack, high fracture toughness, and good tribological and wear properties. A good review of the processing and anisotropic properties of silicon nitride for various automotive and aerospace applications can be found in [Zhu and Sakka 2008] (see also [Kitayama et al. 1999; Vogelgesang et al. 2000; Yokota and Ibukiyama 2003]). Dispersion of thermoelastic guided waves in Si$_3$N$_4$ plates and cylindrical shells was reported in [Al-Qahtani and Datta 2004; Chitikireddy et al. 2010]. Here, we present results for the frequency dependent displacements, stresses, temperature, and thermal flux at an observation point due to the laser generated heating represented by (39).

We consider an infinite cylindrical tube with thickness 0.1 mm and inner radius 0.95 mm. Thus, the nondimensional inner and outer radii are $r_i = 9.5$ and $r_o = 10.5$. The shell is discretized into N equally

Quantity	Units	Silicon nitride
$\rho = \bar{\rho}$	kg/m^3	3.20×10^3
c_{11}	N/m^2	4.33×10^{11}
c_{12}	N/m^2	1.95×10^{11}
c_{13}	N/m^2	1.27×10^{11}
c_{33}	N/m^2	5.74×10^{11}
c_{44}	N/m^2	1.08×10^{11}
β_{rr}	N/m$^2\,°$K	2.71×10^6
β_{zz}	N/m$^2\,°$K	3.22×10^6
K_{rr}	W/m $°$K	43.5
$K_{zz} = \bar{k}$	W/m $°$K	55.4
c_E	J/kg $°$K	0.67×10^3
$T_0 = \bar{T}$	$°$K	296
τ_0	s	4.32×10^{-13}
a	μm	100
t_0	μs	1.0
H	mm	0.1
γ	m^{-1}	$1 \quad \times 10^5$
I_0	Nm	1.0
$\bar{c}$	N/m^2	$2 \quad \times 10^8$

Table 1. Thermomechanical properties of Si$_3$N$_4$.

thick coaxial circular cylinders, where N is allowed to take different values (25 and 50) in order to check for the convergence of the results. As discussed later, numerical results indicate that the heat flux in the radial direction, $|q_r|$, has sharp gradients near the inner and outer surfaces of the cylinder. So, numerical results are obtained also when the thickness is subdivided into 10 layers (5 at the outer end and 5 at the inner), each of thickness 0.00175 mm, and 25 layers of thickness 0.0033 mm each in the middle.

As mentioned above, the material of the cylinder is Si_3N_4, whose properties, from [Al-Qahtani and Datta 2004], are listed in Table 1. The symmetry axis of the material is aligned with the axis of the cylinder. The localized heat source is represented by (39) with $g_z(z)$ taken as $\delta(z)$. As noted before, once the response due to this source is known, that due to any other function $g_z(z)$ can be found by convolution; see (43). In the frequency domain, the frequency spectrum of $f(t)$ is shown in Figure 3. It is seen that the dominant contribution comes from the interval $0 \leq \omega^* \leq 5$, that is, the dominant frequency is between 0 and 1.99 MHz.

All the numerical results obtained here are in nondimensional forms. The nondimensional material properties of the tube are given by

$$
\begin{array}{ccccc}
\rho^* = 1.0, & T_0^* = 1.0, & \tau_0^* = 1.081 \times 10^{-6}, & \beta_{rr}^* = 70.03, & \beta_{\theta\theta}^* = 70.03, \\
\beta_{zz}^* = 83.21, & k_{rr}^* = 0.785, & k_{\theta\theta}^* = 0.785, & k_{zz}^* = 1.00, & c_E^* = 967.50.
\end{array}
\tag{52}
$$

The nondimensional elastic stiffness tensor is given by

$$
C^* = \begin{bmatrix}
2165 & 975 & 635 & 0 & 0 & 0 \\
975 & 2165 & 635 & 0 & 0 & 0 \\
635 & 635 & 2870 & 0 & 0 & 0 \\
0 & 0 & 0 & 540 & 0 & 0 \\
0 & 0 & 0 & 0 & 540 & 0 \\
0 & 0 & 0 & 0 & 0 & 765
\end{bmatrix}.
\tag{53}
$$

Dispersion curves for circumferential modes $n = 0$ are shown in Figure 4, and for $n = 1$ in Figure 5, where γ_R and γ_I are the real and imaginary parts of the wavenumber k_n^*. It is found that the elastic modes are not affected significantly by the thermal effects within the frequency range considered here. This is consistent with the findings reported in [Chitikireddy et al. 2010].

The computation of $V(\theta, z, \omega)$ using (51) involves two summations: one over the number of axial modes M for a fixed circumferential mode n and the other over circumferential modes n. The first is determined by the number of finite elements N_e used to discretize the thickness of the cylinder. The value of θ_0, appearing in (45), has been chosen as 0.01 radians. The observation location is at $\theta = \pi/4$ and $z = r_{\text{mean}}/4$. Furthermore, $g_z(z)$ is taken to be $\delta(z)$ and γ appearing in (42) is chosen as $105 \, \text{m}^{-1}$ (see Table 1).

To test the convergence of the series with the number of elements N_e for an appropriate choice of $|n|$, results were obtained for N_e as 15, 20, 35, and 50 keeping $|n|$ fixed at 50. The value of ω^* was taken as 5. Figure 6 shows the variation of displacement, temperature, stress, and radial heat flux through the thickness with the number N_e. It is seen that 35 elements are sufficient to assure convergence. It is interesting to note the variation of $|T|$ and $|q_r|$ through the thickness. The former decreases in a nearly linear manner from the boundaries towards the middle of the shell. Note the sharp change near the boundaries so that the slopes vanish at the boundaries to satisfy the zero heat flux conditions.
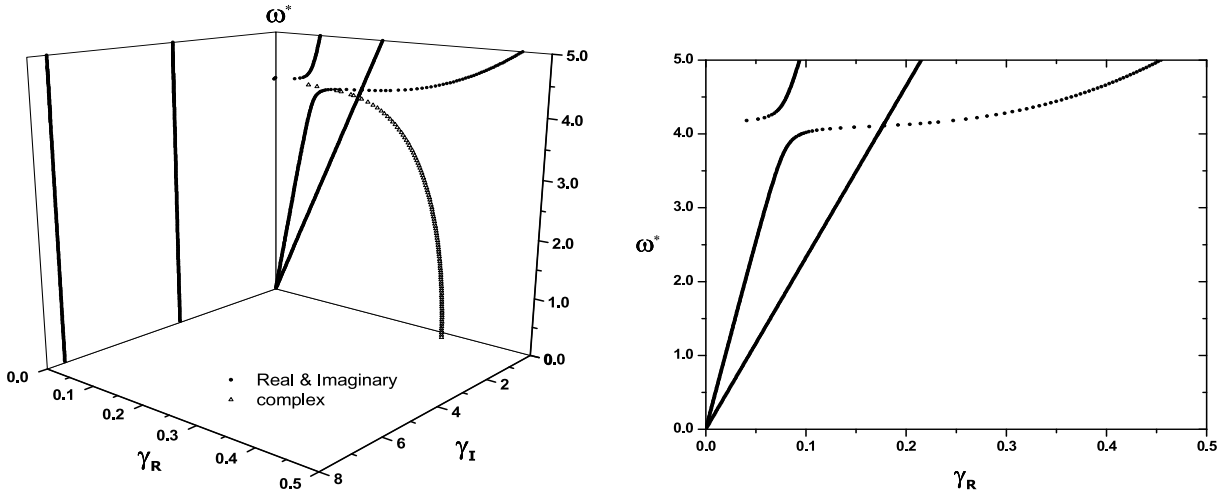
Figure 4. Frequency spectra of silicon nitride cylinder for $n = 0$. The two-dimensional diagram, corresponding to the plane $\gamma_I = 0$, shows the propagating (purely real) modes.
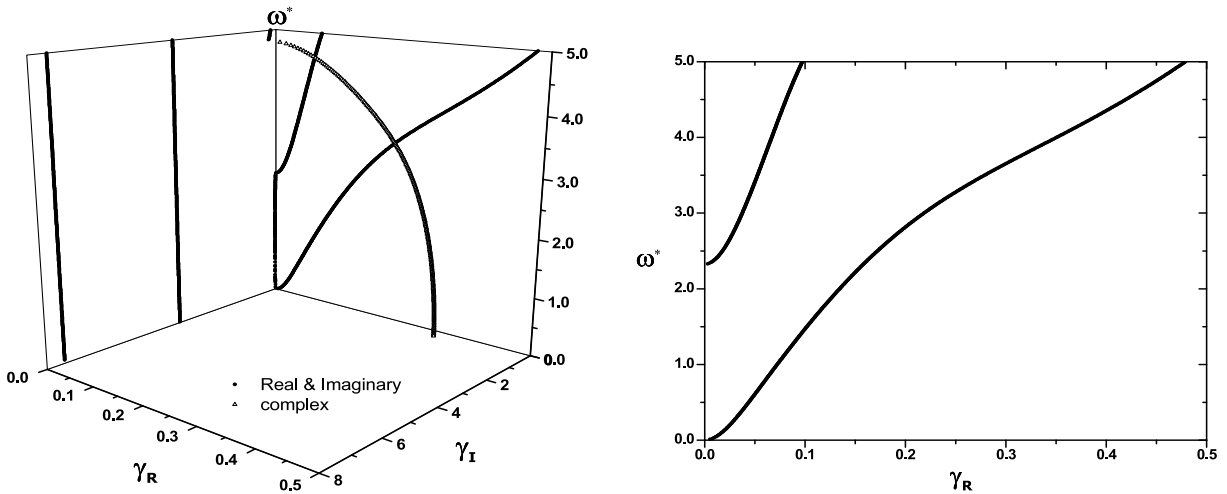
Figure 5. Frequency spectra of silicon nitride cylinder for $n = 1$. The two-dimensional diagram, corresponding to the plane $\gamma_I = 0$, shows the propagating (purely real) modes.

Consequently, the radial heat flux is nearly flat through most of the thickness and then drops steeply to zero at the boundaries. It is seen that convergence is achieved with both 35 and 50 elements.

Figures 7–9 show the changes in the distribution of the field quantities through the thickness at frequencies $\omega^* = 1$, 3, and 5. They also show variations with increasing $|n|$, which was assumed to take the values 20, 40, 50, and 60. The number of elements was kept at 35. It is seen that results converged when $|n|$ was 50. It is interesting to note that the rate of convergence increases with ω^*. This may be explained by noting the fact that the amplitudes decrease (because of diffusion) as the frequency increases. This is seen clearly in these figures.

Other features of these figures are as follows.
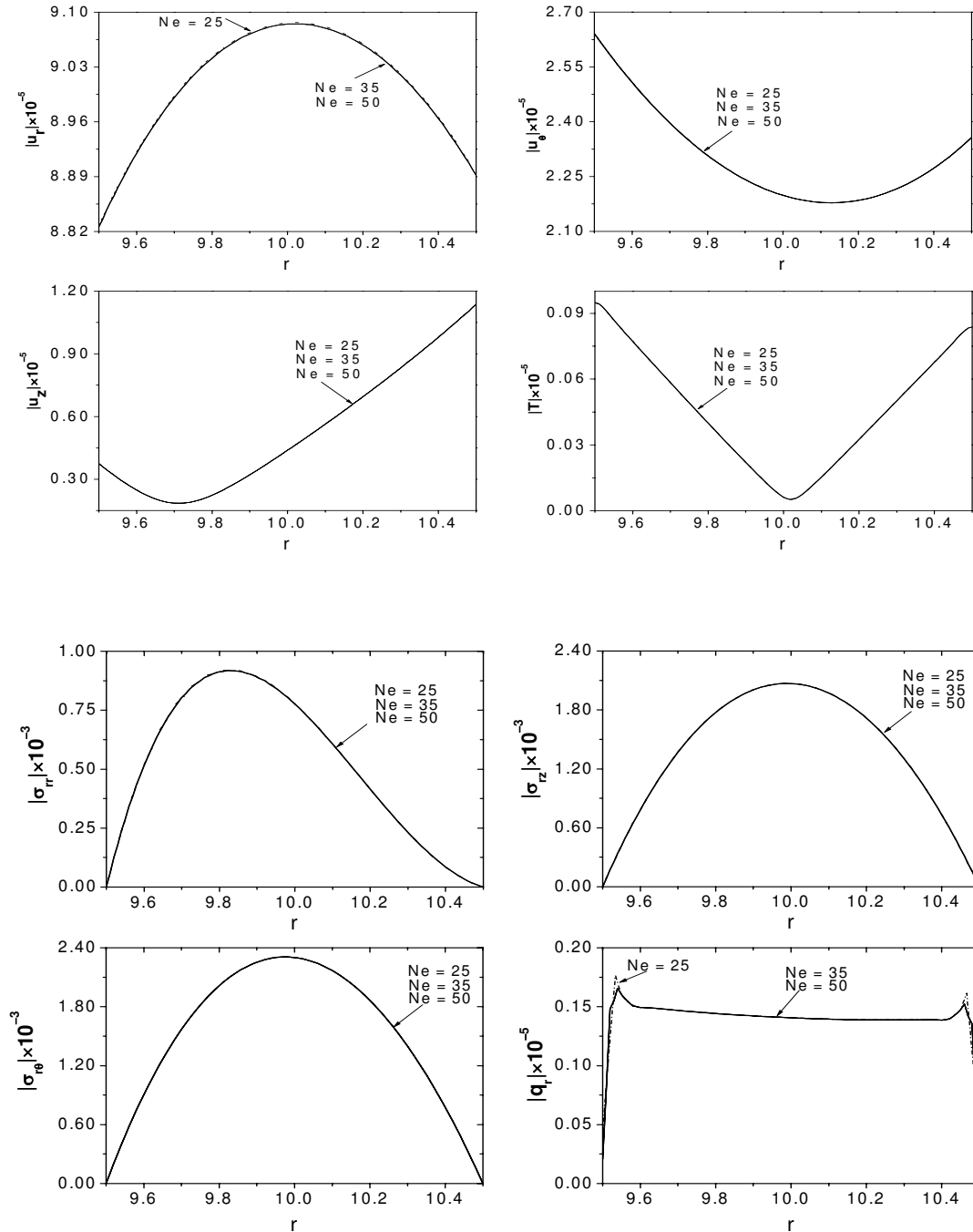
Figure 6. Through-the-thickness dependence of numerically calculated quantities for a silicon nitride cylinder at $\omega^* = 5.0$ with 25, 35, and 50 elements. The top two rows show the displacement components and the temperature distribution; the bottom two rows, the stress components and the heat flux.
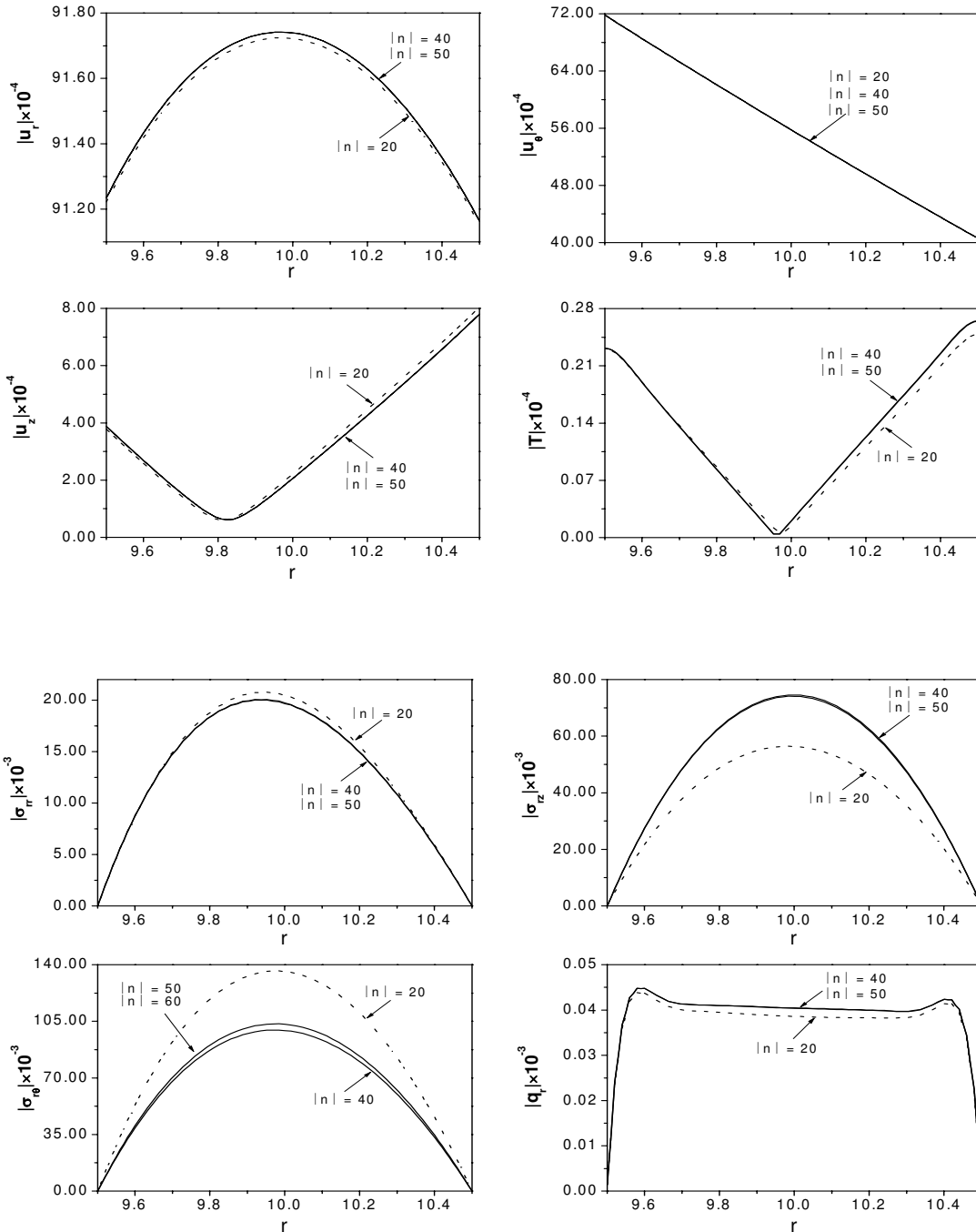
Figure 7. Through-the-thickness dependence of numerically calculated quantities for a silicon nitride cylinder at $\omega^* = 1.0$ with $|n| = 20$, 40, and 50. The top two rows show the displacement components and the temperature distribution; the bottom two rows, the stress components and the heat flux.
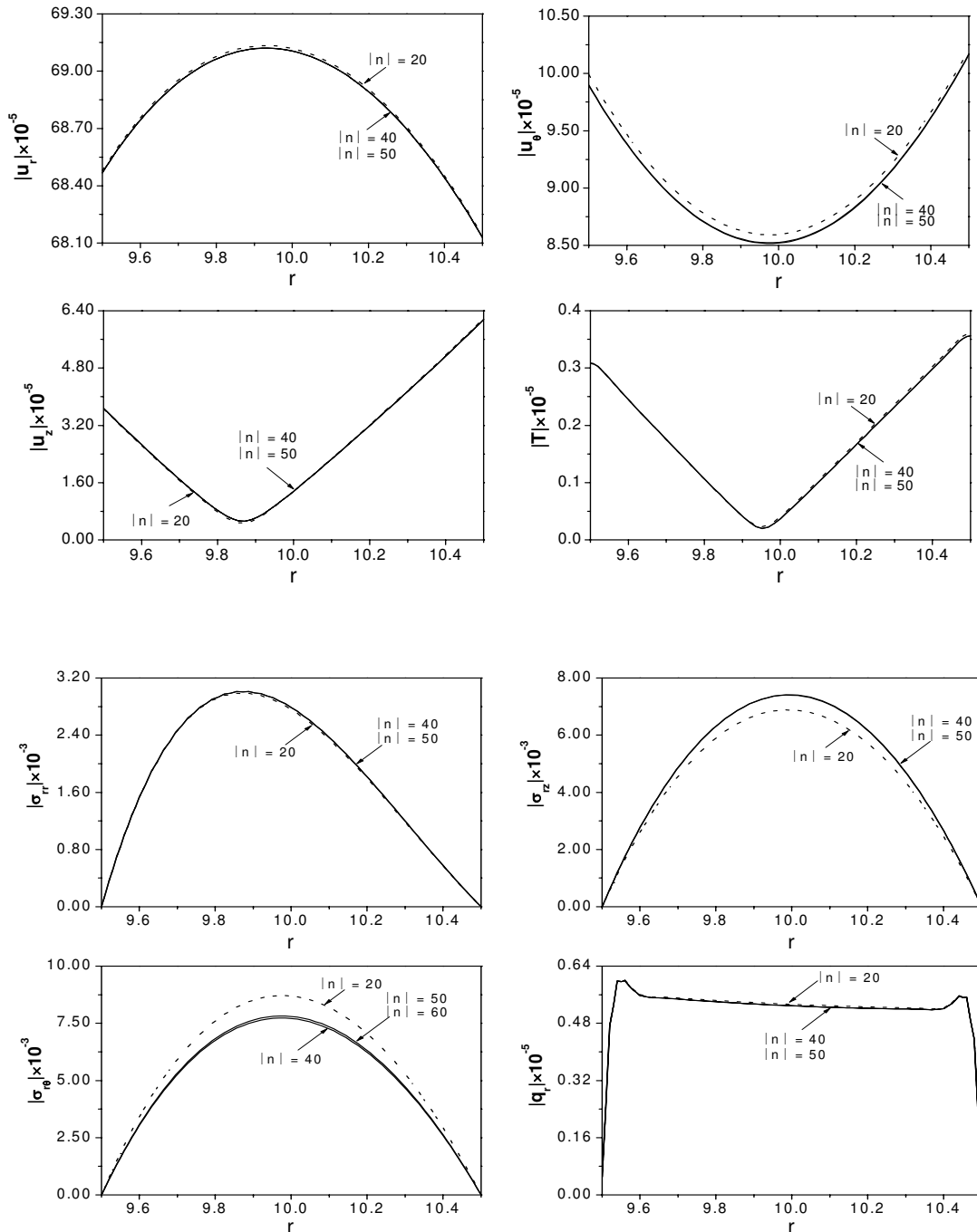
Figure 8. Through-the-thickness dependence of numerically calculated quantities for a silicon nitride cylinder at $\omega^* = 3.0$ with $|n| = 20$, 40, and 50. The top two rows show the displacement components and the temperature distribution; the bottom two rows, the stress components and the heat flux.
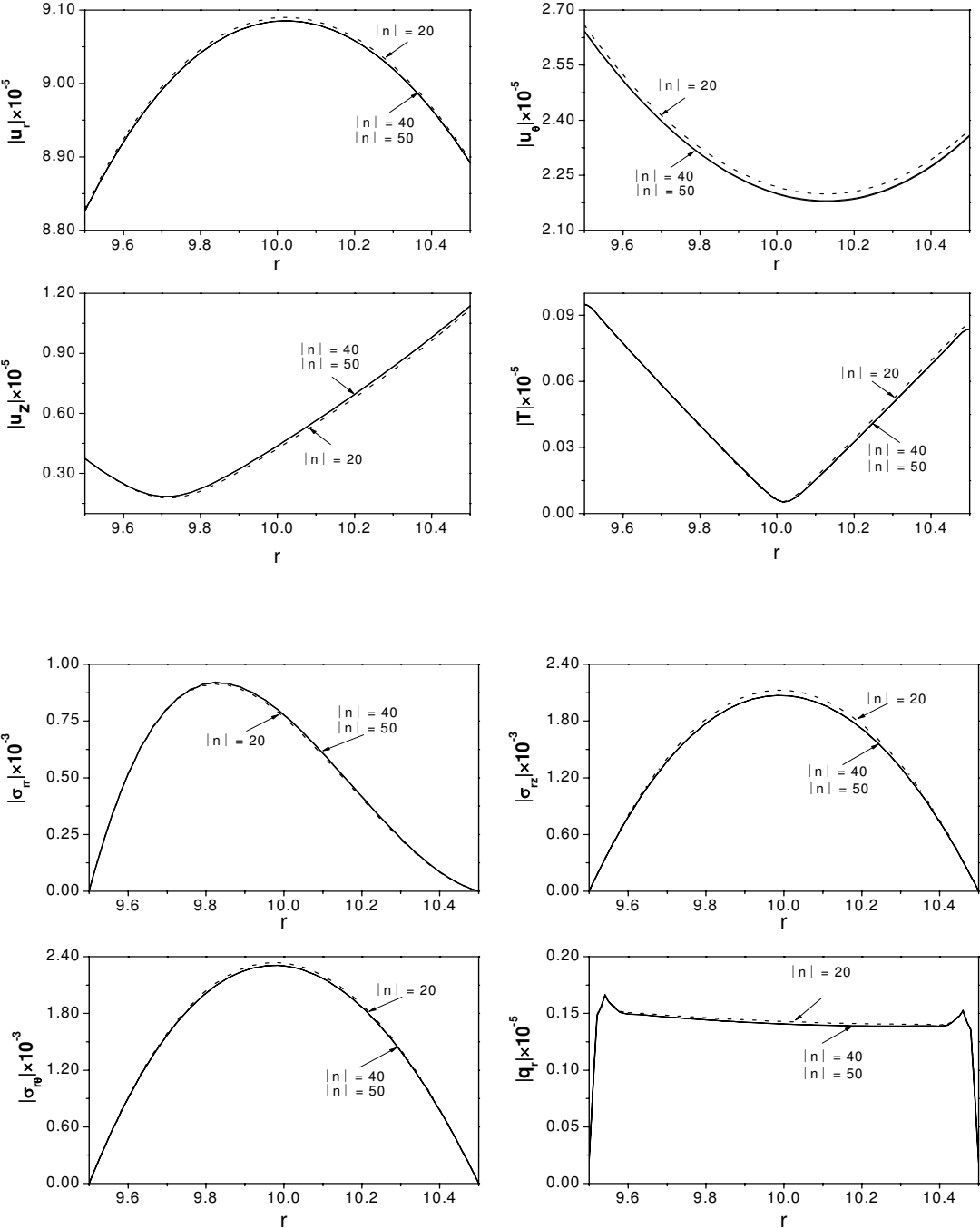
Figure 9. Through-the-thickness dependence of numerically calculated quantities for a silicon nitride cylinder at $\omega^* = 5.0$ with $|n| = 20$, 40, and 50. The top two rows show the displacement components and the temperature distribution; the bottom two rows, the stress components and the heat flux.

(1) At low frequencies, $|u_\theta|$ decreases in a linear fashion from the inner to the outer surface. As the frequency increases, it decreases first from the inner surface, reaches a minimum, and then increases with increasing radius. Both $|u_z|$ and $|T|$ also decrease first starting from the inner surface and then increase as the radius increases. The former attains the highest value at the outer surface, whereas the latter has the highest value at the outer surface at low frequencies; this feature is reversed at high frequencies.

(2) The radial displacement component and all the traction components increase from the inner surface (the latter being zero at both boundaries), reach maxima, and then decrease as the radius increases.

(3) Radial heat flux has the distinctive feature that it increases steeply from zero at the inner surface, reaches a plateau (slightly sloping towards the outer surface), and then decreases steeply to zero.

6. Conclusion

Propagation of thermoelastic waves in a circular cylindrical shell excited by a concentrated heat input has been studied in this paper. The generalized thermoelasticity theory developed by Lord and Shulman has been adopted. This theory includes a single thermal relaxation time. The material of the cylinder has been taken to be transversely isotropic Si_3N_4 as an illustrative example. However, the semianalytical finite element (SAFE) method that has been used here can also be used for generally anisotropic materials.

The heat input resembles that due to heating by a pulsed laser. It is focused at $r = r_o$, $\theta = 0$, $z = 0$, with a radial dependence that decays with decreasing r and acts on a small circular arc $2r_o\theta_0$, where $\theta_0 = .01$ radians. The solution can then be used for more general dependence on z and θ.

Although the solution presented here is for a concentrated heat source only, a concentrated mechanical load can easily be incorporated in the SAFE computations. This will be communicated in a future paper.

The results presented here are for discrete frequencies and the emphasis is on a study of convergence and distribution of the field quantities over the thickness of the cylinder. However, transient wave forms can be obtained by converting the continuous frequency spectrum using an inverse Fourier transform.

Appendix

$$
\boldsymbol{B}_1 = \begin{bmatrix} n_{1,r} & n_{2,r} & n_{3,r} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad
\boldsymbol{B}_2 = \begin{bmatrix} 0 & 0 & 0 \\ \dfrac{n_1}{r} & \dfrac{n_2}{r} & \dfrac{n_3}{r} \\ 0 & 0 & 0 \end{bmatrix}, \quad
\boldsymbol{B}_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ n_1 & n_2 & n_3 \end{bmatrix},
$$

$$
\boldsymbol{D}_1 = \begin{bmatrix}
n_{1,r} & 0 & 0 & n_{2,r} & 0 & 0 & n_{3,r} & 0 & 0 \\
\dfrac{n_1}{r} & 0 & 0 & \dfrac{n_2}{r} & 0 & 0 & \dfrac{n_3}{r} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & n_{1,r} & 0 & 0 & n_{2,r} & 0 & 0 & n_{3,r} \\
0 & n_{1,r} - \dfrac{n_1}{r} & 0 & 0 & n_{2,r} - \dfrac{n_2}{r} & 0 & 0 & n_{3,r} - \dfrac{n_3}{r} & 0
\end{bmatrix},
$$

$$\boldsymbol{D}_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \dfrac{n_1}{r} & 0 & 0 & \dfrac{n_2}{r} & 0 & 0 & \dfrac{n_3}{r} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \dfrac{n_1}{r} & 0 & 0 & \dfrac{n_2}{r} & 0 & 0 & \dfrac{n_3}{r} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \dfrac{n_1}{r} & 0 & 0 & \dfrac{n_2}{r} & 0 & 0 & \dfrac{n_3}{r} & 0 & 0 \end{bmatrix}, \quad \boldsymbol{D}_3 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & n_1 & 0 & 0 & n_2 & 0 & 0 & n_3 \\ 0 & n_1 & 0 & 0 & n_2 & 0 & 0 & n_3 & 0 \\ n_1 & 0 & 0 & n_2 & 0 & 0 & n_3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\boldsymbol{K}_{jk} = \int_r \boldsymbol{D}_j^T \boldsymbol{C} \boldsymbol{D}_k r\, dr \quad (j,k=1,2,3), \qquad \boldsymbol{K}_{0k} = \int_r \boldsymbol{D}_j^T \boldsymbol{\beta} \boldsymbol{N}_2 r\, dr \quad (k=1,2,3),$$

$$\boldsymbol{M} = \int_r \rho \boldsymbol{N}_1^T \boldsymbol{N}_1 r\, dr, \qquad \boldsymbol{g}_{jj} = \int_r \boldsymbol{B}_j^T \boldsymbol{K} \boldsymbol{B}_j r\, dr \quad (j=1,2,3),$$

$$\boldsymbol{f}_j = \int_r T_0 \boldsymbol{N}_2^T \boldsymbol{\beta}^T \boldsymbol{D}_j r\, dr \quad (j=1,2,3), \qquad \boldsymbol{m}_0 = \int_r \rho c_E \boldsymbol{N}_2^T \boldsymbol{N}_2 r\, dr, \qquad \boldsymbol{f}^e = \int_r \boldsymbol{N}_2^T \boldsymbol{f} r\, dr,$$

$$\boldsymbol{H}_1 = \begin{bmatrix} \boldsymbol{M} & 0 \\ \tau_0 \boldsymbol{f}_1 & \tau_0 \boldsymbol{m}_0 \end{bmatrix}, \quad \boldsymbol{H}_2 = \begin{bmatrix} 0 & 0 \\ \tau_0 \boldsymbol{f}_2 & 0 \end{bmatrix}, \quad \boldsymbol{H}_3 = \begin{bmatrix} 0 & 0 \\ \tau_0 \boldsymbol{f}_3 & 0 \end{bmatrix}, \quad \boldsymbol{H}_4 = \begin{bmatrix} 0 & 0 \\ \boldsymbol{f}_1 & \boldsymbol{m}_0 \end{bmatrix}, \quad \boldsymbol{H}_5 = \begin{bmatrix} 0 & 0 \\ \boldsymbol{f}_2 & 0 \end{bmatrix},$$

$$\boldsymbol{H}_6 = \begin{bmatrix} 0 & 0 \\ \boldsymbol{f}_3 & 0 \end{bmatrix}, \quad \boldsymbol{H}_7 = \begin{bmatrix} \boldsymbol{K}_{11} & -\boldsymbol{K}_{01} \\ 0 & \boldsymbol{g}_{11} \end{bmatrix}, \quad \boldsymbol{H}_8 = \begin{bmatrix} \boldsymbol{K}_{12} - \boldsymbol{K}_{21} & \boldsymbol{K}_{02} \\ 0 & 0 \end{bmatrix}, \quad \boldsymbol{H}_9 = \begin{bmatrix} \boldsymbol{K}_{13} - \boldsymbol{K}_{31} & \boldsymbol{K}_{03} \\ 0 & 0 \end{bmatrix},$$

$$\boldsymbol{H}_{10} = \begin{bmatrix} -\boldsymbol{K}_{22} & 0 \\ 0 & -\boldsymbol{g}_{22} \end{bmatrix}, \quad \boldsymbol{H}_{11} = \begin{bmatrix} -(\boldsymbol{K}_{23} + \boldsymbol{K}_{32}) & 0 \\ 0 & 0 \end{bmatrix}, \quad \boldsymbol{H}_{12} = \begin{bmatrix} -\boldsymbol{K}_{33} & 0 \\ 0 & -\boldsymbol{g}_{33} \end{bmatrix}.$$

References

[Al-Qahtani and Datta 2004] H. M. Al-Qahtani and S. K. Datta, "Thermoelastic waves in an anisotropic infinite plate", *J. Appl. Phys.* **96**:7 (2004), 3645–3658.

[Al-Qahtani et al. 2005] H. M. Al-Qahtani, S. K. Datta, and O. M. Mukdadi, "Laser-generated thermoelastic waves in an anisotropic infinite plate: FEM Analysis", *J. Thermal Str.* **28**:11 (2005), 1099–1122.

[Chitikireddy et al. 2010] R. Chitikireddy, H. Bai, A. H. Shah, and S. K. Datta, "Thermoelastic waves in an anisotropic cylinder", *J. Thermal Str.* **33**:2 (2010), 97–120.

[Clorennec and Royer 2003] D. Clorennec and D. Royer, "Analysis of surface acoustic wave propagation on a cylinder using laser ultrasonics", *Appl. Phys. Letters* **82**:25 (2003), 4608–4610.

[Datta and Shah 2009] S. K. Datta and A. H. Shah, *Elastic waves in composite media and structures: with applications to ultrasonic nondestructive evaluation*, CRC Press, 2009.

[Elnagar and Abd-Alla 1987] A. M. Elnagar and A. M. Abd-Alla, "On a generalized thermo-elastic problem in an infinite cylinder under initial stress", *Earth, Moon and Planets* **37**:3 (1987), 213–223.

[Erbay and Şuhubi 1986] S. Erbay and E. S. Şuhubi, "Longitudinal wave propagation in a generalized thermoelastic cylinder", *J. Thermal Str.* **9**:3 (1986), 279–295.

[Green and Lindsay 1972] A. E. Green and K. A. Lindsay, "Thermoelasticity", *J. Elast.* **2**:1 (1972), 1–7.

[Green and Naghdi 1993] A. E. Green and P. M. Naghdi, "Thermoelasticity without energy dissipation", *J. Elast.* **31**:3 (1993), 189–208.

[Ignaczak and Ostoja-Starzewski 2010] J. Ignaczak and M. Ostoja-Starzewski, *Thermoelasticity with finite wave speeds*, Oxford University Press, New York, 2010.

[Kitayama et al. 1999] M. Kitayama, K. Hirao, M. Toriyama, and S. Kanazaki, "Thermal conductivity of β-Si$_3$N$_4$, I: effects of various microstructural factors", *J. Amer. Ceram. Soc.* **82**:11 (1999), 3105–3112.

[Lord and Shulman 1967] H. W. Lord and Y. Shulman, "A generalized dynamical theory of thermoelasticity", *J. Mech. Phys. Solids* **15**:5 (1967), 299–309.

[Pan et al. 2004] Y. Pan, C. Rossignol, and B. Audoin, "Acoustic waves generated by a laser point source in an isotropic cylinder", *J. Acoust. Soc. Am.* **116**:2 (2004), 814–820.

[Pan et al. 2006] Y. Pan, M. Perton, B. Audoin, and C. Rossignol, "Acoustic waves generated by a laser point source in a transversely isotropic cylinder", *J. Acoust. Soc. Am.* **119**:1 (2006), 243–250.

[Ponnusamy 2007] P. Ponnusamy, "Wave propagation in a generalized thermoelastic solid cylinder of arbitrary cross-section", *Int. J. Solids Str.* **44**:16 (2007), 5336–5348.

[Scruby and Drain 1990] C. B. Scruby and L. E. Drain, *Laser ultrasonics: techniques and applications*, Adam Hilger, New York, 1990.

[Sharma 2001] J. N. Sharma, "Three-dimensional analysis of a homogeneous transversely isotropic thermoelastic cylindrical panel", *J. Acoust. Soc. Am.* **110**:1 (2001), 254–259.

[Sharma and Sharma 2002] J. N. Sharma and P. K. Sharma, "Free vibration analysis of homogeneous transversely isotropic thermoelastic cylindrical panel", *J. Thermal Str.* **25**:2 (2002), 169–182.

[Sharma et al. 2000] J. N. Sharma, R. Singh, and R. Kumar, "Generalized thermoelastic waves in homogeneous isotropic plates", *J. Acoust. Soc. Am.* **108**:2 (2000), 848–851.

[Verma 2002] K. L. Verma, "On the propagation of waves in layered anisotropic media in generalized thermoelasticity", *Internat. J. Engrg. Sci.* **40**:18 (2002), 2077–2096.

[Verma and Hasebe 2001] K. L. Verma and N. Hasebe, "Dispersion of thermoelastic waves in a plate with and without energy dissipation", *Int. J. Thermophys.* **22**:3 (2001), 957–978.

[Vogelgesang et al. 2000] R. Vogelgesang, M. Grimsditch, and J. S. Wallace, "The elastic constants of single crystal β-Si$_3$N$_4$", *Appl. Phys. Lett.* **76**:8 (2000), 982–984.

[Yokota and Ibukiyama 2003] H. Yokota and M. Ibukiyama, "Microstructure tailoring for high thermal conductivity of β-Si$_3$N$_4$ ceramics", *J. Amer. Ceram. Soc.* **86**:1 (2003), 197–199.

[Zhu and Sakka 2008] X. Zhu and Y. Sakka, "Textured silicon nitride: processing and anisotropic properties", *Sci. Technol. Adv. Mater.* **9**:3 (2008), 1–47.

[Zhuang et al. 1999] W. Zhuang, A. H. Shah, and S. B. Dong, "Elastodynamic green's function for laminated anisotropic circular cylinders", *J. Appl. Mech.* (*ASME*) **66**:3 (1999), 665–673.

HAO BAI: hbai@lakeheadu.ca
Department of Mechanical Engineering, Lakehead University, 955 Oliver Rd, Thunder Bay, ON P7B 5E1, Canada

RAVI CHITIKIREDDY: umchitik@cc.umanitoba.ca
Department of Civil Engineering, University of Manitoba, Winnipeg, MB R3T 2N2, Canada

ARVIND H. SHAH: shah@cc.umanitoba.ca
Department of Civil Engineering, University of Manitoba, Winnipeg, MB R3T 2N2, Canada

SUBHENDU K. DATTA: Subhendu.Datta@colorado.edu
Department of Mechanical Engineering, University of Colorado, Boulder, CO 80309-0427, United States

■■
■msp

DYNAMICAL CHARACTERIZATION OF MIXED FRACTAL STRUCTURES

Luiz Bevilacqua and Marcelo M. Barros

It is because of people like Marie-Louise and Charles that it is worth fighting for a better world.

We present a new technique to determine the fractal or self-similarity dimension of a sequence of curves. The geometric characterization of the sequence is obtained from the mechanical properties of harmonic oscillators with the same shape of the terms composing the given sequence of curves. The definition of "dynamical dimension" is briefly introduced with the help of simple examples. The theory is proved to be valid for a particular type of curves as those of the Koch family. The method is applied to more complex plane curves obtained by superposing two generators of the Koch family with different fractal dimensions. It is shown that this structure is composed by two series of objects one of which is fractal and the other which is not rigorously a fractal sequence but approaches asymptotically a fractal object. The notion of quasifractal structures is introduced. The results are shown to provide good information about the structure formation. It is shown that the dynamical dimension can identify randomness for certain fractal curves.

1. Introduction

The correlation between the form and the physical properties of certain objects and the fractal characteristics of their geometry has called the attention of several researchers [Feder 1988; Gouyet 1996; Mandelbrot 1982; Mauroy et al. 2004]. However the determination of the geometric fractal dimension of a given sequence of objects using the associated sequence of a selected physical property has not yet been explored as far as we know. In previous papers we have shown that coupling between physics and geometry of fractal objects can be used to determine the fractal dimension of curves belonging to the Koch family.

It was shown that dynamical properties of curves belonging to a fractal sequence can also be fractal. Namely, the periods of a sequence of simple oscillators associated to a given Koch sequence have been successfully used to determine the geometric fractal dimension of the given sequence [Bevilacqua et al. 2008]. We present below the more important results obtained for Koch curves that will help to understand the numerical approach used here to deal with complex curves as explained later on.

We say that a sequence of curves belongs to the Koch family if the k-th order term contains N_k segments with the same length λ_k. The number of segments and the respective lengths are given by $N_k = p^k$ and $\lambda_k = L_0/q^k$ where L_0 is the initial basis or the initiator and p, q are integers. We will use this definition throughout this paper.

Consider a sequence of springs consisting of wires folded in such a way as to reproduce the same geometric shape of the corresponding terms of a given sequence of fractal curves. With these springs is
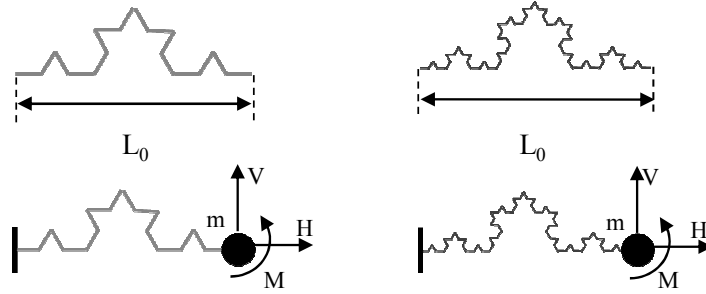
Figure 1. Top: two terms of the fractal geometric sequence. Bottom: Corresponding dynamic oscillators used to characterize the geometric sequence. The mass attached at the extremity of the spring is represented by m.

then possible to construct a simple spring-mass harmonic oscillators sequence by clamping one of the extremities and attaching a mass m at the free extremity. This oscillator sequence as shown in Figure 1 will be called the dynamic characteristic sequence.

It can be shown that it is possible to derive relatively simple relations between the geometric fractal dimension of Koch curves and the fundamental periods of the related harmonic oscillators. Note that the fundamental periods refer to each oscillator as a single isolated spring mass system. That is for a sequence comprising n terms each one defined by the pair (N_k, λ_k) $k = 1, 2, \ldots, n$ where N_k and λ_k stand for the number of terms and their lengths respectively there will be n simple harmonic oscillators each one characterized by the corresponding fundamental period T_k^f. The superscript f stays for the respective degree of freedom excited by the initial conditions induced by a horizontal force H, a vertical force V or a moment M. Since N_k can be written as a function of λ_k it is therefore possible to plot the elementary length λ_k against the fundamental frequency T_k^f. Figure 2 illustrates the polygonal curve representing the relation between the logarithm of the normalized variables λ_k/L_0 and T_k^H/T_0^H up to the sixth term for a given Koch sequence corresponding to an initial excitation induced by a horizontal force H.

We claim that the slopes of the segments composing the polygonal curve tend to a fixed value s which is correlated with the fractal dimension D of the Koch sequence. That is $\lim_{k \to \infty} s_k = s$ and consequently
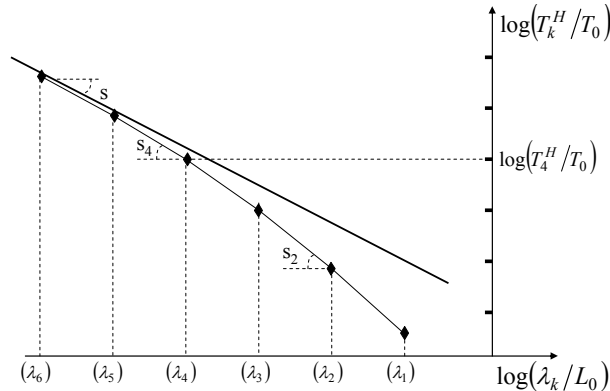
Figure 2. The vertices of the polygonal curve represent the relation between the terms of order k and the normalized frequencies of the corresponding oscillators.
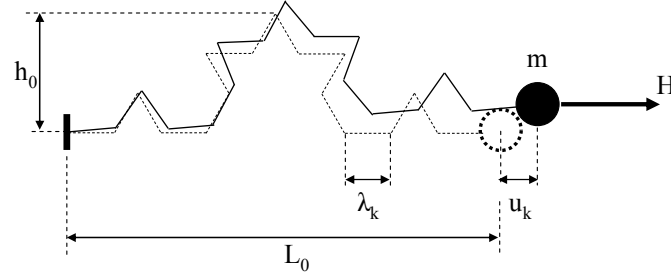
Figure 3. Oscillator corresponding to a term k of the Koch triadic carrying a mass m at the free end and excited by an initial displacement induced by a horizontal force H.

the slope of the last segment in the polygonal curve leads to the fractal dimension of the Koch sequence with increasing precision.

In order to keep this paper as self-contained as possible we will reproduce the proof of the convergence of the slopes s_k to the slope s as $k \to \infty$ or equivalently as $\lambda_k \to 0$.

Consider the Koch triadic sequence. Figure 3 represents the simple oscillator corresponding to a general term k in the sequence.

Let us assume that all the independent oscillators carry an equal mass m at the free end. Imposing an initial displacement generated by a horizontal force H applied at the free end the motion is governed by the elementary equation

$$m\frac{d^2 w_k}{dt^2} + \frac{w_k}{c_{11}^{(k)}} = 0, \tag{1}$$

where w_k is the generalized displacement and $c_{11}^{(k)}$ is the compliance or the inverse of the rigidity. For linear elastic structures the compliance is given by

$$c_{11}^{(k)} = \frac{\partial W_k}{\partial H}\bigg|_{H=1} \tag{2}$$

W_k stands for the stored elastic energy. For the system under consideration the elastic energy stored in a general term k is primarily due to the bending moment distributed along the N_k segments with length λ_k composing the term of order k in the Koch triadic. Therefore the stored bending energy for the k-th order term is

$$W_k = \frac{1}{2} \int_0^{L_t} \frac{M_k^2(s)}{EI} ds = \frac{1}{2} \frac{1}{EI} \sum_{i=0}^{N_k} \int_0^{\lambda_k} (M_{i-1,i}^{(k)}(s))^2 ds, \tag{3}$$

where E is the Young modulus of the wire material and I the moment of inertia of the wire cross section. Both will be assumed constant for all the oscillators. $M_{i-1,i}^{(k)}$ is the bending moment acting on the elementary segment $(i-1, i)$ as shown in Figure 4 and N_k is the total number of segments in the k-th order term. We are disregarding the contribution of the shear and normal forces to the strain energy. Now all the oscillators, for all k, fit into a box $L_0 \times h_0$ as can be seen from Figure 4. The bending moment along a segment $(i-1, i)$ is

$$M_{i-1,i}^{(k)}(s) = H[y_{i-1}^{(k)} + (y_i^{(k)} - y_{i-1}^{(k)})s] \quad \text{where } 0 \le s \le 1. \tag{4}$$
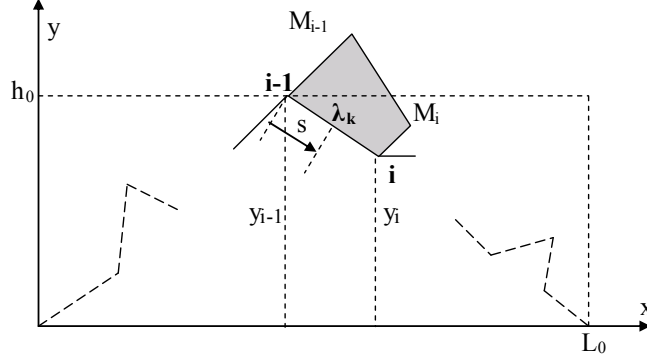
Figure 4. Bending moment along the segment $(i - 1, i)$ for a general term k of the oscillator sequence.

Now introducing (4) into (3), integrating over all segments λ_k and summing up we get

$$W_k = \frac{1}{2} \frac{H^2}{E_k I_k} \lambda_k h_0^2 N_k \Omega_k,$$

where

$$\Omega_k = \frac{1}{N_k} \sum_{i=1}^{N_k} \alpha_i(k)$$

and $\alpha_i(k) = \frac{1}{3}[z_{i-1}^2 + z_{i-1}z_i + z_i^2]$ with $z_j = y_j/h_0$.

From the definition of h_0 clearly $z_j \leq 1$ for all j, and consequently $\alpha_i(k) \leq 1$.

The compliance can now be derived from the stored energy function W_k:

$$c_{11}^{(k)} = \frac{\partial W_k}{\partial H}\bigg|_{H=1} = \frac{h_0^2}{E_k I_k} N_k \lambda_k \Omega(k). \tag{5}$$

Introducing this expression into (1) we obtain

$$\frac{d^2 u_k}{dt^2} + \frac{1}{(T_k^H)^2} u_k = 0, \tag{6}$$

where the period T_k^H is given by

$$T_k^H = \left(\frac{m h_0^2 L_0}{EI} \frac{\lambda_k}{L_0} N_k \Omega_k \right)^{1/2}. \tag{7}$$

Now from the definitions of N_k and λ_k for curves of the Koch family there comes immediately

$$\log N_k = -\log\left(\frac{\lambda_k}{L_0}\right) \frac{\log p}{\log q}. \tag{8}$$

Introducing the value of N_k given by the equation above into (7) and after some straightforward calculations we obtain

$$\log \frac{T_k^H}{T_0^I} = \frac{1}{2} \log \Omega_k + \frac{1}{2}(1 - D) \log \frac{\lambda_k}{L_0}, \tag{9a}$$

where $D = \log p / \log q$ and T_0^I is a reference period:

$$(T_0^I)^2 = \frac{m_0 h_0^2 L_0}{E_0 I_0}.$$

The parameter D is the dynamical fractal dimension. It coincides with the box and the Hausdorff fractal dimensions provided that the mass, the Young modulus and the diameter of the wire cross section are all constant.

Now, if the geometric fractal dimension of the Koch sequence can be determined through the sequence of the periods of the corresponding oscillators, it is necessary that the (9a) plotted on the plane $Y_k \times X_k$, with $Y_k = \log(T_k/T_0)$ and $X_k = \log(\lambda_k/L_0)$, approaches a straight line whose angular coefficient is equal to $(1-D)/2$ as shown in Figure 2. Define the functional relation $Y_k \Leftrightarrow X_k$ as a polygonal curve composed by straight segments connecting the points (X_k, Y_k); (X_{k+1}, Y_{k+1}). Let us prove the asymptotic behavior of the polygonal curve. The following lemma is proved in the Appendix.

Lemma. *For curves belonging to the Koch family — class of curves defined by $N_k = p^k$ and $\lambda_k/L_0 = 1/q^k$ — the first order differential form of the quadratic term Ω_k with respect to λ_k is finite for increasing values of k, or equivalently decreasing values of λ_k. That is,*

$$\lim_{k \to \infty} (\Delta\Omega_k/\Delta\lambda_k) = \lim_{\lambda_k \to 0} (\Delta\Omega_k/\Delta\lambda_k)$$

is finite.

Now recalling (9a) and with $Y_k = \log(T_k/T_0)$ and $X_k = \log(\lambda_k/L_0)$ the calculation of the differential ratio $\Delta Y_k/\Delta X_k$ after some simple operations gives

$$\frac{\Delta Y_k}{\Delta X_k} = \frac{1}{2\Omega_k}\frac{\Delta\Omega_k}{\Delta\lambda_k}\lambda_k + \tfrac{1}{2}(1-D).$$

Therefore from the lemma and since Ω_k is finite and nonzero we have

$$\lim_{\lambda_k \to 0} \frac{\Delta Y_k}{\Delta X_k} = \tfrac{1}{2}(1-D).$$

Proposition 1. *As $k \to \infty$ the curve given by* (9a) *approaches asymptotically a straight line with slope equal to $(1-D)/2$.*

It was shown that the oscillation period sequence approaches asymptotically a fractal sequence whose fractal dimension exhibits a simple correlation with the geometric fractal dimension for the case of an excitation induced by a horizontal force. Similarly it can be shown that the sequences corresponding to the other two initial conditions, triggered by a vertical force or a moment, are governed by similar laws, namely

$$\log \frac{T_k^V}{T_0^{II}} = \tfrac{1}{2}\log \Psi_k + \tfrac{1}{2}(1-D)\log\frac{\lambda_k}{L_0} \tag{9b}$$

for a vertical force and

$$\log \frac{T_k^M}{T_0^{III}} = \tfrac{1}{2}(1-D)\log\frac{\lambda_k}{L_0} \tag{9c}$$

for a moment.

The parameter D is the fractal dimension of the primordial geometric sequence provided that the mechanical properties of the wires acting as springs are the same for all elements.

Note that for the initial condition induced by a moment the sequence of the normalized periods of the simple oscillators follow exactly a power law. The reason is that for this case the bending moment is the same for all the elementary segments, that is, the strain energy is uniformly distributed along the elements composing the respective curve.

Numerical experiments have clearly shown that the fractal dimension of a generalized Koch curve can be determined by the periods of a related oscillator sequence provided that the wire cross section and Young modulus remain the same for all oscillators. It can be shown that if the mass m at the free end is not constant but proportional to the total curve length, for each oscillator, that is, for the k-th oscillator $m_k = \rho N_k \lambda_k$ the factor multiplying $\log(\lambda_k / L_0)$ in (9a)–(9c) should be $1 - D$ instead of $(1 - D)/2$.

The fractal dimension determined with the method introduce above will be called dynamical fractal dimension irrespectively of the value of the mass, constant or not.

Now suppose we are given just one term, sufficiently large, of a hypothetically fractal sequence. The problem now is to find out if the given sample really belongs to some fractal sequence and if so to determine the respective fractal dimension. Let us call this first curve the reference term. Following the same technique exposed before build a spring-mass system using a folded wire with the same shape as the reference term carrying a mass m at one of the extremities and keeping the other clamped (Figure 5, left). It is possible, as already discussed [Bevilacqua et al. 2008] to find three fundamental frequencies corresponding to three selected excitation introduced by a horizontal force H, a vertical force V and a concentrated moment M.

From the reference term — Figure 5, left — cut a piece off the extremity to obtain a new sample with length $L_{n+1} = bL_n$ where $b < 1$ is the scale factor. This operation may be repeated successively to obtain a sequence of samples. Now to each sample corresponds a simple harmonic oscillator and the corresponding periods for all three types of excitation can be determined. Performing this operation in successive steps (Figure 5) it is possible to find a correlation between the periods and the lengths of the sample projections on the horizontal axis. It is convenient, in order to simplify the calculations, to agree on a constant reduction factor $b = L_m / L_{m-1}$ to cut the successive samples.

Call b_m the variable representing the ratio L_m / L_n where L_m is the length of the horizontal projection of the sample m and L_n is the length of the horizontal projection of the reference term. Therefore we have $b_m = b^m$. For the classical Koch curves it is possible to show that:
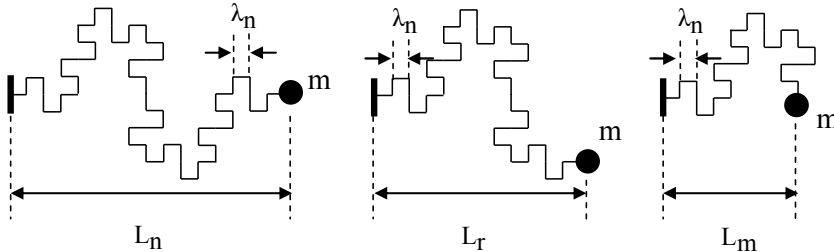
Figure 5. Sequence of three general samples ($m > r > n$) taken from the reference term whose projections on the horizontal axis are L_m, L_r, L_n. The scale b is defined by $L_{r-1} = bL_r$.

(A) For oscillations induced by the action of a concentrated moment:

$$\left(\log \frac{T_{n/m}}{T_n} \right)^{(M)} = \frac{D}{2} \log(b_m) + \Phi_m^{(M)}. \tag{10a}$$

(B) For oscillations induced by the action of a horizontal force or a vertical force we obtain the equations

$$\left(\log \frac{T_{n/m}}{T_n} \right)^{(H)} = \left(1 + \frac{D}{2} \right) \log b_m + \Phi_m^{(H)} \tag{10b}$$

$$\left(\log \frac{T_{n/m}}{T_n} \right)^{(V)} = \left(1 + \frac{D}{2} \right) \log b_m + \Phi_m^{(V)}. \tag{10c}$$

The parameters $\Phi_m^{(M,H,V)}$ in these equation can be interpreted as a kind of noise perturbation intrinsic to the method. The interval of variation of the perturbation $\Phi_m^{(M,H,V)}$ as function of m depends on several conditions other than the scale factor b, as the type of the initial excitation and the geometry of the curves. The analysis of $\Phi_m^{(M,H,V)}$ as a function of m is rather complex even for simple Koch curves. It is possible to show that a first estimation of the relative deviation of the sequence of normalized periods $(T_{n/m}/T_n)$ from the theoretical power law given by $(b_m)^{(1+D/2)}$ is

$$\frac{T_{n/m}/T_n}{(b_m)^{(1+D/2)}} \propto \sqrt{\frac{1 - 2m\varepsilon_0}{1 - mD\varepsilon_0}}, \quad \text{where } \varepsilon_0 = 1 - b > 0.$$

Several numerical experiments were tried for different types of Koch curves. It has been observed that for oscillations induced by a horizontal force the local perturbation is large while for the other two types of excitation induced by vertical force and concentrated moment the noise is very small. In any case the average values obtained with this technique for all boundary conditions are very good particularly for ratios $L_m/L_n > 0.6$.

This paper is intended to show that the dynamical approach is equally applicable for a new class of more complex curves that will be called mixed fractals. Numerical experiments with mixed fractal geometries confirm that the theory developed for Koch curves presents just as good results. In other words, Equations (9a)–(9c) and (10a)–(10c) can be extended to more complex curves.

2. Mixed fractals and quasifractal structures

Figure 6 shows two types of Koch curves C_A and C_B with self-similar or fractal dimensions D_A and D_B respectively. From the definition of Koch curves we may write

$$\log N_k = k \log p$$

and

$$\log \frac{\lambda_k}{L_0} = -k \log q.$$

Eliminating k we obtain

$$\log N_k = -D \log \frac{\lambda_k}{L_0}, \quad \text{where } D = \frac{\log p}{\log q}.$$
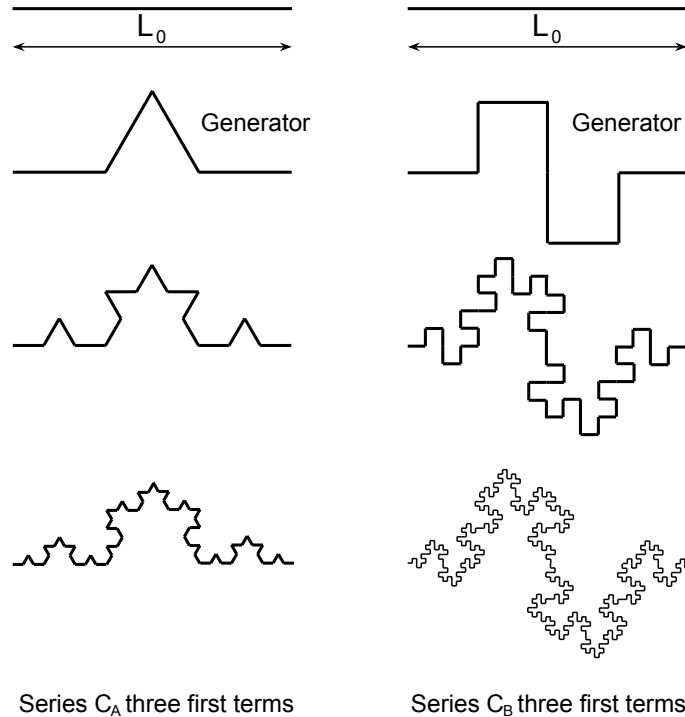
Figure 6. Koch curves C_A and C_B used to generate a new object O. The fractal dimensions are $D_A = \log 4/\log 3$ and $D_B = 1.5$.

The parameter D is the self-similarity dimension or the fractal dimension. For the Koch triadic C_A and the Koch quadratic C_B, we get $D_A = \log 4/\log 3 = 1.261859$ and $D_B = \log 8/\log 4 = 1.5$ respectively.

This paper deals with a series of objects $O : \{O_1, O_2, \ldots, O_n, \ldots\}$ generated with the help of a particular arrangement of two or more curves of the Koch type. We will call this class of objects *mixed fractal curves*. In order to illustrate these ideas let us build a mixed fractal sequence with the help of two particular curves, namely the Koch triadic (C_A) and the Koch quadratic (C_B).

The generation process of a *mixed fractal curve* using the Koch curves C_A and C_B with fractal dimensions D_A and D_B respectively is governed by the following law of formation.

The first term O_1 coincides exactly with the first term, the generator, of the C_A series assembled on a basis with length L_0. That is O_1 consists of $N_1^A = p_A$ segments with length $\lambda_1^A = L_0/q_A$. The second term of the O series is obtained by selecting p_A generators of the series C_B properly scaled such that all of them fit to a basis of length λ_1^A. This means that O_2 consists of $p_A p_B$ segments with length $\lambda_1^B = \lambda_1^A/q_B = L_0/q_A q_B$. The next curves can be obtained repeating the procedure described above, that is by using the elementary segments of the current curve as basis for the following generator of C_A or of C_B properly downscaled. Switching these generators in successive steps all terms of the mixed fractal sequence may be obtained.

Let C_A be the Koch triadic and C_B the Koch quadratic. Taking as initiator the generator of the Koch triadic, ignoring the trivial initiator L_0, the process is simply to switch step by step the generators of both
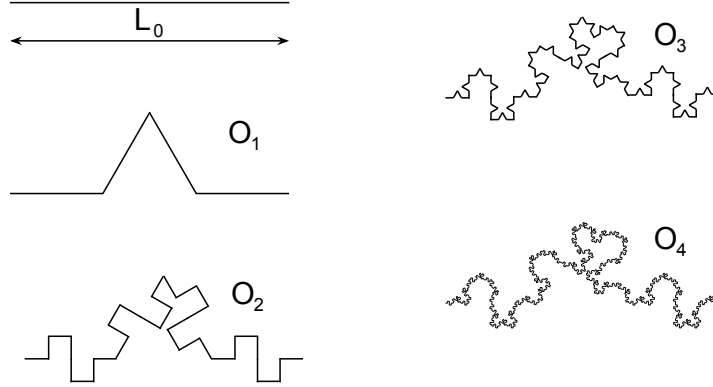
Figure 7. Terms of the series O built up with the Koch curves C_A and C_B.

curves properly scaled to build up a new term of the series O as described above. Figure 7 depicts four terms of the series.

Now, let us investigate the fractal characteristics of this new series built by overlaying alternatively the generators of the triadic and of the quadratic Koch curves on each other properly scaled. Let us assume that the mixed fractal curve belongs to the same class as Koch curves. Then we may write

$$\bar{D}_k = -\frac{\log \bar{N}_k}{\log(\bar{\lambda}_k/L_0)}. \tag{11}$$

If the mixed curve is really a self-similar fractal then $\bar{D}_k$ is independent of k.

Recalling the formation law it is not difficult to calculate the number of segments $\bar{N}_k = (p_A)^i (p_B)^j$ and the segment length $\bar{\lambda}_k = L_0 q_A^{-i} q_B^{-j}$ corresponding to the k-th order term. The integers i and j are either equal or differ by one. That is,

$$\{i, j\} : [i - j = 0 \ \text{ or } \ i - j = 1].$$

Clearly $i + j = k$. This can be expressed analytically by

$$i = \lfloor (k + 1)/2 \rfloor, \quad j = \lfloor k/2 \rfloor,$$

where the $\lfloor \ \rfloor$ indicates the floor function (returning the integer part of a rational number). The expressions for N_k and λ_k can be rewritten using these expressions:

$$\log \bar{N}_k = i \log p_A + j \log p_B \tag{12}$$

and

$$\log(\bar{\lambda}_k/L_0) = -(i \log q_A + j \log q_B). \tag{13}$$

Now using (12) and (13) after some simple operations, the expression (11) becomes

$$\bar{D}_k = \frac{k \log p_A + \lfloor k/2 \rfloor (\log p_B - \log p_A)}{k \log q_A + \lfloor k/2 \rfloor (\log q_B - \log q_A)}. \tag{14}$$

Note that the dimension $\bar{D}_k$ depends on the iteration order k. So it is not possible to say that a mixed curve is fractal, further investigation is necessary. Let us separate the series O into two subsets, the odd series O_{odd} and the even series O_{even}.

For even iterations (14) reads

$$\bar{D}^{\text{even}} = \frac{\log p_A + \log p_B}{\log q_A + \log q_B}. \tag{15}$$

The sequence of even terms has a classical fractal structure with $\bar{D}^{\text{even}}$ independent of k. For odd iterations we have

$$\bar{D}_k^{\text{odd}} = \frac{(k-1)\log p_A + (k-2)\log p_B}{(k-1)\log q_A + (k-2)\log q_B} \quad \text{for } k > 1. \tag{16}$$

Then odd iterations are not strictly fractals. Clearly in the limit, when $k \to \infty$, the expression (16) tends to expression (15) and we may say that the odd sequence approaches asymptotically a fractal sequence. It is remarkable that the series O can be split into two series, O^{even} which is a fractal series and O^{odd} which is not strictly a fractal series, but we could say that it tends to a fractal object when $k \to \infty$. It is possible to generalize the result above to include several fractal curves composing a mixed fractal.

Definition. An infinite set of curves $\{C_1, C_2, \ldots, C_k, \ldots\}$ with the same basis L_0 is said to be a *quasifractal Koch sequence* if, letting $\bar{N}_k$ be the number of segments — all of same length $\bar{\lambda}_k$ — corresponding to the term C_k, the sequence $\{\bar{D}_1, \bar{D}_2, \ldots, \bar{D}_k, \ldots\}$ given by

$$\bar{D}_k = -\frac{\log \bar{N}_k}{\log(\bar{\lambda}_k/L_0)}.$$

satisfies $\bar{D}_i \neq \bar{D}_j$ and has a limit

$$\bar{D} = \lim_{k\to\infty} \bar{D}_k.$$

Suppose that the number of elementary segments and the respective lengths corresponding to a general term C_k are given by

$$\bar{N}_k = p_1^{i_1} p_2^{i_2} \cdots p_m^{i_m} \quad \text{and} \quad (\bar{\lambda}_1/L_0) = q_1^{-i_1} q_2^{-i_2} \cdots q_m^{-i_m} \quad i_1 + i_2 + \cdots + i_m = k,$$

where the pair $(p_j, L_0 q_j^{-1})$ $j = 1, 2, \ldots, m$ corresponds to a simple generator K_j belonging to the Koch family and the exponents i_α are integers function of k, $i_\alpha = f_\alpha(k)$. Then the definition of $\bar{D}_k$ gives

$$\bar{D}_k = \frac{(f_1(k)/f_\beta(k))\log p_1 + (f_2(k)/f_\beta(k))\log p_2 + \cdots + \log p_\beta + \cdots + (f_m(k)/f_\beta(k))\log p_m}{(f_1(k)/f_\beta(k))\log q_1 + (f_2(k)/f_\beta(k))\log q_2 + \cdots + \log q_\beta + \cdots + (f_m(k)/f_\beta(k))\log q_m}.$$

Now if at least one $f_j(k)/f_\beta(k) \neq 1$ and $\lim_{k\to\infty}[f_j(k)/f_\beta(k)] \to 1$ then according to the previous definition the sequence of curves C_k, $k = 1, 2, \ldots, n, \ldots$ is quasifractal. $\qquad\square$

According to this definition the odd sequence obtained with the process described above is clearly quasifractal since from (16) we immediately obtain

$$\lim_{k\to\infty} \bar{D}_k^{\text{odd}} = \frac{\log p_A + \log p_B}{\log q_A + \log q_B} = \bar{D}^{\text{even}}.$$
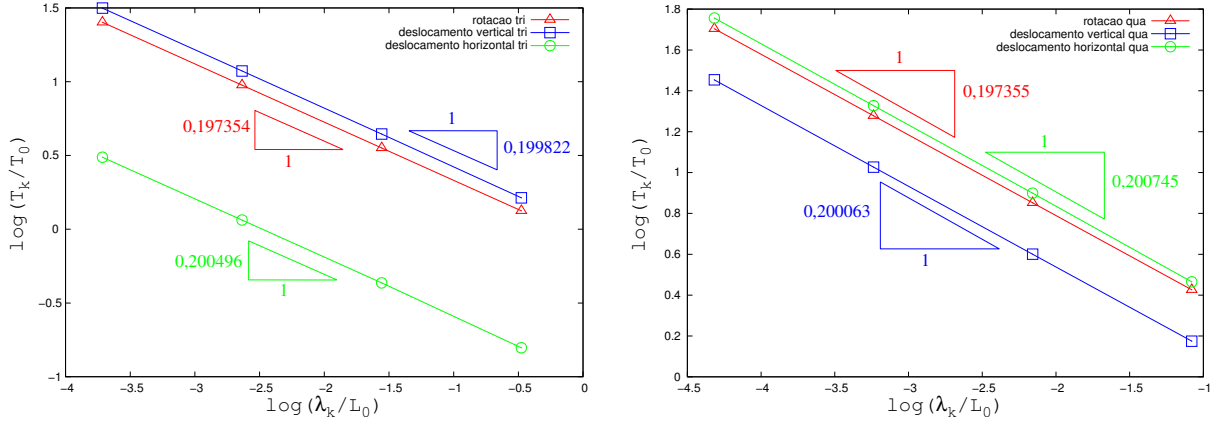
Figure 8. Normalized period $\log T_k/T_0$ versus the normalized length of segments $\log \lambda_k/L_0$ for odd (left) iterations and even (right) iterations of the triadic-quadratic mixed curve.

It is interesting to note that if q_A is close to q_B, that is, if

$$\frac{\log q_B}{\log q_A} = 1 + \varepsilon,$$

the expression (15) can be written as

$$\bar{D}_k^{\text{even}} = \frac{D_A + (1+\varepsilon)D_B}{1 + (1+\varepsilon)}.$$

When $\varepsilon \to 0$ we find that the fractal dimension tends to the mean value

$$\bar{D}_k^{\text{even}} \approx (D_A + D_B)/2.$$

3. Dynamical fractal dimension of mixed fractals

In this section we will use the technique presented in the previous section to investigate the fractal characteristic of the mixed curve presented in the Section 2. The results obtained here, despite the fact that we are dealing with a particular case, suggest that we may expect equally good results for other mixed fractals belonging to the Koch family.

3.1. *Test with a finite subset of consecutive terms removed from a mixed fractal sequence.* Let us test the dynamic fractal dimension method for mixed curves generated by the triadic-quadratic process given the first eight terms. Let eight simple harmonic oscillators be built after the geometry presented in the Figure 7. The eight terms are separated into two sets each one consisting of four terms corresponding respectively to the even sequence and the odd sequence. For the first numerical experiment the oscillators carry a constant mass and are excited by a moment, a horizontal force or a vertical force separately. The results for four terms of each series, the odd series O_{odd} and the even series O_{even} are presented in the Figure 8-a and Figure 8-b. The approximated fractal dimension D_{approx} was obtained with the slope D_{dyn} of the segment joining the points on the graph of Figure 8 corresponding to the two highest order

Exciting force	Mass M_0	Odd iterations		Even iterations	
		D_{dyn}	D_{approx}	D_{dyn}	D_{approx}
Couple	Constant	−0.197354	1.394708	−0.197355	1.394710
Vertical force		−0.199822	1.399644	−0.200063	1.400126
Horizontal force		−0.200496	1.400992	−0.200745	1.401490
Couple	Variable	−0.394709	1.394709	−0.394711	1.394711
Vertical force		−0.397130	1.397130	−0.395132	1.395132
Horizontal force		−0.398592	1.398592	−0.398576	1.398576

Table 1. Characterization using four terms of each sequence. Dynamic fractal dimension of odd and even iterations for the mixed triadic-quadratic Koch. Results for two distinct assemblages: constant mass and variable mass proportional to the curve length.

terms in the sequence. According to the theoretical predictions (9a)–(9c), those slopes should relate with the geometric fractal dimension as $D_{\mathrm{approx}} = 1 - 2D_{\mathrm{dyn}}$. The results are displayed on the Table 1 and agree satisfactorily with the correct value of the geometric fractal dimension given by (15). Note that the approximation of the fractal dimension for the odd sequence is indistinguishable from the approximation for the even fractal sequence for practical proposes.

The results for a second numerical experiment with variable mass proportional to the total spring length of the corresponding geometric term are also presented in the Table 1. For this case the fractal geometric dimension expected from the theoretical results should be related to the dynamical fractal dimension according to $D_{\mathrm{approx}} = 1 - D_{\mathrm{dyn}}$ as stated before.

It is important to mention that the oscillator periods were calculated taking into account only the elastic energy stored by the bending moment, disregarding shear and normal forces. The fractal dimension relative to the even iterations is $\bar{D}_0^{\mathrm{even}} = 1.39471$ independent of k. As shown in the Table 1, the deviation from this value is not more than 1.5%. The results are very consistent. It is also interesting to notice that the results for the odd iterations are quite satisfactory if we think of the limit value as $k \to \infty$. The numerical experiment corroborate the conjecture that the technique that has proved to work out for regular Koch curves are also efficient to calculate the dynamical fractal dimension for mixed fractal objects. We believe that this approach opens up a very rich topic for both theoretical and numerical investigation.

3.2. *Test with a single term removed from the mixed fractal sequence.* For this identification problem just one term of the reference sequence is given to find out the respective fractal dimension. Let us take the eighth order term of the Koch triadic-quadratic mixed fractal series. The projection of this curve on the horizontal axis is therefore equal to the initiator length, that is $L_n = L_0$. Figure 6 shows how the normalized periods for 14 samples obtained from the reference curve as explained in the introduction varies with respect to the successive length ratios. The cuts were made following a rather large scale $b = 1/2$ reducing each sample successively according to the rate $L_k = (L_{k-1})/2$.

Figure 9 shows the normalized periods versus the sample lengths for the three excitation types. The Table 2 shows the slope D_{dyn} of interpolated straight lines corresponding to the points on the figure representing the three excitation types and the derived approximated fractal dimension D_{approx}. According to the theoretical prediction (9a)–(9c), the relations between the geometric fractal dimension and the

Exciting force	D_{dyn}	D_{approx}
Couple	0.608081	1.397961
Vertical force	1.699156	1.398312
Horizontal force	1.710820	1.421658

Table 2. Slope of the lines which minimizes the sum of the square of the errors of the respective data.
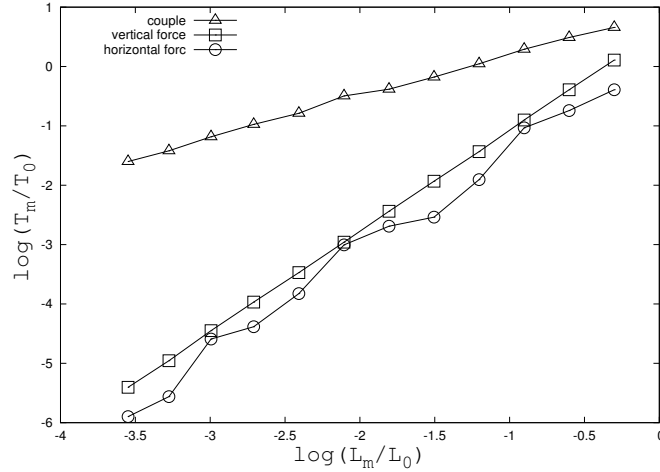
Figure 9. The normalized period $\log(T_k/T_0)$ versus the normalized horizontal projection of the length of the curve $\log(L_k/L_0)$ for 14 samples obtained from the eighth order term of the mixed triadic-quadratic Koch sequence, with $b = 1/2$.

dynamical fractal dimension should give $D_{\mathrm{approx}} = 2D_{\mathrm{dyn}}$ for the case of excitation induced by a moment and $D_{\mathrm{approx}} = 2(D_{\mathrm{dyn}} - 1)$ for the case of excitation induced by a horizontal or vertical force. The interpolated straight lines were adjusted to fit the point sets $\{T_k/T_0, L_k/L_0\}$ displayed on the Figure 6 with the least mean square deviation method.

The mass was assumed constant for all oscillators. As shown in the Table 2 the maximum error obtained with this technique is less than 3% and corresponds to the horizontal excitation. Clearly the points corresponding to the horizontal excitation presents a large dispersion. For excitations corresponding to a couple or a vertical force the errors do not exceed 1.2%.

3.3. *Random mixed structures.* It has been shown that the dynamical approach provides very good and consistent results for regular, deterministic curves belonging to the Koch family. The method has also proved to be successful in determining the fractal characteristics of mixed fractal curves.

For the classical Koch curves, besides providing means to determine the classical fractal dimension, the dynamical approach tell us if the geometry is random or not. For the regular generation process, the n-th order term is obtained from the $(n - 1)$-th term following a well determined rule. What we call random geometry is that one obtained in a similar way except that the regular rule is replaced by a random orientation procedure.
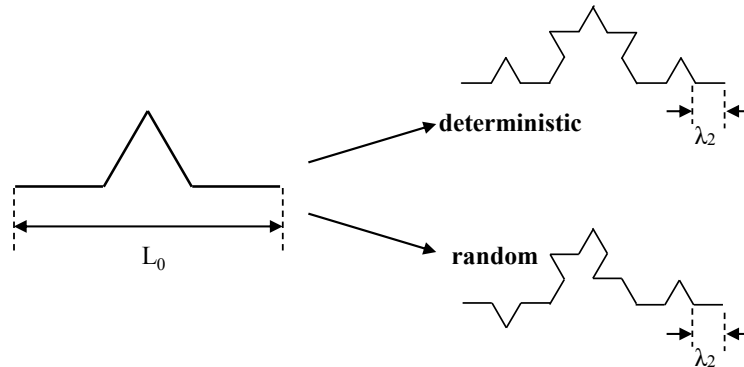
Figure 10. The first two terms of two Koch triadic curves: the classical deterministic curve (top) and a random generation (bottom). Both have the same Hausdorff dimension.

Figure 10 illustrates what we mean by a random fractal, taking as an example the classical Koch triadic. Note that the cover set for both curves is the same and therefore they have the same Hausdorff dimension. However note that the random series doesn't present the self-similarity property. Also, if the box counting technique were applied to compute the fractal dimension of these curves the results would be the same for all curves. The dynamical approach however is able to identify the random character of the curve.

Consider four terms of the triadic-quadratic odd series O^{odd} corresponding to five different random generations. Consider the corresponding simple oscillators excited by a horizontal force. The results correlating the normalized periods T_k/T_0 and the relative lengths λ_k/L_0 are displayed in Figure 11. The curves obtained by connecting the point set $\{T_k/T_0, \lambda_k/L_0\}$ for each random series clearly do not coincide. However similar construction connecting the point set derived from the excitation induced by
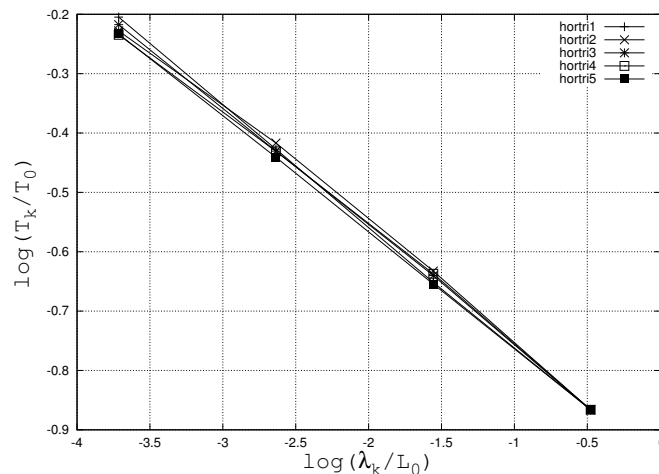
Figure 11. Logarithm of the normalized period versus the logarithm of the relative length of the elementary segment for five random triadic-quadratic sequences taking only odd iterations with initial conditions induced by a horizontal force.

a concentrated moment would lead to just one and the same curve. The reason is that for this particular type of excitation, that is, a concentrated moment, the strain energy is uniformly distributed along the curve which is the dynamical equivalent to the set covering introduced in the Hausdorff theory [Falconer 1990].

Indeed, recall that we are considering only the elastic energy induced by the bending moment. For the case of a moment this energy is independent of the orientation of the segments in the oscillator and consequently the results are the same for all random curves of the same order k. Now for excitations due to a horizontal or a vertical force the bending moment distribution depends on the position of each segment in the oscillator. This indicates that the elastic energy stored in the oscillator is sensitive to the position of the segments in the structure and this is translated in the point set distribution identifying the randomness of the structure formation. The slopes of the interpolated straight lines representing the points $\{T_k/T_0, L_k/L_0\}$ for each random series lead as expected to a unique value related closely to the Hausdorff dimension. That is on the average the result coincides with the other classical methods but the slight dispersion of the points reveals the random character of the series.

4. Conclusions

The technique proposed in a series of papers [Bevilacqua and Barros 2007; Bevilacqua et al. 2008] is applied here to mixed structures. The results reproduce the expected output of the method. All the experiments lead to the conclusion that this method is powerful and justify further exploration, encompassing both theoretical and computational fields. What the method suggests is that the dynamical properties of fractal, self-similar structures and random structures hide very rich information that need to be further investigated. In order to detect details as that associated to multi-fractals structures, random formation, mixed fractals and the like it would be convenient to use more complex oscillators for the identification problem with several added masses, that is extend the method to multi-mass systems. Possibly the frequency spectrum of those more complex structures will provide the information needed for the corresponding characterization. Application can be found in the determination of fractal dimension, if any, of protein chains, tissues and biological membranes [Bassingthwaighte et al. 1994]. It is also important to remark that this technique can be applied to physical objects, that is, characterization through laboratory experimentation. Therefore laboratory experiments can be designed to determine the dynamical properties of biological tissues or fibers and consequently the fractal characterization of the sample. The fractal characterization of composite materials may also be obtained using samples of the material to be analyzed and applying the procedure described above.

In [Bassingthwaighte et al. 1994, Chapter 12] the question is raised of "*fractals where the physical mechanism must be different at different scales*", referring to problems related to neural networks. The analysis introduced here may give some clues to explain this kind of puzzling behavior. Indeed if the mechanical properties of the oscillators are different at different scales then the sequence of periods of the oscillators may follow a power law quite different from that characterizing the geometric sequence, or even not be fractal at all. This means that for the general case the physical behavior may generate a sequence of physical properties quite different from the geometric characteristics.
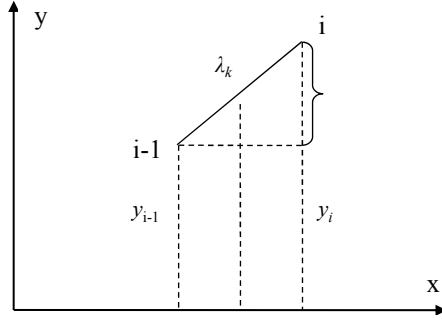
Figure 12. Mean value $\bar{y}_{i-1 \cdot i}$ and difference $\Delta y_{i-1,i}$.

Finally we would like to remark the interesting property of mixed fractals formed by two Koch curves that split up into two different sets. One is fractal and the other is quasifractal. To the best of our knowledge this is a new result that needs further investigation and generalization for structures constructed with more than two Koch curves.

Appendix

We show that the first derivative of the bilinear term Ω_k with respect to λ_k is finite for increasing values of k. Recall that

$$\Omega_k = \frac{1}{N_k} \sum_{i=0}^{N_k} \alpha_i(k) \quad \text{and} \quad \alpha_i(k) = \tfrac{1}{3}[z_{i-1}^2 + z_{i-1}z_i + z_i^2].$$

First let us write $\alpha_i(k)$ under the form

$$\alpha_i(k) = \bar{z}_{i-1,i}^2 + \tfrac{1}{3}\Delta z_{i-1,i}^2, \tag{17}$$

where

$$\bar{z}_{i-1,i} = (z_{i-1} + z_i)/2 \quad \text{and} \quad \Delta z_{i-1.i} = (z_i - z_{i-1})/2.$$

Introducing (17) in the expression for Ω_k we get

$$\Omega_k = \frac{1}{N_k} \sum_{i=0}^{N_k} (\bar{z}_{i-1,i}^2 + \tfrac{1}{3}\Delta z_{i-1,i}^2). \tag{18}$$

Clearly $\bar{z}_i \leq 1$ and $|\Delta z_{i-1,i}| \leq 1$.

Consider the second term on the right-hand side of (18). By definition $\Delta y_{i-1,i} = y_i - y_{i-1}$ as shown in Figure 12 and therefore it is possible to write

$$\Delta y_{i-1,i} = \gamma_{i-1,i}\lambda_k \quad \text{with} \quad \gamma_{i-1,i} \leq 1.$$

From this it follows that

$$\Delta z_{i-1,i} = \frac{1}{2h_0}\Delta y_{i-1,i} = \frac{1}{2h_0}\gamma_{i-1,i}\lambda_k = \beta_{i-1,i}\lambda_k.$$

Introducing this expression in (18) we get

$$\Omega_k = \frac{1}{N_k} \sum_{i=0}^{N_k} (\bar{z}_{i-1,i}^2 + \tfrac{1}{3}\beta_{i-1,i}^2 \lambda_k^2). \tag{19}$$

Now define the vector functions

$$\bar{z}_k^T = \left[\bar{z}_{0,1}^{(k)} \bar{z}_{1,2}^{(k)} \cdots \bar{z}_{i-1,i}^{(k)} \cdots \bar{z}_{N_k-1,N_k}^{(k)} \right]$$

and

$$\boldsymbol{\beta}_k^T = \frac{1}{2\sqrt{3}h_0} \left[\beta_{0,1}^{(k)} \beta_{1,2}^{(k)} \cdots \beta_{i-1,i}^{(k)} \cdots \beta_{N_k-1,N_k}^{(k)} \right].$$

Then (19) reads, in vector notation,

$$\Omega_k = \frac{1}{N_k} (z_k^T z_k + \lambda_k^2 \boldsymbol{\beta}_k^T \boldsymbol{\beta}_k).$$

Since $|z_{i-1,i}| \le 1$, $|\beta_{i-1,i}| \le 1$ and $\lambda_k \le M$ for all k, we conclude that Ω_k remains bounded as $k \to \infty$. Similarly the term of order $k+1$ can be written as

$$\Omega_{k+1} = \frac{1}{N_{k+1}} \left(\bar{z}_{k+1}^T \bar{z}_{k+1} + \tfrac{1}{3}\lambda_{k+1}^2 \boldsymbol{\beta}_{k+1}^T \boldsymbol{\beta}_{k+1} \right),$$

where the components of z_{k+1} are proportional to the ordinates of the corners of the curve corresponding to the term of order $k+1$. In general we may write

$$z_{k+1}^T = \frac{1}{h_0} \left[y_0^{(k+1)} y_1^{(k+1)} y_2^{(k+1)} \cdots y_{N_k+1}^{(k+1)} \right].$$

Referring to the preceding term in the sequence as shown in Figure 13 we have

$$z_{k+1}^T = \frac{1}{h_0} \left[y_0^{(k)} y_{01}^{(k)} \cdots y_{0(p-1)}^{(k)} y_1^{(k)} \cdots y_{1(p-1)}^{(k)} \cdots y_{N_k}^{(k)} \right]$$

Note that

$$y_{N_k}^{(k)} = y_{pN_k}^{(k+1)},$$
$$y_i^{(k)} = y_{pi}^{(k+1)},$$
$$y_{ij}^{(k)} = y_{pi+j}^{(k+1)}, \quad i = 0, \ldots, N_k-1; \; j = 1, \ldots, p-1$$

represent respectively the ordinates of the corners of the k-th curve in the sequence and the ordinates of the added corners for the $(k+1)$-th curve. It is possible then to decompose the vector z_{k+1} as

$$z_{k+1} = \frac{U_1}{h_0} \left[y_0^{(k)} y_1^{(k)} \cdots y_{N_k}^{(k)} \right]^T + \frac{U_2}{h_0} \left[y_{01}^{(k)} \cdots y_{0(p-1)}^{(k)} y_1^{(k)} \cdots y_{1(p-1)}^{(k)} \cdots y_{(N_k-1)(p-1)}^{(k)} \right],$$

where U_1 and U_2 are Boolean matrices.

Now, with this decomposition it is not difficult to show that the vector $\bar{z}_{k+1}$ can be written as

$$\bar{z}_{k+1} = R\bar{z}_k + R(\Delta z_k) + \lambda_k \boldsymbol{\rho}_{k+1}, \quad \text{where } |\rho_i^{(k+1)}| \le 1$$
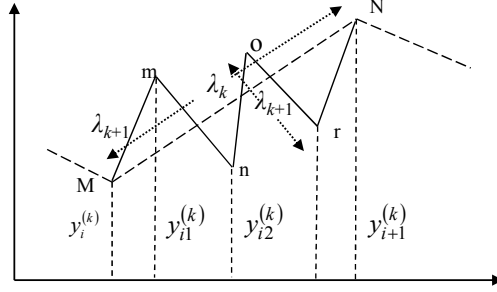
Figure 13. Term of order $k+1$ attached to the previous term of order k. $MN = \lambda_k Mm = mn = no = or = rN = \lambda_{k+1}$.

and $\boldsymbol{R} = [r_{ij}]$ is a Boolean matrix defined as follows (where i, j, p are integers):

$$r_{ij} = \begin{cases} 1 & \text{if } p(j-1) < i \leq pj, \\ 0 & \text{otherwise.} \end{cases}$$

Now using the definition of Ω_{k+1} and recalling that $(\boldsymbol{\Delta z_k}) = \lambda_k \boldsymbol{\beta}_k$ we get

$$\Omega_{k+1} = \Omega_{k+1}^* + \frac{1}{3N_{k+1}} \lambda_{k+1}^2 \boldsymbol{\beta}_{k+1}^T \boldsymbol{\beta}_{k+1},$$

where

$$\Omega_{k+1}^* = \frac{1}{N_{k+1}} (\bar{z}_k^T \boldsymbol{R}^T \boldsymbol{R} \bar{z}_k + \lambda_k^2 \boldsymbol{\beta}_k^T \boldsymbol{R}^T \boldsymbol{R} \boldsymbol{\beta}_k)$$
$$+ \frac{2}{N_{k+1}} (\lambda_k \boldsymbol{\beta}_k^T \boldsymbol{R}^T \boldsymbol{R} \bar{z}_k + \lambda_k \boldsymbol{\rho}_{k+1}^T \boldsymbol{R} \bar{z}_k + \lambda_k^2 \boldsymbol{\rho}_{k+1}^T \boldsymbol{R} \boldsymbol{\beta}_k) + \frac{1}{N_{k+1}} (\lambda_k^2 \boldsymbol{\rho}_{k+1}^T \boldsymbol{\rho}_{k+1}).$$

Recalling that

$$|z_i^{(k)}| \leq 1, \quad |\beta_i^{(k)}| \leq 1, \quad |\beta_i^{(k+1)}| \leq 1, \quad |\rho_i^{(k+1)}| \leq 1, \quad N_{k+1} = pN_k, \quad \lambda_{k+1} = \lambda_k/q$$

and that $\boldsymbol{R}$ according to the definition above has the property $\boldsymbol{R}^T \boldsymbol{R} = p\boldsymbol{I}$ we arrive at

$$\Omega_{k+1} = \frac{1}{N_{k+1}} \left(p(\bar{z}_k^T \bar{z}_k + \tfrac{1}{3}\lambda_k^2 \boldsymbol{\beta}_k^T \boldsymbol{\beta}_k) \right) + \lambda_k R_1(k, k+1) + \lambda_k^2 R_2(k, k+1),$$

where $R_1(k, k+1)$ and $R_2(k, k+1)$ are finite for all k, $\max(R_1, R_2) < M$ (finite). Finally recalling that $N_{k+1} = pN_k$ we get

$$\Omega_{k+1} = \Omega_k + \lambda_k R_1(k, k+1) + \lambda_k^2 R_2(k, k+1).$$

Now noting that

$$\Delta\lambda_k = \lambda_{k+1} - \lambda_k = \lambda_k \left(\frac{1}{q} - 1 \right).$$

We may write

$$\frac{\Delta\Omega_k}{\Delta\lambda_k} = \left(\frac{1}{q} - 1 \right)^{-1} (R_1(k, k+1) + \lambda_k R_2(k, k+1)).$$

Lemma. *For curves belonging to the Koch family — the class of curves defined by $N_k = p^k$ and $\lambda_k/L_0 = 1/q^k$ — the first order differential form of the quadratic term Ω_k with respect to λ_k is finite for increasing values of k, or equivalently decreasing values of λ_k. That is, the limit*

$$\lim_{k \to \infty} (\Delta\Omega_k/\Delta\lambda_k) = \lim_{\lambda_k \to 0} (\Delta\Omega_k/\Delta\lambda_k)$$

is finite.

References

[Bassingthwaighte et al. 1994] J. B. Bassingthwaighte, L. S. Liebovitch, and B. J. West, *Fractal physiology*, Oxford University Press, 1994.

[Bevilacqua and Barros 2007] L. Bevilacqua and M. M. Barros, "Dynamical fractal dimension: direct and inverse problems", pp. 127–136 in *IUTAM Symposium on dynamics and control of nonlinear systems with uncertainty* (Nanjing, 2006), edited by H. Y. Hu and E. Kreuzer, Springer, Dordrecht, 2007.

[Bevilacqua et al. 2008] L. Bevilacqua, M. M. Barros, and A. C. N. R. Galeão, "Geometry, dynamics and fractals", *J. Brazilian Soc. Mech. Sci. Eng.* **30** (2008), 11–21.

[Falconer 1990] K. Falconer, *Fractal geometry*, Wiley, 1990.

[Feder 1988] J. Feder, *Fractals*, Plenum Press, New York and London, 1988.

[Gouyet 1996] J. F. Gouyet, *Physics and fractal structures*, Masson, Springer, New York, 1996.

[Mandelbrot 1982] B. Mandelbrot, *The fractal geometry of nature*, Freeman, N.Y., 1982.

[Mauroy et al. 2004] B. Mauroy, M. Filoche, E. R. Weibel, and B. Sapoval, "The optimal bronchial tree is dangerous", *Nature* **427** (2004), 633–636.

LUIZ BEVILACQUA: bevilacqua@coc.ufrj.br
COPPE, Universidade Federal do Rio de Janeiro, Centro de Tecnologia, Bloco B, s/ 101, Cidade Universitária, 21945-970 Rio de Janeiro, Brazil

MARCELO M. BARROS: barros@lncc.be
Laboratório Nacional de Computação Científica, Rua Getúlio Vargas, 333, 25651-075 Petropolis, Brazil

msp

TAPPING DYNAMICS FOR A COLUMN OF PARTICLES AND BEYOND

DENIS BLACKMORE, ANTHONY ROSATO, XAVIER TRICOCHE,
KEVIN URBAN AND VISHAGAN RATNASWAMY

This paper is respectfully dedicated to Marie-Louise and Charles Steele.

The dynamics of a vertical stack of particles subject to gravity and a sequence of small, periodically applied taps is considered. First, the motion of the particles, assumed to be identical, is modeled as a system of ordinary differential equations, which is analyzed with an eye to observing connections with finite-dimensional Hamiltonian systems. Then, two approaches to obtaining approximate continuum models for large numbers of particles are described: the long-wave approximation that yields partial differential equations and the BSR method that employs integro-partial differential models. These approximate continuum models, which comprise infinite-dimensional dynamical systems, are studied with a focus on nonlinear wave type behavior, which naturally leads to investigating links to infinite-dimensional Hamiltonian systems. Several examples are solved numerically to show similarities among the solution properties of the finite-dimensional (lattice-dynamics), and the approximate long-wave and BSR continuum models. Extensions to higher dimensions and more general dynamically driven particle configurations are also sketched.

1. Introduction

There are several extant explanations of the dynamical behavior of configurations of particles subjected to small periodic tapping forces, most of which are aimed at conforming to a number of empirically derived formulas currently in use. Our primary intention is to apply fundamental principles of physics to obtain dynamical models for the evolution of systems of particles subjected to such tapping perturbations, and from them deduce simpler dynamical models - accessible to modern dynamical systems analysis — that are capable of predicting some of the most important aspects of the evolution of these systems with acceptable accuracy. Two of our main tools shall be the long-wave limit procedure, which has been successfully applied to particle dynamics by researchers such as Nesterenko, Daraio and their collaborators [Daraio et al. 2006; Nesterenko 1983; Nesterenko et al. 2005; Porter et al. 2009], and simplified integro-differential dynamical models developed in [Blackmore and Dave 1997; 1999; 2000]. In this regard, one should mention that there also is the related direct lattice-dynamics of FPU-chains approach of [Sen and Manciu 1999; 2001], which deals directly with the dynamics of long chains of particles, without going to a limit, and provides many opportunities for comparison with the results of the continuum approximation methods especially with regard to nonlinear wave type behavior of solutions. In addition, we shall briefly consider possible generalizations of these approximate continuum modeling

approaches to higher-dimensional tapping problems and even more general granular flow phenomena. Along the way, we shall indicate some of the most interesting dynamical systems aspects of the models that are introduced and analyze them from a more theoretical perspective.

We begin in Section 2 with a continuous one-dimensional dynamical model for the vertical column of particles (assuming Walton–Braun–Mindlin type interaction laws) expressed as a system of ordinary differential equations (ODEs) in the usual Newtonian way. The initial system of second order equations is converted to a system of first order equations emphasizing the comparison with the Hamiltonian system obtained for perfectly elastic collisions. In Section 3, we illustrate the dynamics with numerical solutions for a couple of cases with relatively small numbers of particles, and also deduce a theorem about the existence of chaotic dynamics.

Then, in Sections 4 and 5, we use this system in two different ways to obtain approximate continuum models for the tapping dynamics of a monodisperse column of particles in the form of infinite-dimensional dynamical systems comprised of partial differential equations (PDEs) or integro-partial differential equations (IPDEs). In particular, the PDEs are obtained using a standard long-wave limit, such as in [Daraio et al. 2006; Nesterenko 1983; Nesterenko et al. 2005; Porter et al. 2009; Zabusky and Kruskal 1965], as the number of particles goes to infinity, and then using the simplified model in [Blackmore et al. 1999]. A monodisperse system of particles is chosen in the interest of simplicity, since the dynamics of polydisperse systems are considerably more difficult to approximate using continuum models, especially for the long-wave limit method. After developing approximate continuum models for the tapping dynamics via the *long-wave* (L-W) and Blackmore–Samulyak–Rosato (BSR) methods, we also formulate the associated boundary-initial value problems corresponding to the one-dimensional vertical tapping regime under investigation in these sections, respectively.

In Section 6, we turn our attention to some of the mathematical properties of the L-W and BSR dynamics, with a focus on the Hamiltonian and near Hamiltonian infinite-dimensional dynamical systems associated with these continuum models. Some of the results for the approximate L-W and BSR models have already been proved in the literature, and some related theorems shall only be stated since their proofs, which shall be treated in forthcoming work, are a bit too technical to include in this paper. We follow this in Section 7 with numerical solutions of the L-W and BSR approximations corresponding to the numerical simulations of the tapping regimes considered in Section 3, and find good qualitative agreement among the solutions of the exact and both approximate continuum models. Finally, in Section 7, we summarize the work presented and outline some of our plans for future research — both with regard to granular flow applications and infinite-dimensional dynamical system theory - for higher-dimensional systems and more general particle configurations and forcing scenarios.

2. Newtonian model

We begin with a mathematical model for the tapping motion of the column of particles obtained using Newtonian and Hamiltonian principles assuming that the particle-particle and particle-floor interaction forces are of Walton–Braun–Mindlin type [Blackmore et al. 1999; MacKay 1999; Blackmore and Dave 1997; Daraio et al. 2006; Sen and Manciu 1999; Nesterenko 1983]. The results obtained have certain aspects in common with the approaches of the last four works just cited and of [Blackmore et al. 2000; Fermi et al. 1965; Nesterenko et al. 2005; Porter et al. 2009; Sen and Manciu 2001], among many

other works. In particular, we consider a configuration of N particles p_i, $1 \leq i \leq N$, aligned along the (vertical) positive y-axis, stacked one above the other starting with p_1, under the action of gravity — with constant gravitational acceleration g — and interacting inelastically (according to the Walton–Braun–Mindlin model) with neighboring particles and a "container" bottom, denoted as y_0, initially at the origin that is moving in a manner that simulates a periodic nearly impulsive tapping force applied vertically to the floor. The moving bottom and particle centers are located, respectively, at the points

$$0 \leq y_0(t) < y_1 < \cdots < y_N \tag{1}$$

in the semi-infinite interval $I := \{y : 0 \leq y\}$, and we assume that the particles have masses and radii $m_1, r_1, \ldots, m_N, r_N$, respectively.

In order to model the periodic tapping, we assume that $y_0(t)$ is a periodic function of period $T > 0$ represented as

$$y_0(t) := \begin{cases} a \sin \omega t & \text{if } 0 < t \leq \pi/\omega, \\ 0 & \text{if } t = 0 \text{ or } \pi/\omega \leq t \leq T \end{cases} \tag{2}$$

for $0 \leq t \leq T$, where $\pi/\omega \ll T$, and the amplitude a is a small positive number. Note that the derivative of (2) with respect to t, denoted by a dot over the variable, is the discontinuous function given as

$$\dot{y}_0(t) := \begin{cases} a\omega \cos \omega t & \text{if } 0 < t \leq \pi/\omega, \\ 0 & \text{if } t = 0 \text{ or } \pi/\omega \leq t \leq T. \end{cases} \tag{3}$$

Our intention is to determine the motion of the particles after each of a periodic sequence of taps, and compare the positions of the particles after being given time to settle following each of the taps — say at times $.99T, 1.99T, 2.99T, \ldots$, in order to determine trends after many such taps, with the evolution of the linear particle density being of particular interest.

We shall assume that the particles and bottom interact inelastically according to a simplified Walton–Braun–Mindlin law, so that the equations of motion obtained from Newton's second and third laws take the form of a system of N second-order ODEs:

$$m_i \ddot{y}_i = F_i \quad (1 \leq i \leq N), \tag{4}$$

where the forces on the particles are expressed as

$$F_i := -m_i g + f_i^{i-1} + f_i^{i+1} \tag{5}$$

for $1 \leq i \leq N$, where f_i^{i-1} is the force exerted by p_{i-1} (or the bottom when $i = 1$) on p_i and f_i^{i+1} is the force exerted by p_{i+1} on p_i when $1 \leq i \leq N-1$. These interaction forces are given by

$$\begin{aligned}
f_1^0 &:= \left(K_1^0 - \hat{K}_1^0 \sigma(\dot{y}_1 - \dot{y}_0)\right)\left(r_1 - (y_1 - y_0(t))\right)\chi\left(r_1 - (y_1 - y_0(t))\right), \\
f_i^{i+1} &:= -\left(K_i^{i+1} - \hat{K}_i^{i+1}\sigma(\Delta \dot{y}_i)\right)(r_i + r_{i+1} - \Delta y_i)\chi(r_i + r_{i+1} - \Delta y_i), \\
f_N^{N+1} &:= 0,
\end{aligned} \tag{6}$$

for $1 \leq i \leq N-1$, with

$$f_i^{i-1} = -f_{i-1}^i \tag{7}$$

for $2 \leq i \leq N$. Here

$$\Delta y_i := y_{i+1} - y_i, \quad \Delta \dot{y}_i := \dot{y}_{i+1} - \dot{y}_i, \tag{8}$$

for $1 \leq i \leq N-1$, the K_i^{i+1} and $\hat{K}_i^{i+1}$ are constants with $0 < \hat{K}_i^{i+1} \leq K_i^{i+1}$ for all $1 \leq i \leq N-1$, and σ and χ denote the signum and step functions defined, respectively, by

$$\sigma(\tau) := \begin{cases} -1 & \text{if } \tau < 0, \\ 0 & \text{if } \tau = 0, \\ 1 & \text{if } \tau > 0, \end{cases} \quad \text{and} \quad \chi(\tau) := \begin{cases} 0 & \text{if } \tau < 0, \\ \frac{1}{2} & \text{if } \tau = 0, \\ 1 & \text{if } \tau > 0. \end{cases} \tag{9}$$

With the definitions above, the system (4) may now be rewritten as

$$\ddot{y}_i = Y_i := \frac{1}{m_i} F_i = -g + \frac{1}{m_i}(f_i^{i-1} + f_i^{i+1}) \quad (1 \leq i \leq N), \tag{10}$$

which can be recast in vector form for $\boldsymbol{y} := (y_1, \ldots y_N)$ as

$$\ddot{\boldsymbol{y}} = \boldsymbol{Y}(\boldsymbol{y}, \dot{\boldsymbol{y}}, t; \boldsymbol{\mu}), \tag{11}$$

where $\boldsymbol{Y} := (Y_1, \ldots Y_N) = \left(-g + m_1^{-1}(f_1^0 + f_1^2), \ldots, -g + m_{N-1}^{-1}(f_{N-1}^{N-2} + f_{N-1}^N), -g + m_N^{-1} f_N^{N-1}\right)$, and $\boldsymbol{\mu}$ is a parameter (vector) incorporating a, ω, T, all the particle masses and radii, and all of the interaction parameters K_i^{i+1} and $\hat{K}_i^{i+1}$. As is usual for such second-order systems, we shall find it convenient to recast it as the following system of $2N$ first-order ODEs, which is better suited to direct numerical solution by such schemes as a Runge–Kutta solver:

$$\dot{y}_i = v_i, \quad \dot{v}_i = Y_i, \tag{12}$$

for $1 \leq i \leq N$. Of course, we can also write this succinctly in (the Hamiltonian related) vector form as

$$\dot{\boldsymbol{x}} = \boldsymbol{X}(\boldsymbol{x}, t; \boldsymbol{\mu}), \tag{13}$$

where

$$\boldsymbol{x} := (x_1, x_2, \ldots, x_{2N-1}, x_{2N}) := (y_1, v_1, \ldots, y_N, v_N),$$
$$\boldsymbol{X} := (X_1, X_2, \ldots, X_{2N-1}, X_N) := (v_1, Y_1, \ldots, v_N, Y_N). \tag{14}$$

Observe that only the second coordinate (component) of $\boldsymbol{X}$ depends explicitly on t, and by definition, $\boldsymbol{X}$ is periodic of period T in the independent time variable t. Also, X_2 is the only coordinate that depends explicitly on the parameters a and ω. The forces defined above are chosen so that the inelasticity of the particle-particle and particle-bottom interactions is manifested by a loss of energy upon impact that is essentially represented by a spring constant of $K_i^{i+1} + \hat{K}_i^{i+1}$ when the particles are approaching one another or the bottom and a spring constant of $K_i^{i+1} - \hat{K}_i^{i+1}$ when the particles are moving away from one another or the bottom after impact. From this perspective, the case $\hat{K}_i^{i+1} = 0$ represents a perfectly elastic interaction. We also note that the discontinuities in (13) are somewhat inconvenient from a theoretical standpoint, but they can easily be handled by a standard numerical scheme, such as a Runge–Kutta solver, and the forces can also be approximated to any degree of accuracy by smooth ($= C^\infty$) functions if necessary.

So in summary, we want to solve (13), and this can be accomplished numerically via say a variable step size Runge–Kutta scheme subject to the initial condition

$$\boldsymbol{x}(0) = (y_1(0), 0, y_2(0), \ldots, 0, y_N(0), 0), \tag{15}$$

which represents a stacked configuration of particles initially at rest. Here we have to determine the values of the $y_k(0)$ by requiring that the stack of particles is initially at rest and in equilibrium, and because the particles are assumed to be non-rigid we shall have to determine these values in a way that guarantees that $y_0(t) < y_1(t) < y_2(t) < \cdots < y_N(t)$ for all $t \geq 0$, which we shall demonstrate in the sequel for the monodisperse case. We also want to make comparisons of the positions of the particles at times, say at $.99T, 1.99T, 2.99T, \ldots, k.99T$ for rather large positive integer values of k. In these simulations it makes sense to simplify matters by taking all the $K's$ equal and all the $\hat{K}'s$ equal, with $0 \leq \hat{K} = eK$, with $0 < e < 1$, and all the radii and masses equal or having at most a pair of possible values. Note that in this context, the coefficient of restitution ϱ, which is the standard measure of elasticity with $0 \leq \varrho \leq 1$ and the left (0) and right (1) extremes representing the perfectly inelastic and elastic cases, respectively, is given as

$$\varrho^2 = \frac{1-e}{1+e}, \quad \text{or, equivalently,} \quad e = \frac{1-\varrho^2}{1+\varrho^2}. \tag{16}$$

For the other parameters, some good choices are as follows: $20 \leq N$, $T = 1$, $\pi/\omega < 1/10$, $a \leq 1/50$, $r_i = r$ and $m_i = m$, with $0 < r < 1/100$ and $0 < m < 1/20$, and it makes sense to run several cases for (13)–(15) for different parameter values satisfying these inequalities.

With the assumptions that all $K's$ and $\hat{K}'s$ are the same and $\hat{K} = eK$, the equations of motion take the form

$$\ddot{y}_i = Y_i := -g + \frac{K}{m_i}(f_i^{i-1} + f_i^{i+1}) \quad (1 \leq i \leq N), \tag{17}$$

where

$$\begin{aligned} f_1^0 &:= [1 - e\sigma(\dot{y}_1 - \dot{y}_0(t))](r_1 - y_1 + y_0(t))\chi(r_1 - y_1 + y_0(t)), \\ f_i^{i+1} &:= -[1 - e\sigma(\Delta\dot{y}_i)](r_i + r_{i+1} - \Delta y_i)\chi(r_i + r_{i+1} - \Delta y_i), \\ f_N^{N+1} &:= 0. \end{aligned} \tag{18}$$

Whence, the Newtonian equations of motion are

$$\begin{aligned} \ddot{y}_1 &= -g - \frac{K}{m_1}\Big([1 - e\sigma(\Delta\dot{y}_1)](r_1 + r_2 - \Delta y_1)\chi(r_1 + r_2 - \Delta y_1) \\ &\qquad\qquad - [1 - e\sigma(\dot{y}_1 - \dot{y}_0(t))](r_1 - y_1 + y_0(t))\chi(r_1 - y_1 + y_0(t))\Big), \\ \ddot{y}_i &= -g - \frac{K}{m_i}\Big([1 - e\sigma(\Delta\dot{y}_i)](r_i + r_{i+1} - \Delta y_i)\chi(r_i + r_{i+1} - \Delta y_i) \\ &\qquad\qquad - [1 - e\sigma(\Delta\dot{y}_{i-1})](r_{i-1} + r_i - \Delta y_{i-1})\chi(r_{i-1} + r_i - \Delta y_{i-1})\Big) \quad (2 \leq i \leq N-1), \\ \ddot{y}_N &= -g + \frac{K}{m_N}[1 - e\sigma(\Delta\dot{y}_{N-1})](r_{N-1} + r_N - \Delta y_{N-1})\chi(r_{N-1} + r_N - \Delta y_{N-1}). \end{aligned} \tag{19}$$

The natural initial conditions for the tapping dynamics are then

$$\begin{aligned} &y_1(0), y_2(0), \ldots, y_N(0) \quad \text{determined by assuming initial equilibrium of column,} \\ &\dot{y}_1(0) = \dot{y}_2(0) = \dot{y}_3(0) = \cdots = \dot{y}_N(0) = 0. \end{aligned} \tag{20}$$

It is instructive to observe that (13) can be written in the form

$$\dot{x} = \Phi(x) + \Psi(x, t; \mu), \tag{21}$$

where

$$\dot{\boldsymbol{x}} = \Phi(\boldsymbol{x}) \tag{22}$$

is an autonomous Hamiltonian system that is completely integrable in the Liouville–Arnold sense (*L-A integrable*) and Ψ may be viewed as a (non-Hamiltonian) perturbation of the integrable system (22) that goes to zero as $e, a \to 0$; see, for example, [Arnold 1978; Błaszak 1998; Guckenheimer and Holmes 1983; Katok and Hasselblatt 1995; Prykarpatsky et al. 1999]. We also note that the periodicity in t can be used to recast (21) in the autonomous form

$$\dot{\boldsymbol{x}} = \Phi(\boldsymbol{x}) + \Psi(\boldsymbol{x}, \theta; \boldsymbol{\mu}), \quad \dot{\theta} = 2\pi/T \tag{23}$$

on the $(2N+1)$-dimensional cylinder $\mathbb{R}^{2N} \times \mathbb{S}^1$, where $\mathbb{S}^1$ is the unit circle.

2.1. *Monodisperse particle configuration.*

In the interest of transparency and simplicity, we shall confine our attention to a monodisperse column of particles, so that we assume now and hereafter that $m_1 = \cdots = m_N := m$ and $r_1 = \cdots = r_N := r$. This assumption naturally simplifies each of the equations (17), (18) and (19); for example, (19) becomes

$$\ddot{y}_1 = -g - \frac{K}{m}\Big([1-e\sigma(\Delta\dot{y}_1)](2r-\Delta y_1)\chi(2r-\Delta y_1)-[1-e\sigma(\dot{y}_1-\dot{y}_0(t))](r-y_1+y_0(t))\chi\big(r-y_1+y_0(t)\big)\Big),$$

$$\ddot{y}_i = -g - \frac{K}{m}\Big([1-e\sigma(\Delta\dot{y}_i)](2r-\Delta y_i)\chi(2r-\Delta y_i)[1-e\sigma(\Delta\dot{y}_{i-1})](2r-\Delta y_{i-1})\chi(2r-\Delta y_{i-1})\Big) \tag{24}$$
$$(1 < i < N),$$

$$\ddot{y}_N = -g + \frac{K}{m}[1-e\sigma(\Delta\dot{y}_{N-1})](2r-\Delta y_{N-1})\chi(2r-\Delta y_{N-1}),$$

and the initial conditions (20) reduce to

$$y_1(0), y_2(0), \ldots, y_N(0) \text{ to be determined} \quad \text{and} \quad \dot{y}_1(0) = \cdots \dot{y}_N(0) = 0, \tag{25}$$

and (21)–(23) are modified analogously.

To determine the initial positions, we assume that the configuration is initially at rest and in equilibrium, so (24) implies that the following equations must be satisfied:

$$-2y_1(0) + y_2(0) = r + \mathfrak{M},$$
$$y_{k-1}(0) - 2y_k(0) + y_{k+1}(0) = \mathfrak{M} \quad (1 < k < N-1), \tag{26}$$
$$y_{N-1}(0) - y_N(0) = -2r + \mathfrak{M},$$

where

$$\mathfrak{M} := \frac{mg}{K}.$$

It is straightforward to show that (26) has the solution

$$y_k(0) = (2k-1)r - (k/2)(2N-k+1)\mathfrak{M} \quad (1 \le k < N-1),$$
$$y_N(0) = (2N-1)r - (1/2)[(N-1)N+2]\mathfrak{M}. \tag{27}$$

As we naturally require that $y_k(0) - y_{k-1}(0) > 0$ for all $1 \le k \le N$, this leads to the condition

$$\mathfrak{M} < \frac{r}{N}, \quad \text{or, equivalently,} \quad \frac{Ng}{r} < \frac{K}{m}, \tag{28}$$

which must be imposed in order to maintain physical realism. We note here that the maximum height S of the column of particles where they are all just touching one another is

$$S = 2Nr. \tag{29}$$

We shall study the solutions of (24)–(25) directly, and also analyze the dynamics of the corresponding first-order system (21)–(23), with the monodisperse assumption imposed to simplify the expressions. One slightly troubling feature from a theoretical standpoint is the fact is that although (20), or (24), is piecewise linear, it has jump discontinuities, which we should add can be easily handled using a numerical scheme such as the Runge–Kutta method. Theoretically, we can always adjust the systems to be smooth ($= C^\infty$) by using appropriate approximations of the signum and step functions. A rather good choice for smooth approximations is

$$\sigma(s) \cong \sigma_\alpha(s) := \tanh \alpha s, \quad \chi(s) \cong \chi_\alpha(s) := \tfrac{1}{2}(1 + \tanh \alpha s), \tag{30}$$

where $\alpha \gg 1$.

3. Dynamics of Newtonian model

In this section we shall briefly investigate the dynamical properties of the system (21) using both numerical simulation and analytical means. We begin with some numerical simulations of the solutions of (24)–(25) employing a standard ODE solver of the type used in the molecular dynamics simulation codes developed by Rosato and his collaborators, which have been applied to advantage in many studies of granular flows, such as in simulations of examples in [Blackmore et al. 1999].

For these simulations, with monodisperse particle stacks, we fix $T = 50$, $m = 1$, $r = 1.0$ cm, $a/2r = 0.53$, $e = 0.09$, $g = 980.67$ cm/sec^2, and following (29), $K = (2Ng/r)$, and vary the number of particles N over the values $N = 5$, 10. Note that the frequency $\omega = 2\pi\Gamma$ is varied for the two cases, with $\Gamma = 2.4019$ for $N = 5$ and $\Gamma = 3.8476$ for $N = 10$, mainly to obtain better resolution for the particle configurations. The dynamics for the five-particle and the ten-particle cases are illustrated in Figure 1, which plots of $y_1(t)$ through $y_N(t)$ for all choices N over a time interval long enough to include four periodic taps.
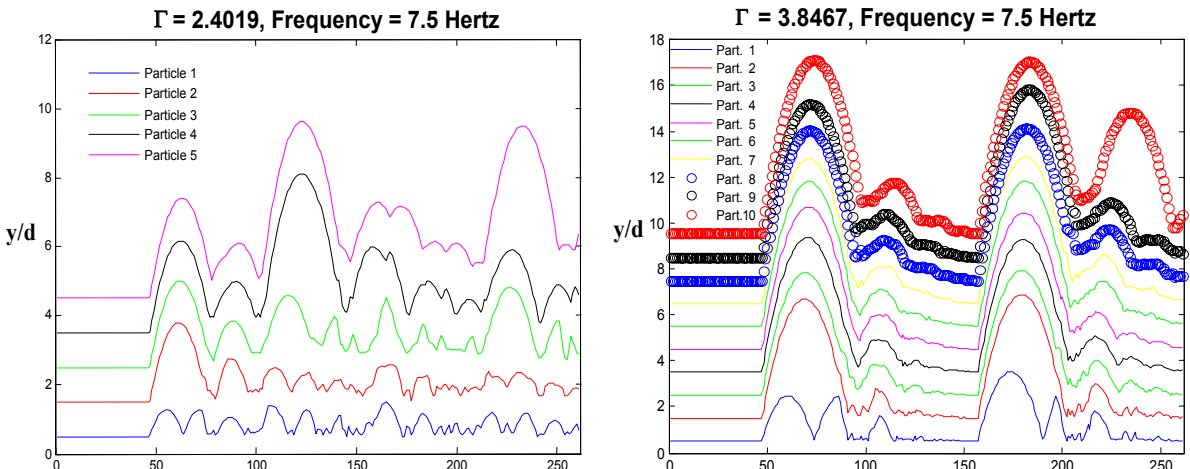

Figure 1. Particle trajectories for tapping of five (left) and ten (right) particle stacks.

Observe that in order to more clearly establish the quiescent state, the formulas for the initial conditions above are translated by 50 time steps in the numerical simulations. One can see from these time series plots for the positions of the particles how the density of the particles varies with time by observing the evolution of the interparticle distances.

From the perspective of modern dynamical systems theory, one can use the same arguments as enlisted in [Blackmore and Dave 1997; Blackmore et al. 2000] to prove the following result (see also [Guckenheimer and Holmes 1983; Prykarpatsky et al. 1999]).

Theorem 1. *Fixing all parameters save the amplitude of the taps, there exist sufficiently large values of a for which* (21) *has periodic solutions of arbitrarily large period, and ultimately there exist large values of a such that the system of particles exhibits chaotic motion. Moreover, if all the parameters are fixed but the period T of the taps, then there exist chaotic orbits when the period is sufficiently small.*

Before moving on to continuum approximations for the system (24) for very large N, we should mention a very important related body of work in the realm of lattice dynamics, which deals with a direct analysis of these equations of motion — usually focusing on finding solutions of a particular wave related form. Important representatives of this type of one-dimensional lattice dynamics research (which has yielded the existence of solitary wave solutions, solitons and breathers, to name a few examples) can be found in [Friesecke and Wattis 1994; MacKay 1999; MacKay and Aubry 1994; Poggi and Ruffo 1997; Sen and Manciu 1999; 2001; Treschev 2004; Yoshimura and Doi 2007; Zolotaryuk et al. 2000]. There are also some useful studies of the finite systems, such as in [Blackmore and Dave 1997; Blackmore et al. 2000; Yang and Wylie 2010] devoted to finding special properties of certain orbits of the dynamical systems of N particles, where N need not be particularly large.

4. Long-wave continuum approximation

We shall first use the standard long-wave continuum limit method to find an approximation for our discrete system of particles in the form of a nonlinear partial differential equation (PDE). One of the most famous examples of this approach is the continuum approximation of the Fermi–Pasta–Ulam (FPU) model [Fermi et al. 1965] obtained by Zabusky & Kruskal [Zabusky and Kruskal 1965] that established a connection between the FPU system and the Korteweg–de Vries equation that was a vital link in the chain leading to an explanation of the existence of solitons and their connection with integrable PDEs.

Toward this end, we focus on the middle equation of (24) in the monodisperse case; namely

$$\ddot{y}_i = -g - \frac{K}{m}\big([1 - e\sigma(\Delta\dot{y}_i)](2r - \Delta y_i)\chi(2r - \Delta y_i) - [1 - e\sigma(\Delta\dot{y}_{i-1})](2r - \Delta y_{i-1})\chi(2r - \Delta y_i)\big). \quad (31)$$

Next we introduce the alternative variable z for y to avoid confusion in our notation, and define $h := 2r$ to be the (mean) distance between the particles ("springs"). The underlying assumption in the long-wave approximation method is that the wavelength L of the wave phenomena embodied in the motion of the particles satisfies $h \ll L$, which as it turns out is quite a reasonable supposition for a wide range of granular flow regimes, including the one under consideration here. Our intention is to find an approximate PDE representing the dynamics of (31) as $N \to \infty$. For this we assume that y_i represents the function $y(z, t)$, which is at least of class C^4 in all of its variables, and

$$y_{i\pm1} := y(z \pm h, t), \quad (32)$$

which can be expanded in Taylor series up to the fourth order in h as

$$y_{i\pm1} := y(z \pm h, t) = y_i \pm y_z h + \tfrac{1}{2}y_{zz}h^2 \pm \tfrac{1}{6}y_{zzz}h^3 + \tfrac{1}{24}y_{zzzz}h^4 + \cdots, \tag{33}$$

and we also have

$$\dot{y}_{i\pm1} := y_t(z \pm h, t) = \dot{y}_i \pm y_{tz}h + \tfrac{1}{2}y_{tzz}h^2 \pm \tfrac{1}{6}y_{tzzz}h^3 + \tfrac{1}{24}y_{tzzzz}h^4 + \cdots. \tag{34}$$

Substituting (32), (33) and (34) in (31), we find after a long and tedious but straightforward calculation that up to the fourth order in h we have the fourth order PDE on $(z, t) \in \mathbb{R}^+ \times \mathbb{R}^+ = [0, \infty) \times [0, \infty)$:

$$y_{tt} = -g + \frac{Kh^2}{m}\Big(\tfrac{1}{3}A_{(+)}y_{zz} - \tfrac{1}{3}h[By_z + e\sigma'(0)A_{(+)}y_{tz}]y_{zz}$$
$$+ h^2\big(\tfrac{1}{36}A_{(-)}y_{zzzz} + 2r\sigma'(0)y_z^2 y_{zz} + e\sigma'(0)By_z y_{zz}y_{tz} + \tfrac{1}{6}e\sigma'(0)A_{(+)}y_{zz}y_{tzz}\big)\Big), \tag{35}$$

where

$$A_{(\pm)} := 3\chi(0) + 6r\sigma'(0) \pm 8r^3\sigma'''(0) \quad \text{and} \quad B := \sigma'(0) + 4r^2\sigma'''(0), \tag{36}$$

and we have assumed that σ and χ are convenient smooth approximations of the signum and step functions such as given in (30) and such that all even derivatives of σ and χ vanish at the origin. If we use the approximation for σ in (31), then (35) takes the form

$$y_{tt} = -g - \frac{Kh^2}{m}\Big(\tfrac{1}{3}A_{(+)}y_{zz} - \tfrac{1}{3}h[By_z + e\alpha A_{(+)}y_{tz}]y_{zz}$$
$$+ h^2\big(\tfrac{1}{36}A_{(-)}y_{zzzz} + \alpha(2ry_z^2 y_{zz} + eBy_z y_{zz}y_{tz} + \tfrac{1}{6}eA_{(+)}y_{zz}y_{tzz})\big)\Big), \tag{37}$$

with

$$A_{(\pm)} := 3\chi(0) + 6r\alpha \mp 16r^3\alpha^3 \quad \text{and} \quad B := \alpha(1 - 8r^2\alpha^2). \tag{38}$$

We can take (37) as our long-wave continuum approximation for the dynamics of the particle system as $N \to \infty$. Note that if we choose $\chi(0) = \tfrac{1}{2}$ and $\alpha = 1/h = 1/(2r)$, then

$$A_{(+)} = 2.5, \quad A_{(-)} = 13.5 \quad \text{and} \quad B = -(1/h).$$

Observe that this is a highly nonlinear wave type equation. Of course, we shall need to impose appropriate auxiliary conditions on (35) or (37) in order to obtain well-posedness, and also include the periodic external (tapping) motion given by (2) and (3). Suitable initial and boundary conditions are

$$y(z, 0) = z, \quad y(0, t) = y_0(t), \tag{39}$$

but more has to be done in order to (approximately) model the actual evolution of the system primarily due to the vertical nature of the particle stack and the influence of gravity; something that makes analysis of a horizontal particle configuration considerably easier to handle.

We shall now briefly explain these attendant difficulties, and describe the simplified method used in this paper to deal with the intricacies of numerically solving (37). This problem is defined on the semi-infinite first quadrant of the t, z-plane in a way that actually makes it analogous to a free boundary problem that also necessitates an adjustment of the right-hand side of (37) consistent with evolution of the lower and upper boundaries of the initial "particle interval", $0 \le y(z, 0) = z \le S$. What is required is to find the evolution of the top point S of the vertical continuum of particle matter, $0 \le z \le S$, defining a

curve $z = \varphi(t)$ in the in the t, z-plane, above which the force on the right-hand side of (37) has to be set to zero. In addition, the gravity term g on the right-hand side of (37) must be set equal to zero along the moving bottom $y_0(t)$, which defines the motion of the lower boundary point of the initial particle interval. We choose to deal with these mathematical difficulties completely in a forthcoming paper dedicated to just these matters, and use a much simpler hybrid method for our numerical treatment of the solution of (37) in the sequel. In particular, we use our numerical solutions of the finite-dimensional initial value problem (24)–(25) to track the (upper) "free boundary" of the evolution of the initial particle continuum of height S, and then make the adjustments in (37) that we have just indicated.

As noted earlier, we also need to impose constraints on the parameters so that the initial stack of particles, now approximated by a continuum of height S does not collapse. To this end, noting that we are assuming $h = 2r$ to be the average distance between the particles, it follows from (27) and (28) that we must have

$$\frac{Kh^2}{m} = \frac{4Kr^2}{m} > 4Ngr = 2Sg. \tag{40}$$

So, for example, $Kh^2/m = 2.5\, Sg$ would be a safe choice.

With regard to the PDEs (35) and (37), note that they represent (infinite-dimensional) integrable Hamiltonian dynamical systems when $e = 0$ (see [Ablowitz and Segur 1981; Błaszak 1998; Blackmore et al. ≥ 2011; Dickey 1991; Faddeev and Takhtajan 1987; Prykarpatsky and Mykytiuk 1998]), which is not all that surprising since they are obtained as continuum approximations of the finite-dimensional system (21), which can be rewritten as an L-A integrable Hamiltonian system when $e = 0$.

It is useful to observe that if (37) is solved imposing the auxiliary conditions described above, we can then use the continuity equation to determine the particle density ρ via the equation

$$\frac{\partial \rho}{\partial t} + y_t \frac{\partial \rho}{\partial z} + y_{tz}\rho = 0, \tag{41}$$

subject to the auxiliary (Cauchy) condition

$$\rho(z, 0) = \text{constant} = \frac{m}{2r}. \tag{42}$$

5. Alternative continuum approximation: BSR model

There is another, more direct, method for finding continuum approximations of granular flow equations that was developed in [Blackmore et al. 1999] and used in [Blackmore and Dave 1997] to approximate the dynamics of a large number of particles moving along a horizontal line with one fixed boundary (wall) and subject to a periodic forcing function. This approach leads first to the following integro-PDE (IPDE) for the dynamics of the granular motion being considered:

$u_t + uu_z =$

$$-g - \frac{K}{m} \int_0^\infty \left([1 - e\sigma(\Delta \dot{y}_i)](2r - \Delta y_i)\chi(2r - \Delta k y_i) - [1 - e\sigma(\Delta \dot{y}_{i-1})](2r - \Delta y_{i-1})\chi(2r - \Delta y_{i-1}) \right) dz, \tag{43}$$

where u is the velocity $y_t = z_t$. In order to obtain the correct expression for the integral, we need to reinterpret the local force (appearing as the integrand) on a particle at the point y in the manner of [Blackmore et al. 1999].

The actual force expression at $z \in [0, \infty)$, which must be "averaged" over a neighborhood $(z - y, z + y)$ corresponding to the near-neighbors in the lattice dynamics formulation in Section 2.1 is

$$-\frac{K}{m}[1 - e\sigma(u(z + y) - u(z))](2r - |y|)\chi(2r - |y|),$$

and this suggests the following as a first approximation for the integrated force in (42):

$$-\frac{K}{m}\int_{-h}^{h}[1 - e\sigma(u(z + y) - u(z))](2r - |y|)\chi(2r - |y|)\,dy.$$

However, this does not account for the probabilities of an abutting particle (or bottom) that actually causes the force on a particle centered at z. A reasonable choice is that of a normal distribution that depends on z in a way that the force goes rapidly to zero as $z \to \infty$ having the form

$$\Theta(y, z; b, c) := \frac{1}{b\sqrt{2\pi}}\exp\left(-\frac{((z + y) - b)^2}{2c^2}\right), \tag{44}$$

where a and b are positive constants that can be tuned to match results obtained by simulation or experiment. For the sequel, we choose $b = 7h$ and $c = 3h$. Whence, our IPDE becomes

$$u_t + uu_z = -g - \frac{K}{m}\int_{-h}^{h}\Theta(y, z; 7h, 3h)[1 - e\sigma(u(z + y) - u(z))](2r - |y|)\chi(2r - |y|)\,dy. \tag{45}$$

Again we require auxiliary conditions for well-posedness, and these take the form

$$u(z, 0) = 0, \qquad u(0, t) = \dot{y}_0(t). \tag{46}$$

Once again, it is worth noting that when $e = 0$, (45) represents an infinite-dimensional Hamiltonian dynamical system, inasmuch as it can be written in the form

$$u_t = \theta \circ \nabla H(u), \tag{47}$$

where H is a smooth function, called the Hamiltonian function, ∇ is the standard variational gradient operator for functions and θ is what is called an implectic operator, which plays a role for infinite-dimensional Hamiltonian systems analogous to that of the $2N \times 2N$ matrix

$$J = \begin{pmatrix} 0 & I_N \\ -I_N & 0 \end{pmatrix},$$

where I_N is the identity matrix of order N. These notions are all standard in the theory of Hamiltonian dynamical systems, as found in texts such as [Ablowitz and Segur 1981; Arnold 1978; Blackmore et al. $\geq$ 2011; Błaszak 1998; Dickey 1991; Faddeev and Takhtajan 1987; Guckenheimer and Holmes 1983; Katok and Hasselblatt 1995; Prykarpatsky and Mykytiuk 1998], but are a bit too mathematical to delve into here.

Observe that in a manner analogous to that in the L-W formulation above, one also can use the solution $u = (z, t)$ to determine the density via

$$\frac{\partial \rho}{\partial t} + u\frac{\partial \rho}{\partial z} + u_z\rho = 0, \tag{48}$$

subject to the Cauchy data (46).

6. Dynamics of continuum models

If, as indicated by our preliminary study here, the infinite-dimensional dynamical systems in the form of PDEs and IPDEs obtained using the L-W and BSR methods, respectively, provide good approximations to the behavior of a fairly rich variety of granular flows, it follows that we may be able to bring many of the tools of nonlinear dynamics to bear on a multitude of related problems of practical interest. Here we shall only give a couple of examples - in the form of theorems — related to wave propagation in granular media.

First, we consider the L-W approximation (37) and the possible existence of the kind of persistent stable wave structure associated with solitons, which were first described by Zabusky and Kruskal [1965] and have since become an important fixture of the modern theory of dynamical systems, as indicated by such treatments as [Ablowitz and Segur 1981; Blackmore et al. $\geq$ 2011; Błaszak 1998; Dickey 1991; Faddeev and Takhtajan 1987; Prykarpatsky and Mykytiuk 1998]. And these solitons associated with integrable systems have become increasingly important in applied research in general, and granular flow investigations in particular. Nesterenko [1983] was the first to analytically show the existence of solitons in L-W approximations of a horizontal chain of particles with perfectly elastic Hertzian interactions, which he validated by observing soliton like behavior in actual physical experiments. His equations were obtained by fourth order L-W approximations analogous to (37), except for the gravity term. He showed that his equations can be transformed into the Korteweg–de Vries equation, which is well known to be integrable and so admit soliton solutions. The transformation that he used can be readily modified for our vertical, Walton–Braun rather than Hertzian interactions, to prove the following result.

Theorem 2. *The infinite-dimensional Hamiltonian system* (37) *is integrable when* $e = 0$; *therefore it admits soliton solutions.*

Several years after Nesterenko's pioneering work on solitons in L-W approximations of one-dimensional Hertzian particle chains, Blackmore and Dave [1997] analyzed approximate models of the same kind of particle chains assuming Walton–Braun particle-particle and particle-wall interactions using the BSR method. They found that a further approximation using the BSR approach yielded Burgers' equation, which is another well-known example of an integrable PDE admitting soliton solutions. The Hamiltonian system obtained from (45) when $e = 0$ is also integrable, and this can be proved by showing that there exists another independent representation of the form (47) satisfying certain assumptions guaranteeing that the system is bi-Hamiltonian (cf. [Błaszak 1998]) in a way that implies integrability. We shall only state this result here, leaving the proof to a forthcoming paper.

Theorem 3. *If* $e = 0$, *the BSR model* (45) *is an integrable Hamiltonian dynamical system.*

7. Illustrative examples

Here we present examples of a numerical solution of the one-dimensional tapping equations (together with auxiliary data) obtained using both the L-W and BSR methods. In each case we use the simplified hybrid method — described for the L-W model in Section 4 and the BSR model in Section 5. The partial differential operators in equations (37) and (45) are approximated using central differences with equal time and distance steps of 0.01 and the integral in (45) is computed using the trapezoidal approximation.
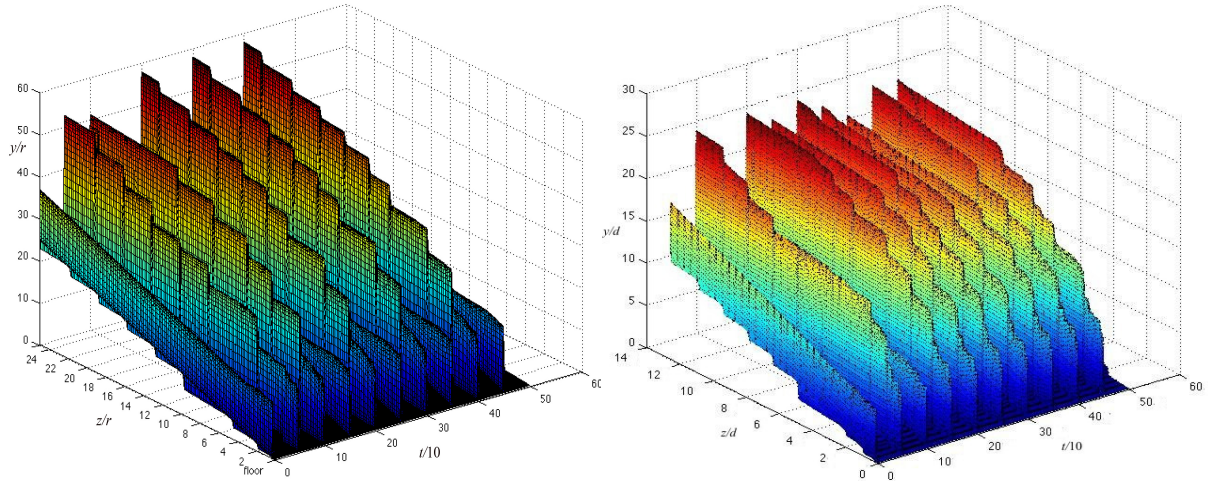
Figure 2. L-W (left) and BSR (right) approximations of tapping solution surface with parameters in Figure 1, right.

We use the same parameters as in the ten-particle example in Figure 1, right, but consider a much longer tap duration, so that we can compare the approximate numerical solutions of the continuum models to the numerical solutions of the exact Newtonian initial value problem (24)–(25).

The solution surfaces of the approximate L-W and BSR solutions are shown in the two halves of Figure 2. One can see that both approximate continuum solutions are very close to one another — both qualitatively and quantitatively — and they show good qualitative agreement with the (numerical) solution of the exact system of ODEs as depicted in Figure 1, right.

8. Concluding remarks and future research

We have shown that both the L-W and BSR approximate continuum models for the one-dimensional tapping problem have solutions that are in rather good qualitative agreement with exact lattice solutions. It seems reasonable to assume then that both of the approximate continuum approaches employed in this paper are viable predictive tools for higher-dimensional and more varied granular flows. Consequently, it just may be that a wealth of useful information can be mined from a dynamical systems oriented investigation of these relatively simple continuum models. With this in mind, we plan, in our future research, to investigate the potential of these infinite-dimensional models using a combined dynamical systems, simulation and visualization approach.

From the dynamical systems perspective, we shall first generalize the two continuum approaches to two- and three-dimensional tapping flows. Extension of the L-W method to two and three dimensions is inherently more difficult than the BSR approximation, because the lattice equations from which it is derived get quite a bit more complicated due to the increasing number of possible regular lattices (square, hexagonal, cubic, tetrahedral and more) to choose from and the increased complexity of nearest-neighbor dynamical equations, as indicated in [Ruiz-Ramírez and Macías-Díaz 2010; Zolotaryuk et al. 2000]. Naturally, we shall have to perfect our numerical treatment of the resulting boundary-initial value problems (which we conveniently simplified using a hybrid method in this paper) in order to validate

the dynamical predictions obtained from the continuum models. And having a rigorously formulated and coded numerical simulation approach will be a must as we extend the continuum approximation approach to additional types of (higher-dimensional) granular flow regimes.

We also expect the nonlinear dynamics approach to be especially useful in detecting and analyzing such phenomena as jamming and force chains. Roughly speaking, jamming can be associated with nearly invariant subregions of the dynamical systems, and force chains can be identified with connected nearly invariant networks in the resulting stress tensor fields. From a more mathematical perspective, we envisage obtaining new theoretical results linking the integrability of the limiting continuum approximations with the integrability of the exact finite systems, possibly yielding new examples of higher-dimensional integrable systems and novel KAM theory type results for systems that are only nearly integrable to begin with (see [Arnold 1978; Katok and Hasselblatt 1995; Kuksin 1993], for example) or results akin to those in [Prykarpatsky et al. 1999] when slight inelasticity renders the dynamics only nearly Hamiltonian.

Our approach using dynamical systems theory is intended to go hand-in-hand with novel refinements and extensions of existing simulation capabilities in a manner that both informs and validates the analysis. One of the applications intended for the models we have discussed in this paper is that of "density relaxation", in which a vessel of granular materials subjected to vibrations or tapping experiences an increase in its bulk density. Because the ability of granular materials to undergo this change in density is an inherent property that is not well-understood, this remains a critical impediment in developing predictive models of flowing bulk solids. We intend to address this issue in future studies by applying the dynamical systems models described in this paper to a system of uniform spheres in an attempt to determine how the various features of the motion scale with the total mass, and the details (e.g., amplitude and frequency) of the floor motion. In particular, our recent stochastic and discrete element findings on density relaxation, reported in [Dybenko et al. 2007; Rosato et al. 2010a; 2010b], have suggested the importance of the tap amplitude in the evolution of granular microstructure, as well as a strong effect of the bulk mass of the system.

An indispensable component of our future research is the kind of sophisticated dynamical systems oriented computer visualization that has come to the fore in recent year in work such as [Haller 2001; Lindeberg 1998; Tricoche et al. 2008]. As our theoretical and computational investigation moves toward more complex, higher-dimensional, and larger-scale granular systems, the interpretation and analysis of our models will become significantly more difficult and require dedicated tools that do not yet exist. We are therefore initiating a research thrust in visual data analysis aimed at supporting the characterization of salient structural and quantitative properties of the systems. Our work will explore an approach connecting the dynamical systems perspective to differential geometric structure definitions. The latter will then form the basis of abstract visual representations that address the visual clutter of large particulate assemblies.

Specifically we will apply a topological framework [Tricoche et al. 2008] to characterize the structural skeleton of the stress field acting upon the system and study its relationship with relaxation and jamming behaviors. Moreover we will study invariant manifolds of the system's dynamics by extending to the discrete setting the notion of Lagrangian coherent structures that has been recently introduced in the nonlinear dynamics and fluid mechanics literature [Haller 2001]. Finally, we will devise a novel scale space framework [Lindeberg 1998] suitable for the detection of salient core and edge manifolds in granular assemblies and explore their relationship with observed physical properties of granular flows.

References

[Ablowitz and Segur 1981] M. J. Ablowitz and H. Segur, *Solitons and the inverse scattering transform*, SIAM Studies in Applied Mathematics **4**, SIAM, Philadelphia, 1981.

[Arnold 1978] V. I. Arnold, *Mathematical methods of classical mechanics*, Graduate Texts in Mathematics **60**, Springer, New York, 1978.

[Blackmore and Dave 1997] D. Blackmore and R. Dave, "Chaos in one-dimensional granular flows with oscillating boundaries", pp. 409–412 in *Powders & grains 97* (Durham, NC, 1997), edited by R. P. Behringer and J. T. Jenkins, Balkema, Rotterdam, 1997.

[Blackmore et al. 1999] D. Blackmore, R. Samulyak, and A. Rosato, "New mathematical models for particle flow dynamics", *J. Nonlinear Math. Phys.* **6**:2 (1999), 198–221.

[Blackmore et al. 2000] D. Blackmore, R. Samulyak, R. Dave, and A. Rosato, "Dynamics of a two species oscillating particle system", pp. 255–268 in *Segregation in granular flows*, edited by A. Rosato and D. Blackmore, Kluwer, Dordrecht, 2000.

[Blackmore et al. ≥ 2011] D. Blackmore, A. Prykarpatsky, and V. Samoylenko, *Nonlinear dynamics of mathematical physics: spectral and symplectic integrability analysis*, World Scientific, Hackensack, NJ. In press.

[Błaszak 1998] M. Błaszak, *Multi-Hamiltonian theory of dynamical systems*, Springer, Berlin, 1998.

[Daraio et al. 2006] C. Daraio, V. F. Nesterenko, E. B. Herbold, and S. Jin, "Tunability of solitary wave properties in one-dimensional strongly nonlinear phononic crystals", *Phys. Rev. E* **73**:2 (2006), 026610.

[Dickey 1991] L. A. Dickey, *Soliton equations and Hamiltonian systems*, Advanced Series in Mathematical Physics **12**, World Scientific, River Edge, NJ, 1991. 2nd ed. published in 2003 as Volume 26.

[Dybenko et al. 2007] O. Dybenko, A. D. Rosato, and D. J. Horntrop, "Three-dimensional Monte Carlo simulations of density relaxation", *Kona Powder Part.* **25** (2007), 133–144.

[Faddeev and Takhtajan 1987] L. D. Faddeev and L. A. Takhtajan, *Hamiltonian methods in the theory of solitons*, Springer, Berlin, 1987.

[Fermi et al. 1965] E. Fermi, J. Pasta, and S. Ulam, *Collected works of Enrico Fermi*, vol. 2, Chicago University Press, Chicago, 1965.

[Friesecke and Wattis 1994] G. Friesecke and J. A. D. Wattis, "Existence theorem for solitary waves on lattices", *Commun. Math. Phys.* **161**:2 (1994), 391–418.

[Guckenheimer and Holmes 1983] J. Guckenheimer and P. Holmes, *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields*, Applied Mathematical Sciences **42**, Springer, New York, 1983.

[Haller 2001] G. Haller, "Distinguished material surfaces and coherent structures in three-dimensional fluid flows", *Physica D* **149**:4 (2001), 248–277.

[Katok and Hasselblatt 1995] A. Katok and B. Hasselblatt, *Introduction to the modern theory of dynamical systems*, Encyclopedia of Mathematics and its Applications **54**, Cambridge University Press, Cambridge, 1995.

[Kuksin 1993] S. B. Kuksin, *Nearly integrable infinite-dimensional Hamiltonian systems*, Lecture Notes in Mathematics **1556**, Springer, Berlin, 1993.

[Lindeberg 1998] T. Lindeberg, "Edge detection and ridge detection with automatic scale selection", *Int. J. Comput. Vis.* **30**:2 (1998), 117–156.

[MacKay 1999] R. S. MacKay, "Solitary waves in a chain of beads under Hertz contact", *Phys. Lett. A* **251**:3 (1999), 191–192.

[MacKay and Aubry 1994] R. S. MacKay and S. Aubry, "Proof of existence of breathers for time-reversible or Hamiltonian networks of weakly coupled oscillators", *Nonlinearity* **7**:6 (1994), 1623–1643.

[Nesterenko 1983] V. Nesterenko, "Propagation of nonlinear compression pulses in granular media", *Z. Prikladn. Mekh. Tekh. Fiz.* **5** (1983), 136–148. In Russian; translated in *J. Appl. Mech. Tech. Phys.* **24**:5 (1984), 733–743.

[Nesterenko et al. 2005] V. F. Nesterenko, C. Daraio, E. B. Herbold, and S. Jin, "Anomalous wave reflection at the interface of two strongly nonlinear granular media", *Phys. Rev. Lett.* **95**:15 (2005), 158702.

[Poggi and Ruffo 1997] P. Poggi and S. Ruffo, "Exact solutions in the FPU oscillator chain", *Physica D* **103**:1-4 (1997), 251–272.

[Porter et al. 2009] M. A. Porter, C. Daraio, I. Szelengowicz, E. B. Herbold, and P. G. Kevrekidis, "Highly nonlinear solitary waves in heterogeneous periodic granular media", *Physica D* **238**:6 (2009), 666–676.

[Prykarpatsky and Mykytiuk 1998] A. K. Prykarpatsky and I. V. Mykytiuk, *Algebraic integrability of nonlinear dynamical systems on manifolds: classical and quantum aspects*, Mathematics and its Applications **443**, Kluwer, Dordrecht, 1998.

[Prykarpatsky et al. 1999] Y. Prykarpatsky, A. M. Samoilenko, and D. Blackmore, "Imbeddings of integral submanifolds and associated adiabatic invariants of slowly perturbed integrable Hamiltonian systems", *Rep. Math. Phys.* **44**:1-2 (1999), 171–182.

[Rosato et al. 2010a] A. D. Rosato, O. Dybenko, D. J. Horntrop, V. Ratnaswamy, and L. Kondic, "Microstructure evolution in density relaxation by tapping", *Phys. Rev. E* **81**:6 (2010), 061301.

[Rosato et al. 2010b] A. D. Rosato, V. Ratnaswamy, D. J. Horntrop, O. Dybenko, and L. Kondic, "Temporal dynamics in density relaxation", pp. 89–99 in *IUTAM-ISIMM Symposium on Mathematical Modeling and Physical Instances of Granular Flows* (Reggio-Calabria, 2009), edited by J. D. Goddard et al., AIP Conference Proceedings **1227**, American Institute of Physics, Melville, NY, 2010.

[Ruiz-Ramírez and Macías-Díaz 2010] J. Ruiz-Ramírez and J. E. Macías-Díaz, "On the propagation of binary signals in a two-dimensional nonlinear lattice with nearest-neighbor interactions", *J. Nonlinear Math. Phys.* **17**:2 (2010), 127–136.

[Sen and Manciu 1999] S. Sen and M. Manciu, "Discrete Hertzian chains and solitons", *Physica A* **268**:3–4 (1999), 644–649.

[Sen and Manciu 2001] S. Sen and M. Manciu, "Solitary wave dynamics in generalized Hertz chains: an improved solution of the equation of motion", *Phys. Rev. E* **64**:5 (2001), 056605.

[Treschev 2004] D. Treschev, "Travelling waves in FPU lattices", *Discret. Contin. Dyn. Syst.* **11**:4 (2004), 867–880.

[Tricoche et al. 2008] X. Tricoche, G. Kindlmann, and C.-F. Westin, "Invariant crease lines for topological and structural analysis of tensor fields", *IEEE Trans. Vis. Comput. Graph.* **14**:6 (2008), 1627–1634.

[Yang and Wylie 2010] R. Yang and J. J. Wylie, "Degenerate orbit transitions in a one-dimensional inelastic particle system", *Phys. Rev. E* **82**:1 (2010), 011302.

[Yoshimura and Doi 2007] K. Yoshimura and Y. Doi, "Moving discrete breathers in nonlinear lattice: resonance and stability", *Wave Motion* **45**:1-2 (2007), 83–99.

[Zabusky and Kruskal 1965] N. J. Zabusky and M. D. Kruskal, "Interaction of 'solitons' in a collisionless plasma and the recurrence of initial states", *Phys. Rev. Lett.* **15**:6 (1965), 240–243.

[Zolotaryuk et al. 2000] A. V. Zolotaryuk, A. V. Savin, and P. L. Christiansen, "From the FPU chain to biomolecular dynamics", pp. 393–407 in *Nonlinear science at the dawn of the 21st century*, edited by P. L. Christiansen et al., Lecture Notes in Physics **542**, Springer, Berlin, 2000.

DENIS BLACKMORE: deblac@m.njit.edu
Department of Mathematical Sciences and Center for Applied Mathematics and Statistics, New Jersey Institute of Technology, University Heights, Newark, NJ 07102-1982, United States

ANTHONY ROSATO: anthony.rosato@njit.edu
Department of Mechanical Engineering, New Jersey Institute of Technology, Newark, NJ 07102-1982, United States

XAVIER TRICOCHE: xmt@purdue.edu
Department of Computer Science, Purdue University, West Lafayette, IN 47907-2107, United States

KEVIN URBAN: kdu2@njit.edu
Center for Solar-Terrestrial Research, New Jersey Institute of Technology, Newark, NJ 07102-1982, United States

VISHAGAN RATNASWAMY: vr43@njit.edu
Department of Mechanical Engineering, New Jersey Institute of Technology, Newark, NJ 07102-1982, United States

REFLECTION OF TRANSIENT PLANE STEP-STRESS WAVES: SOME CONSIDERATIONS OF ORTHOTROPY AND THERMOELASTICITY

LOUIS MILTON BROCK

To Charles and Marie-Louise Steele

Governing transient equations and (dimensionless) wave speeds for isotropic and orthotropic solids, and for isothermal and thermoelastic cases, are presented. For the orthotropic solid, propagation occurs in a principal plane. The thermoelastic cases treat Fourier heat flow and thermal relaxation, and stress- and temperature-step waves of a class that does not exhibit attenuation and dispersion. Reflection of incident step-waves by a half-space surface is then treated. Situations vary in the combinations of model features noted above. Isotropic limit cases of orthotropic results are also examined, for isothermal and thermoelastic situations. Finally, restrictions on angles of incidence and reflection due to anisotropy are identified, and some related calculations presented.

1. Introduction

The reflection of transient plane waves by traction-free surfaces of elastic solids is an important consideration in seismology [Cagniard 1962] and in models for layers and layered media [Brekhovskikh 1957; Achenbach 1973; Miklowitz 1978]. The reflection process is more complicated in the orthotropic elastic solid [Lindsay 1960] because elastic wave speeds depend on propagation direction. Similarly, isotropic thermoelastic solids [Chadwick 1960; Lord and Shulman 1960; Green and Lindsay 1972; Ignaczak and Ostoja-Starzewski 2010] exhibit waves both with and without dispersion and attenuation.

Whether isothermal [Scott and Miklowitz 1967] or thermoelastic [Sharma and Sidhu 1986], anisotropic solids are often studied in terms of the plane harmonic wave. Thus effects of step-stress (shock) signals in a reflection process are not as readily discerned as in transient analyses of, for example, isotropic isothermal solids [Achenbach 1973; Miklowitz 1978]. Plane wave propagation without dispersion can occur in an isotropic solid subject to thermal relaxation [Ignaczak and Ostoja-Starzewski 2010]. For the orthotropic solid, a class of plane waves exhibits neither dispersion nor attenuation [Brock 2010], whether Fourier conduction [Chadwick 1960] or thermal relaxation [Lord and Shulman 1960; Green and Lindsay 1972] holds. The class includes the step-stress. For plane wave propagation in an arbitrary direction, displacement and stress are in effect defined by the temperature change. For propagation in a principal plane, out-of-plane displacement uncouples from temperature, and travels as a shear wave of arbitrary form. In the isotropic limit both displacement components parallel to the wave fronts travel as shear waves of arbitrary form.

This study examines problems of reflection of a transient plane step-wave by the traction-free surface of a half-space. Two problems involve isotropic solids that are governed by the thermal relaxation

models of [Lord and Shulman 1960] and [Green and Lindsay 1972], respectively. The former exhibits a single relaxation time, while the latter has two such times. A third problem involves an orthotropic solid governed by the Fourier model [Chadwick 1960], and a fourth problem concerns an isothermal orthotropic situation. For simplicity, orthotropic problems treat plane wave propagation in a principal plane, and include the corresponding isotropic cases as limits. The incident wave in the fourth problem is a step-stress. The first three problems consider incident stress- and temperature-step waves without attenuation or dispersion [Brock 2010].

It should be noted, for the Fourier model in particular, that several problems involve restrictions on the particular combination of natural boundary conditions imposed. Nevertheless, the solutions *de facto* represent nonconventional thermoelastic processes [Ignaczak and Ostoja-Starzewski 2010], and allow insight into propagation without dispersion and attenuation for plane waves with temperature or stress steps.

As indicated above, a substantial literature exists for isothermal wave propagation. For purposes of illustration and the use of uniform definitions of parameters and functions, however, some key results for these problems are included. It is also noted that a plane wave in an infinite isothermal solid is obtained from the solution to an eigenvalue problem. That is, the solution defines the wave speeds and couples components of displacement (and also stress). The work in [Brock 2010] differs only in that a class of solutions — including the step-stress (shock) case — does not exhibit the typical attenuation and dispersion of coupled thermoelasticity.

The study begins with a presentation of governing equations for the orthotropic isothermal and the orthotropic thermoelastic solid. Corresponding equations for propagation of plane waves in a principal plane are then extracted, and characteristic wave speeds for the isothermal and thermoelastic cases examined. The isotropic limit speeds are obtained, and asymptotic formulas for the speeds when orthotropy is weak are then presented. Consideration of the four problems and isotropic limit cases follows.

2. Governing equations for orthotropic elasticity

In terms of principal Cartesian basis $\boldsymbol{x} = (x_1, x_2, x_3)$ and time t, linear momentum balance requires that

$$\partial_k T_{ki} = \bar{\mu} D^2 u_i, \quad D = \frac{\partial_t}{\bar{v}}, \quad \bar{v} = \sqrt{\bar{\mu}/\rho} \qquad (i, k = 1, 2, 3). \tag{1}$$

Here ∂_i and ∂_t signify derivatives with respect to x_i and t, and u_i, T_{ik} are the components of displacement vector $\boldsymbol{u}$ and stress tensor $\boldsymbol{T}(\boldsymbol{x}, t)$. For the isothermal case the components of $\boldsymbol{T}$ are

$$\begin{bmatrix} T_{11} \\ T_{22} \\ T_{33} \end{bmatrix} = \bar{\mu} \begin{bmatrix} C_1 & C_{12} & C_{13} \\ C_{12} & C_2 & C_{23} \\ C_{13} & C_{23} & C_3 \end{bmatrix} \begin{bmatrix} \partial_1 u_1 \\ \partial_2 u_2 \\ \partial_3 u_3 \end{bmatrix}, \tag{2a}$$

$$T_{23} = \bar{\mu} C_4 (\partial_2 u_3 + \partial_3 u_2), \quad T_{31} = \bar{\mu} C_5 (\partial_3 u_1 + \partial_1 u_3), \quad T_{12} = \bar{\mu} C_6 (\partial_1 u_2 + \partial_2 u_1), \tag{2b}$$

$$\bar{\mu} C_i = c_{ii}, \quad \bar{\mu} C_{ik} = c_{ik} \quad (i \neq k). \tag{2c}$$

Here ρ is mass density, the c_{ik} are the nine elastic constants for orthotropic elasticity [Sokolnikoff 1956; Jones 1999], $\bar{\mu}$ is a reference shear modulus chosen for convenience from the set (c_{44}, c_{55}, c_{66}), and $\bar{v}$ is a corresponding reference speed that becomes the shear (rotational) wave speed in the isotropic limit [Achenbach 1973]. A positive-definite elastic strain energy requires [Jones 1999] that the determinant

of the coefficient matrix in (2a) be positive, and that

$$C_k > 0 \qquad (k = 1, 2, 3, 4, 5, 6), \qquad (3a)$$

$$C_i C_k - C_{ik}^2 > 0 \qquad (i, k = 1, 2, 3, \ i \neq k), \qquad (3b)$$

$$|C_{ij} C_{ik} - C_i C_{jk}| < \sqrt{(C_i C_j - C_{ij}^2)(C_i C_k - C_{ik}^2)} \qquad (i, j, k = 1, 2, 3 \ \text{all distinct}). \qquad (3c)$$

The summation convention does not hold in (3c). Equations (1), (2b), (2c) and (3) hold for a thermoelastic solid initially at uniform (absolute) temperature T_0, but coupling of $(\boldsymbol{u}, \boldsymbol{T})$ with the change in absolute temperature $\theta(\boldsymbol{x}, t)$ requires that (1) be augmented by

$$(h_i \partial_i^2 - \tilde{D}D)\theta - \bar{D}D\Big(\frac{\varepsilon_1}{K_1} \partial_1 u_1 + \frac{\varepsilon_2}{K_2} \partial_2 u_2 + \frac{\varepsilon_3}{K_3} \partial_3 u_3 \Big) = 0, \qquad (4)$$

with $\bar{D}, \tilde{D}, \varepsilon_i, h_i$ defined below. The gradient column matrix in (2a) and (3) are also modified:

$$\begin{bmatrix} \partial_1 u_1 - K_1 \hat{D}\theta \\ \partial_2 u_2 - K_2 \hat{D}\theta \\ \partial_3 u_3 - K_3 \hat{D}\theta \end{bmatrix} \quad \text{with} \quad \begin{bmatrix} K_1 \\ K_2 \\ K_3 \end{bmatrix} = \begin{bmatrix} C_1 & C_{12} & C_{13} \\ C_{12} & C_2 & C_{23} \\ C_{13} & C_{23} & C_3 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}, \qquad (5a)$$

$$\alpha_i, K_i > 0 \quad (i = 1, 2, 3). \qquad (5b)$$

For the Fourier model [Chadwick 1960], denoted by F, and the thermal relaxation models of [Lord and Shulman 1960] and [Green and Lindsay 1972], denoted by I and II, respectively, the operators $\bar{D}, \tilde{D}, \hat{D}$ are

$$\text{F}: \quad \bar{D}, \tilde{D}, \hat{D} = 1, \qquad (6a)$$

$$\text{I}: \quad \bar{D}, \tilde{D} = 1 + h_{\text{I}}D, \quad \hat{D} = 1, \qquad (6b)$$

$$\text{II}: \quad \bar{D} = 1, \quad \tilde{D} = 1 + h_{\text{I}}D, \quad \hat{D} = 1 + h_{\text{II}}D. \qquad (6c)$$

Here ε_i, h_i are dimensionless thermal coupling constants and $h_{\text{I}}, h_{\text{II}}$ are thermal characteristic lengths; their expressions are

$$\varepsilon_i = \frac{T_0}{c_V}(\bar{v}K_i)^2, \quad h_i = \frac{k_i}{c_V \sqrt{\bar{\mu}\rho}}, \quad h_{\text{I}} = \bar{v}\, t_{\text{I}}, \quad h_{\text{II}} = \bar{v}\, t_{\text{II}}, \qquad (7)$$

where $t_{\text{I}}, t_{\text{II}}$ are thermal relaxation times, with $t_{\text{II}} > t_{\text{I}}$, while k_i, α_i, c_V are the conductivity, coefficient of linear expansion, and specific heat at constant volume. In the isotropic limit, $\bar{\mu} = \mu$ reduces to the shear modulus μ, while $\bar{v}$ reduces to the isotropic shear (rotational) wave speed $v_S = \sqrt{\mu/\rho}$; we then have, with λ the first Lamé constant and ν Poisson's ratio,

$$C_1, C_2, C_3 = \frac{\lambda}{\mu} + 2, \quad C_4, C_5, C_6 = 1, \quad C_{12}, C_{13}, C_{23} = \frac{\lambda}{\mu} = \frac{2\nu}{1 - 2\nu}, \qquad (8a)$$

$$K_1, K_2, K_3 = K = \Big(3\frac{\lambda}{\mu} + 2\Big)\alpha, \quad \varepsilon_1, \varepsilon_2, \varepsilon_3 = \varepsilon = \frac{T_0}{c_V}\Big[v_r\Big(3\frac{\lambda}{\mu} + 2\Big)\alpha\Big]^2, \qquad (8b)$$

$$h_1, h_2, h_3 = h = \frac{k}{c_V \sqrt{\mu\rho}}. \qquad (8c)$$

3. Plane wave propagation in arbitrary direction: formulation

The formulation for plane wave propagation in an arbitrary direction is given in [Brock 2010] for the thermoelastic case. For purposes of this article, it is sufficient to consider propagation in the $x_1 x_2$-principal plane. Thus the Cartesian basis (x, y, z) is defined with respect to the principal basis $\boldsymbol{x}$ by the transformation

$$x = x_1 \cos \phi - x_2 \sin \phi, \quad y = x_1 \sin \phi + x_2 \cos \phi, \quad z = x_3 \qquad (|\phi| < \pi/2). \tag{9}$$

Because propagation occurs in the $x_1 x_2$-plane it is convenient to choose

$$\bar{\mu} = c_{66}, \quad \bar{v} = v_6 = \sqrt{\frac{c_{66}}{\rho}}, \quad \tau = v_6 t, \quad D = \partial_\tau. \tag{10}$$

The temporal variable τ has dimensions of length, and D now signifies differentiation with respect to τ. When propagation is in the x-direction and (y, z)-dependence is suppressed, (1) gives

$$\partial_x T_{xx} = c_{66} D^2 u_x, \quad \partial_x T_{yx} = c_{66} D^2 u_y, \quad \partial_x T_{zx} = c_{66} D^2 u_z. \tag{11}$$

Similarly, in place of (2),

$$\begin{bmatrix} T_{xx} \\ T_{yx} \\ T_{zx} \end{bmatrix} = c_{66} \begin{bmatrix} C_x & C_{xy} & 0 \\ C_{xy} & C_y & 0 \\ 0 & 0 & C_{zx} \end{bmatrix} \begin{bmatrix} \partial_x u_x \\ \partial_x u_y \\ \partial_x u_z \end{bmatrix}, \tag{12a}$$

$$\begin{bmatrix} T_{yy} \\ T_{zz} \\ T_{yz} \end{bmatrix} = c_{66} \begin{bmatrix} C_y^x & C_y^y & 0 \\ C_z^x & C_z^y & 0 \\ 0 & 0 & C_{yz} \end{bmatrix} \begin{bmatrix} \partial_x u_x \\ \partial_x u_y \\ \partial_x u_z \end{bmatrix}, \tag{12b}$$

$$T_{xy} = T_{yx}, \quad T_{yz} = T_{zy}, \quad T_{xz} = T_{zx}. \tag{12c}$$

In (12a) the dimensionless coefficients are

$$C_x = C_1 \cos^4 \phi + C_2 \sin^4 \phi + (1 + \tfrac{1}{2}C_{12}) \sin^2 2\phi, \tag{13a}$$

$$C_y = \cos^2 2\phi + \tfrac{1}{4}(C_1 + C_2 - 2C_{12}) \sin^2 2\phi, \tag{13b}$$

$$C_{xy} = \tfrac{1}{2}[C_2 \sin^2 \phi - C_1 \cos^2 \phi + (2 + C_{12}) \cos 2\phi] \sin 2\phi, \tag{13c}$$

$$C_{zx} = C_5 \cos^2 \phi + C_4 \sin^2 \phi. \tag{13d}$$

In (12b) the coefficients are

$$C_y^x = C_{12}(\cos^4 \phi + \sin^4 \phi) + [\tfrac{1}{4}(C_1 + C_2) - 1] \sin^2 2\phi, \tag{14a}$$

$$C_y^y = \tfrac{1}{2}[C_2 \cos^2 \phi - C_1 \sin^2 \phi - (2 + C_{12}) \cos 2\phi] \sin 2\phi, \tag{14b}$$

$$C_z^x = C_{13} \cos^2 \phi + C_{23} \sin^2 \phi, \quad C_z^y = (C_{23} - C_{13}) \sin \phi \cos \phi, \tag{14c}$$

$$C_{yz} = (C_4 - C_5) \sin \phi \cos \phi. \tag{14d}$$

The matrix in (12a) is symmetric but that in (12b) is not, and

$$C_{xy} + C_y^y = (C_2 - C_1) \sin \phi \cos \phi, \quad C_x + C_y = 1 + C_1 \cos^2 \phi + C_2 \sin^2 \phi. \tag{15}$$

For the thermoelastic case, (11) is coupled with

$$(h_x \partial_x^2 - \tilde{D} D)\theta - D\bar{D}\partial_x \Big(\frac{\varepsilon_x}{K_x} u_x + \frac{\varepsilon_y}{K_y} u_y \Big) = 0. \tag{16}$$

The modifications of the gradient column matrices in (12a) and (12b) are

$$\begin{bmatrix} \partial_x u_x - K_x \hat{D}\theta \\ \partial_x u_y - K_{yx} \hat{D}\theta \\ \partial_x u_z \end{bmatrix}, \quad \begin{bmatrix} \partial_x u_x - K_y^x \hat{D}\theta \\ \partial_x u_y - K_3 \hat{D}\theta \\ \partial_x u_z \end{bmatrix}. \tag{17}$$

In (16) and (17),

$$h_x = h_1 \cos^2 \phi + h_2 \sin^2 \phi, \quad \varepsilon_x = \frac{T_0}{c_V}(v_6 K_x)^2, \quad \varepsilon_y = \frac{T_0}{c_V}(v_6 K_{yx})^2. \tag{18}$$

The K coefficients in (16)–(18) are given by

$$K_x = K_1 \cos^2 \phi + K_2 \sin^2 \phi, \qquad\qquad K_y^x = K_1 \sin^2 \phi + K_2 \cos^2 \phi, \tag{19a}$$

$$K_{yx} = (K_2 - K_1) \sin \phi \cos \phi, \qquad\qquad K_x + K_{yx} = K_1 + K_2. \tag{19b}$$

4. Isothermal plane waves

Studies of isothermal waves in isotropic, transversely isotropic and orthotropic solids are available in [Achenbach 1973; Payton 1983; Lindsay 1960; Scott and Miklowitz 1967]. For the sake of transparency and notational consistency, however, basic results are presented here: (11) can be uncoupled, and (12a) then gives

$$(C_{zx} \partial_x^2 - D^2) u_z = 0, \tag{20a}$$

$$\begin{bmatrix} C_x \partial_x^2 - D^2 & C_{xy} \partial_x^2 \\ C_{xy} \partial_x^2 & C_y \partial_x^2 - D^2 \end{bmatrix} \begin{bmatrix} u_x \\ u_y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{20b}$$

For plane wave propagation in the x-direction, (20) admits the general solutions

$$u_\xi = U_{\xi+} + (c_+ \tau - x) + U_{\xi-}(c_- \tau - x), \qquad \xi = (x, y), \tag{21a}$$

$$u_z = U_z(c_z \tau - x), \quad c_z = c_z(\phi) = \sqrt{C_5 \cos^2 \phi + C_4 \sin^2 \phi}. \tag{21b}$$

The functions U_z and $U_{\xi\pm}$ are arbitrary but $U_{\xi\pm}$ is subject to either of two restrictions:

$$(C_x - c_\pm^2)U_{x\pm}'' + C_{xy}U_{y\pm}'' = 0, \quad C_{xy}U_{x\pm}'' + (C_y - c_\pm^2)U_{y\pm}'' = 0 \tag{22}$$

Here a prime signifies differentiation. The dimensionless speed c_z is defined in (21b); the dimensionless speeds $c = c_\pm$ come from the roots of the secular equation

$$c^4 - (C_x + C_y)c^2 + C_x C_y - C_{xy}^2 = 0. \tag{23}$$

Properties of quadratics [Abramowitz and Stegun 1972], together with (13), (14) and (15), give

$$\begin{aligned} c_+^2 + c_-^2 &= C_x + C_y = 1 + C_1 \cos^2 \phi + C_2 \sin^2 \phi, \\ c_+^2 c_-^2 &= C_x C_y - C_{xy}^2 = C_1 \cos^2 \phi + C_2 \sin^2 \phi + \Omega \sin^2 \phi \cos^2 \phi. \end{aligned} \tag{24}$$

Solution of (23) gives

$$2c_\pm = 2c_\pm(\phi) = \Gamma^+ \pm \Gamma^-, \tag{25a}$$

$$\Gamma^\pm = \sqrt{1 + C_1 \cos^2\phi + C_2 \sin^2\phi \pm 2\sqrt{C_1 \cos^2\phi + C_2 \sin^2\phi + \Omega \sin^2\phi \cos^2\phi}}, \tag{25b}$$

$$\Omega = (C_1 - 1)(C_2 - 1) - m^2 = \gamma - C_1 - C_2, \quad \gamma = 1 + C_1 C_2 - m^2, \quad m = 1 + C_{12}. \tag{25c}$$

It is noted that $c_\pm$ and c_z are symmetric in ϕ. Parameters (γ, m) and Ω, respectively, are used by [Payton 1983] and [Brock and Georgiadis 2007] for transversely isotropic materials, but prove useful in the present study as well. In the isotropic limit $\Omega = 0$ and (21b) and (25) give

$$v_6 = v_r, \quad c_z = c_- = c_S = 1, \quad c_+ = c_D = \sqrt{\frac{\lambda}{\mu} + 2} > 1 \quad (|\phi| < \pi/2). \tag{26}$$

Subscript D and S signify quantities associated with isotropic dilatational and shear waves. For orthotropic materials, $c_+ > c_- > 0$ in any of these situations:

$$\text{A1:} \quad \gamma^2 - 4C_1 C_2 < 0, \quad |\phi| < \pi/2; \tag{27a}$$

$$\text{A2:} \quad \gamma^2 - 4C_1 C_2 > 0, \quad \gamma > 0, \quad |\phi| < \pi/2; \tag{27b}$$

$$\text{A3:} \quad \gamma^2 - 4C_1 C_2 > 0, \quad \gamma < 0, \quad |\phi| < \Phi_{A-}, \quad \Phi_{A+} < |\phi| < \pi/2, \tag{27c}$$

where

$$\Phi_{A\pm} = \tan^{-1} \frac{1}{\sqrt{2C_2}} \sqrt{-\gamma \mp \sqrt{\gamma^2 - 4C_1 C_2}}. \tag{27d}$$

If in A3 we take instead $|\phi| = \Phi_{A\pm}$, then $\Gamma^+ = \Gamma^- > 0$ and $c_+ > 0$, $c_- = 0$. If $\Phi_{A-} < |\phi| < \Phi_{A+}$, then Γ^+, Γ^- are complex conjugates with positive real parts, and c_+ is positive real, but c_- is imaginary.

5. Thermoelastic plane waves

Equation (1) again uncouples to produce (20a) and (21b). However, (u_x, u_y) and θ are coupled by (16); as a consequence [Chadwick 1960; Achenbach 1973; Ignaczak and Ostoja-Starzewski 2010] we get for isotropic materials a secular equation that gives complex dimensionless speeds c; that is, u_x, u_y, θ exhibit dispersion and exponential decay.

However one can consider a restricted class of forms [Brock 2010]

$$\theta = \sum_{k=1}^{N} \Theta_{k-1}(c\tau - x)^{k-1} \quad (c\tau \geq x) \tag{28}$$

with $N = 4$ for F and $N = 3$ for I, II. Here the Θ_k are constants; the Θ_0 term represents a propagating temperature step. The secular equations then become

$$\text{F:} \quad c^4 - I_1^F c^2 + I_2^F = 0, \tag{29a}$$

$$\text{I:} \quad c^6 - (l_x + I_1^F)c^4 + [l_x(c_+^2 + c_-^2) + I_2^F]c^2 - l_x c_+^2 c_-^2 = 0, \tag{29b}$$

$$\text{II:} \quad c^6 - (l_x + \bar{I}_1)c^4 + [l_x(c_+^2 + c_-^2) + \bar{I}_2]c^2 - l_x c_+^2 c_-^2 = 0. \tag{29c}$$

The coefficients in (29a) are

$$I_1^F = 1 + F_1 \cos^2 \phi + F_2 \sin^2 \phi, \quad I_2^F = F_1 \cos^2 \phi + F_2 \sin^2 \phi + \Omega_F \sin^2 \phi \cos^2 \phi, \tag{30a}$$

with

$$\Omega_F = (F_1 - 1)(F_2 - 1) - m_F^2 = \gamma_F - F_1 - F_2, \quad \gamma_F = 1 + F_1 F_2 - m_F^2, \quad m_F = 1 + F_{12}, \tag{30b}$$

$$F_1 = C_1 + \varepsilon_1, \quad F_2 = C_2 + \varepsilon_2, \quad F_{12} = C_{12} + \sqrt{\varepsilon_1 \varepsilon_2}. \tag{30c}$$

In (29b) and (29c), c_+ and c_- are given by (25), and

$$\bar{I}_1 = l I_1^F + (1 - l)(c_+^2 + c_-^2), \quad \bar{I}_2 = l I_2^F + (1 - l) c_+^2 c_-^2, \quad l_x = \frac{h_x}{h_I} > 1, \quad l = \frac{h_{II}}{h_I} > 1. \tag{31}$$

The inequalities in (31) are based on data [Ignaczak and Ostoja-Starzewski 2010; Brock 2009]. The roots of (29a), (29b) and (29c) give dimensionless speeds $c = c_{F\pm}$, $c = c_{1I}, c_{2I}, c_{3I}$ and $c = c_{1II}, c_{2II}, c_{3II}$. Properties of quadratic and cubic equations [Abramowitz and Stegun 1972] show that

$$c_{F+}^2 + c_{F-}^2 = I_1^F, \quad c_{F+}^2 c_{F-}^2 = I_2^F, \tag{32a}$$

$$c_{1I}^2 + c_{2I}^2 + c_{3I}^2 = l_x + I_1^F, \quad c_{1II}^2 + c_{2II}^2 + c_{3II}^2 = l_x + \bar{I}_1, \tag{32b}$$

$$c_{1I}^2 c_{2I}^2 + c_{2I}^2 c_{3I}^2 + c_{3I}^2 c_{1I}^2 = l_x I_1 + I_2^F, \quad c_{1II}^2 c_{2II}^2 + c_{2II}^2 c_{3II}^2 + c_{3II}^2 c_{1II}^2 = l_x I_1 + \bar{I}_2, \tag{32c}$$

$$c_{1I}^2 c_{2I}^2 c_{3I}^2 = c_{1II}^2 c_{2II}^2 c_{3II}^2 = l_x I_2. \tag{32d}$$

For model F, (29a) and (30) give symmetric real functions of ϕ:

$$2c_{F\pm} = 2c_{F\pm}(\phi) = \Gamma_F^+ \pm \Gamma_F^-, \tag{33a}$$

$$\Gamma_F^\pm = \sqrt{1 + F_1 \cos^2 \phi + F_2 \sin^2 \phi \pm 2\sqrt{F_1 \cos^2 \phi + F_2 \sin^2 \phi + \Omega_F \sin^2 \phi \cos^2 \phi}}. \tag{33b}$$

In the isotropic limit, Ω_F vanishes and (30) and (33) give, for all $|\phi| < \pi/2$,

$$v_6 = v_r, \quad c_z = c_{F-} = c_S = 1, \quad c_{F+} = c_F = \sqrt{c_D^2 + \varepsilon} > 1. \tag{34}$$

The parameter set $(\gamma_F, m_F, \Omega_F)$ is the thermoelastic counterpart of (γ, m, Ω). For the orthotropic material, the behavior of $c_{F\pm}$ is governed by conditions that can be obtained from (27) in Section 4 by replacing $c_\pm$, $\Gamma^\pm$, γ, m, Ω, C_1, C_2 with $c_{F\pm}$, $\Gamma_F^\pm$, γ_F, m_F, Ω_F, F_1, F_2, respectively.

Formal expressions for the roots c_{1I}^2, c_{2I}^2, c_{3I}^2 of (29b) and basic inequalities that guarantee positive real values are given in [Brock 2010] for Model I. The isotropic case is of interest here, and it can be shown that (29b) gives, for all $|\phi| < \pi/2$,

$$v_6 = v_S, \quad c_{1I} = c_{I+}, \quad c_{2I} = c_{I-}, \quad c_{3I} = c_z = c_S = 1, \tag{35a}$$

$$2c_{I\pm} = \Gamma_I^+ \pm \Gamma_I^-, \quad \Gamma_I^\pm = \sqrt{(c_D \pm \sqrt{l_x})^2 + \varepsilon}. \tag{35b}$$

Available data [Brock 2009] suggest that $1 < c_{I-} < c_D < c_F < c_{I+}$. Equation (35a) shows that the components of u_x, u_y corresponding to c_{3I} uncouple from θ, i.e., are shear waves defined by arbitrary functions of $\tau - x$.

Equation (29c) for Model II yields roots $c_z, c_{1\text{II}}, c_{2\text{II}}, c_{3\text{II}}$. Formal expressions for these, and basic inequalities that guarantee positive real values, are given in [Brock 2010]. As with Model I the isotropic case is of interest here:

$$v_6 = v_S, \quad c_{1\text{II}} = c_{\text{II}+}, \quad c_{2\text{II}} = c_{\text{II}-}, \quad c_{3\text{II}} = c_z = c_S = 1, \tag{36a}$$

$$2c_{\text{II}\pm} = \Gamma_{\text{II}}^+ \pm \Gamma_{\text{II}}^-, \quad \Gamma_{\text{II}} = \sqrt{(c_D \pm \sqrt{l_x})^2 + l\varepsilon} \tag{36b}$$

Available data [Brock 2009] suggest that $1 < c_{\text{II}-} < c_D < c_F < c_{\text{II}+}$.

The components of u_x, u_y that correspond to $c_{F-}, c_{3\text{I}}, c_{3\text{II}}$ are seen from equations (34), (35a), (36a), respectively, to uncouple from θ, and become shear waves that are arbitrary functions of $\tau - x$. Moreover (35) and (36) give the dimensionless speeds in transient two-dimensional studies that are valid for short times ($\tau/h_\text{I} \ll 1$ and $\tau/h_\text{II} \ll 1$ respectively); see [Brock 2009].

6. Speeds in weakly orthotropic solids

The limit (26) of (25) can be defined in terms of dimensionless parameters C_1 and Ω, and likewise the limit (34) of (33) in terms of F_1 and Ω_F:

$$C_1 \to C_2, \quad \Omega \to 0: \quad c_+(\phi) = c_D, \quad c_-(\phi) = 1, \tag{37a}$$

$$F_1 \to F_2, \quad \Omega_F \to 0: \quad c_{F+}(\phi) = c_F, \quad c_{F-}(\phi) = 1, \tag{37b}$$

In similar fashion, results from [Brock 2010] show that

$$C_1 \to C_2, \quad F_1 \to F_2, \quad h_1 \to h_2, \quad \Omega, \Omega_F \to 0: \begin{cases} c_{1k\text{I}}(\phi) = c_{\text{I}+}, \quad c_{2\text{I}}(\phi) = c_{\text{I}-}, \quad c_{3\text{I}}(\phi) = 1, & (38a) \\ c_{1\text{II}}(\phi) = c_{\text{II}+}, \quad c_{2\text{II}}(\phi) = c_{\text{II}-}, \quad c_{3\text{II}}(\phi) = 1, & (38b) \end{cases}$$

When orthotropy is weak, i.e., when $C_1, F_1, h_1, \Omega, \Omega_F$ are close to the limits indicated in (37) and (38), asymptotic formulas can be derived. The derivation does not require explicit formulas such as (25a) and (33a). A first-order variation of (24) and (29), under constraints (23) and (32), gives

$$c_+(\phi) \approx c_D + \frac{\delta C_1}{2c_D} \cos^2 \phi - \frac{\Omega}{8mc_D} \sin^2 2\phi, \qquad c_-(\phi) \approx 1 + \frac{\Omega}{8m} \sin^2 2\phi, \tag{39a}$$

$$c_{F+}(\phi) \approx c_F + \frac{\delta F_1}{2c_F} \cos^2 \phi - \frac{\Omega_F}{8m_F c_F} \sin^2 2\phi, \qquad c_{F-}(\phi) \approx 1 + \frac{\Omega_F}{8m_F c_F} \sin^2 2\phi, \tag{39b}$$

$$[c_{1\text{I}}(\phi), c_{2\text{I}}(\phi)] \approx c_{\text{I}\pm} + \frac{1}{2c_{\text{I}\pm}} \frac{\cos^2 \phi}{P_\text{I} c_{\text{I}\pm}^2 + Q_\text{I}} (l_0\Omega - c_{\text{I}\pm}^2 \Omega_F) \sin^2 \phi + (c_{\text{I}\pm}^2 - 1)\left[c_{\text{I}\pm}^2 \delta F_1 - l_0 \delta C_1 + (c_{\text{I}\pm}^2 - c_D^2)\frac{\delta h_1}{h_\text{I}}\right], \tag{39c}$$

$$[c_{1\text{II}}(\phi), c_{2\text{II}}(\phi)] \approx c_{\text{II}\pm} + \frac{1}{2c_{\text{II}\pm}} \frac{\cos^2 \phi}{P_\text{II} c_{\text{II}\pm}^2 + Q_\text{II}} \times [\Omega(l_0 + c_{\text{II}\pm}^2 (l - 1)) - l c_{\text{II}\pm}^2 \Omega_F] \sin^2 \phi$$

$$+ (c_{\text{II}\pm}^2 - 1)\left[l c_{\text{II}\pm}^2 \delta F_1 + (c_{\text{II}\pm}^2 (1 - l) - l_0) \delta C_1 + (c_{\text{II}\pm}^2 - c_D^2)\frac{\delta h_1}{h_\text{I}}\right], \tag{39d}$$

$$c_{3\mathrm{I}}(\phi) \approx 1 + \frac{l_0\Omega - \Omega_{\mathrm{F}}}{1 - c_{\mathrm{F}}^2 + ml_0} \frac{1}{8} \sin^2 2\phi, \tag{39e}$$

$$c_{3\mathrm{II}}(\phi) \approx 1 + \frac{(l_0 + l - 1)\Omega - l\Omega_{\mathrm{F}}}{1 - lc_{\mathrm{F}}^2 + ml_0 + 2c_D^2(l - 1)} \frac{1}{8} \sin^2 2\phi, \tag{39f}$$

where $l_0 = h/h_{\mathrm{I}}$ is the limit value of l_x, the parameters $P_{\mathrm{I}}, P_{\mathrm{II}}, Q_{\mathrm{I}}, Q_{\mathrm{II}}$ are defined by

$$P_{\mathrm{I}} = 1 + c_{\mathrm{F}}^2 - 2l_0, \quad P_{\mathrm{II}} = 1 + c_D^2 + l\varepsilon - 2l_0, \tag{40a}$$

$$Q_{\mathrm{I}} = 3l_0^2 + l_0[c_{\mathrm{F}}^2 + 2(\varepsilon - 1)] - 2c_{\mathrm{F}}^2, \tag{40b}$$

$$Q_{\mathrm{II}} = 3l_0[l_0 + \varepsilon(l - 1)] + l_0[c_{\mathrm{F}}^2 + 2(\varepsilon - 1)] - 2(c_D^2 + l\varepsilon), \tag{40c}$$

and $\delta C_1, \delta F_1, \delta h_1, \Omega, \Omega_{\mathrm{F}}$ are small:

$$\delta C_1 = C_1 - C_2, \quad \delta F_1 = F_1 - F_2, \quad \delta h_1 = h_1 - h_2, \tag{41a}$$

$$\left|\frac{C_1}{C_2} - 1\right| \ll 1, \quad \left|\frac{F_1}{F_2} - 1\right| \ll 1, \quad \left|\frac{h_1}{h_2} - 1\right| \ll 1, \quad \left|\frac{\gamma}{C_1 + C_2} - 1\right| \ll 1, \quad \left|\frac{\gamma_{\mathrm{F}}}{F_1 + F_2} - 1\right| \ll 1. \tag{41b}$$

7. Isotropic case: shear wave reflection for model I

The isotropic half-space $x_1 > 0$ is initially at rest at uniform temperature T_0. An incident plane shear step-wave travels in the negative x-direction, and reaches surface point $(x_1, x_2) = 0$ at time $t = 0$ ($\tau = 0$):

$$(u_x, u_z) = 0, \quad T_{xy} = G_i + G_i'(\tau + x)(\tau + x \geq 0). \tag{42}$$

Here G_i, G_i' are given constants. In view of (9), the wave (42) generates, for $x_1 = 0$ and $\tau + x_2 \sin\phi \geq 0$,

$$\theta = 0, \quad (T_{12}, T_{11}) = (\cos 2\phi, -\sin 2\phi)[G_i + G_i'(\tau + x_2 \sin\phi)] \tag{43}$$

Reflection of (43) generates plane waves governed by (11), (28), (29b) and (34). These travel away from the surface, i.e., in positive x-directions whose angles differ from ϕ. In view of (9), (12) and (17), therefore, we have, for $x_1 = 0$ and $\tau - x_2 \sin\phi_S \geq 0$,

$$\theta = 0, \quad (T_{12}, T_{11}) = (\cos 2\phi_S, -\sin 2\phi_S)[G_S + G_S'(\tau - x_2 \sin\phi_S)]. \tag{44}$$

For $x_1 = 0$, $c_{\mathrm{I}\pm}\tau - x_2 \sin\theta_\pm \geq 0$ we have

$$\theta = \Theta_\pm + \Theta_\pm'(c_{\mathrm{I}\pm}\tau - x_2 \sin\phi_\pm), \tag{45a}$$

$$(T_{12}, T_{11}) = -\frac{\mu K}{d_\pm}(\sin 2\phi_\pm, C_\pm)[\Theta_\pm + \Theta_\pm'(c_{\mathrm{I}\pm}\tau - x_2 \sin\phi_\pm)], \tag{45b}$$

$$C_\pm = c_{\mathrm{I}\pm}^2 - 2\sin^2\phi_\pm, \quad d_\pm = c_{\mathrm{I}\pm}^2 - c_D^2, \quad d_+ d_- = -\varepsilon c_D^2. \tag{45c}$$

Here $G_S, G_S', \Theta_\pm, \Theta_\pm'$ are unknown constants. The half-space surface remains traction-free and is governed by thermal convection [Chadwick 1960]:

$$x_1 = 0: \quad T_{12} = T_{11} = 0, \quad \partial_1\theta - \beta\theta = 0 \tag{46}$$

Parameter β is related to the Biot number [Boley and Weiner 1985]. Satisfaction of (46) by the summation of (43)–(45) requires that

$$\phi_S = -\phi, \quad \phi_\pm = -\sin^{-1}(c_{I\pm} \sin \phi) \quad (c_{I\pm} \sin \phi \leq 1). \tag{47}$$

It follows that (46) produces the equations

$$\Theta'_\pm = \pm \frac{\beta}{\Delta}(\Theta_+ + \Theta_-), \quad \Delta = \cos \phi_+ - \cos \phi_-, \tag{48a}$$

$$\mu K \left(\frac{\Theta_+}{d_+} \sin 2\phi_+ + \frac{\Theta_-}{d_-} \sin 2\phi_- \right) + G_S \cos 2\phi_S + G_i \cos 2\phi = 0, \tag{48b}$$

$$-\mu K \left(\frac{\Theta_+}{d_+} C_+ + \frac{\Theta_-}{d_-} C_- \right) + G_S \sin 2\phi_S - G_i \sin 2\phi = 0, \tag{48c}$$

$$\mu K \frac{\beta}{\Delta} \left(\frac{c_{I+}}{d_+} \sin 2\phi_+ - \frac{c_{I-}}{d_-} \sin 2\phi_- \right)(\Theta_+ + \Theta_-) + G'_S \cos 2\phi_S + G'_i \cos 2\phi = 0, \tag{48d}$$

$$-\mu K \frac{\beta}{\Delta} \left(\frac{c_{I+}}{d_+} C_+ - \frac{c_{I-}}{d_-} C_- \right)(\Theta_+ + \Theta_-) + G'_S \sin 2\phi_S - G'_i \sin 2\phi = 0. \tag{48e}$$

Equation (45a) implies a spike (Dirac) function $\delta(\tau + x_2 \sin \phi)$ at the wave intersection in the heat flux term in (46). Therefore (48) is subject to the restriction

$$c_{I-}\Theta_+ \cos \phi_+ + c_{I+}\Theta_- \cos \phi_- = 0. \tag{49}$$

The equation set (48b)–(48e) is solved for $(G_S, G'_S, \Theta_\pm)$, whereupon (48a) yields $\Theta'_\pm$. The surface temperature change generated by reflection is of particular interest, and it can be shown that, for $x_1 = 0$ and $\tau + x_2 \sin \phi \geq 0$,

$$\theta = \frac{\varepsilon c_D^2}{\mu K} \frac{G'_i N_{2I} \sin 4\phi}{R_{2I} R_{3I}} \left[\frac{\Delta}{\beta} + (c_{I+} - c_{I-})(\tau + x_2 \sin \phi) \right], \tag{50a}$$

$$R_{kI} = d_- c_{I+}^k R_+ - d_+ c_{I-}^k R_-, \quad N_{kI} = d_- c_{I+}^k R_+ + d_+ c_{I-}^k R_- \quad (k = 1, 2, 3), \tag{50b}$$

$$R_\pm = 2 \sin 2\phi \sin \phi \sqrt{s_{I\pm}^2 - \sin^2 \phi} - \cos^2 2\phi, \quad s_{I\pm} = \frac{1}{c_{I\pm}}. \tag{50c}$$

The function $R_\pm$ is of the Rayleigh type [Achenbach 1973] in the isothermal case, and R_{2I} is a thermo-elastic counterpart. Investigation of possible roots of functions (N_{kI}, R_{kI}) is beyond the scope of this article, but is necessary to complete this analysis.

Equation (50a) represents the effects of heat production from a shear wave by mode conversion. It is noted that the step-stress term G_i does not contribute to surface temperature. Moreover, restriction (49) is satisfied only when the incident shear wave parameters (G_i, G'_i) are related by

$$\left[\beta \frac{G_i}{G'_i}(c_{I+} \cos \phi_- - c_{I-} \cos \phi_+) + 1 - \cos \phi_+ \cos \phi_- \right] R_{3I} - c_{I+}^2 c_{I-}^2 R_{1I} \sin^2 \phi = 0 \tag{51}$$

That is, the surface in general exhibits a spike in the heat flux.

The restriction in (47) indicates that reflections moving at speed $c_{I-} v_S$ travel parallel to the surface when $c_{I+} \sin \phi = 1$. Therefore $\phi = \sin^{-1} s_{I+}$ is the minimum grazing angle of incidence. For this angle

$\cos \phi_+ = 0$ and (50) gives

$$R_+ = -(1 - 2s_{I+}^2)^2, \quad R_- = 4s_{I+}^2\sqrt{1 - s_{I+}^2}\sqrt{s_{I-}^2 - s_{I+}^2} - (1 - 2s_{I+}^2)^2. \tag{52}$$

Use of (52) causes distinct changes in the forms of (50a) and (51). However, this does not give correspondingly distinctive behavior.

8. Isotropic case: shear wave reflection for model II

Equations (29c) and (34) now govern. Thus (42)–(44) and (45a) hold, but with $c_{I\pm}$ replaced by $c_{II\pm}$. In place of (45b) and (45c) we have, for $x_1 = 0$ and $c_{II\pm}\tau - x_2 \sin \phi_\pm \geq 0$,

$$(T_{12}, T_{11}) = -\frac{\mu K}{d_\pm}(\sin 2\phi_\pm, C_\pm)[\Theta_\pm + \Theta'_\pm(h_{II}c_{II\pm} + c_{II\pm}\tau - x_2 \sin \phi_\pm)], \tag{53a}$$

$$d_\pm = c_{II\pm}^2 - c_D^2, \quad d_+d_- = -\varepsilon\frac{h_{II}}{h_I}c_D^2 \tag{53b}$$

Equations (48a), (48d), (48e) and (49) still hold, with $c_{I\pm}$ replaced by $c_{II\pm}$. Equations (47), (48b) and (48c) are replaced by

$$\phi_r = -\phi, \quad \phi_\pm = -\sin^{-1}(c_{II\pm}\sin\phi)(c_{II+}\sin\phi \leq 1), \tag{54a}$$

$$\mu K\left(p_+\frac{\Theta_+}{d_+}\sin 2\phi_+ + p_-\frac{\Theta_-}{d_-}\sin 2\phi_-\right) + G_S\cos 2\phi_S + G_i\cos 2\phi = 0, \tag{54b}$$

$$-\mu K\left(p_+\frac{\Theta_+}{d_+}C_+ + p_-\frac{\Theta_-}{d_-}C_-\right) + G_S\sin 2\phi_S - G'_i\sin 2\phi = 0, \tag{54c}$$

$$p_+ = 1 + \frac{\beta}{\Delta}c_{II+}h_{II}, \quad p_- = 1 - \frac{\beta}{\Delta}c_{II-}h_{II}. \tag{54d}$$

In this case, for $x_1 = 0$ and $\tau + x_2 \sin \phi \geq 0$, we have

$$\theta = \frac{\varepsilon c_D^2}{\mu K}\frac{h_{II}}{h_I}\frac{G'_iN_{2II}\sin 4\phi}{R_{2II}R_{3II}}\left[\frac{\Delta}{\beta} + (c_{II+} - c_{II-})(\tau + x_2 \sin \phi)\right], \tag{55a}$$

$$R_{kII} = d_-p_+c_{II+}^kR_+ - d_+p_-c_{II-}^kR_-, \quad N_{kII} = d_-p_+c_{II+}^kR_+ + d_+p_-c_{II-}^kR_- \quad (k = 1, 2, 3). \tag{55b}$$

Equation (50c) still holds, with $c_{I\pm}$ replaced by $c_{II\pm}$. The counterpart to (51) does not hold and a surface heat flux spike will arise unless

$$\beta\frac{G_i}{G'_i}(c_{II-}\cos\phi_+ - c_{II+}\cos\phi_-)R_{3I} + (1 - \cos\phi_+\cos\phi_-)R_{3II} - c_{II+}^2c_{II-}^2R_{1II}\sin^2\phi = 0. \tag{56}$$

Here R_{3I}, R_{3II} both appear, with $c_{I\pm}$ replaced by $c_{II\pm}$ in (50b). Completion of this analysis will require study of possible roots of N_{kII}, R_{kII}. Equation (54a) shows that the minimum grazing angle is $\phi = \sin^{-1} s_{II+}$. Then (52) holds, with s_{I+} replaced by s_{II+}. Equations (55a) and (56) do not exhibit distinctive behavior for this angle.

9. Orthotropic case: thermal wave reflection for model F

Consider the temperature change $\theta = \Theta_i$, traveling as a plane step-wave with speed $c_{F+}v_6$ in the negative x-direction toward the surface $x_1 = 0$ of half-space $x_1 > 0$. Arrival at surface point $(x_1, x_2) = 0$ occurs at time $t = 0$ ($\tau = 0$). The material is orthotropic and satisfies the Fourier model equations, and (33a) for $c_{F\pm} = c_{F\pm}(\phi)$ in particular. This problem is treated in [Brock 2010], so only key steps are presented:

For $x_1 = 0$, $c_{F+}(\phi)\tau + x_2 \sin \phi \geq 0$ the incident wave generates

$$\theta = \Theta_i, \quad T_{12} = -c_{66}\frac{Q(c_{F+}, \phi)}{S(c_{F+}, \phi)}\Theta_i, \quad T_{11} = -c_{66}\frac{P(c_{F+}, \phi)}{S(c_{F+}, \phi)}\Theta_i \tag{57}$$

In (57) the functions Q, P are defined by

$$Q(c, \phi) = (K_1 + K_2)c^2(\phi)\sin\phi\cos\phi + Q_{12}\sin^2\phi(1 + 2\cos 2\phi) + Q_{21}\cos^2\phi(1 - 2\cos 2\phi), \tag{58a}$$

$$Q_{12} = K_1 m \cos^2\phi - K_2(C_1\cos^2\phi + \sin^2\phi), \tag{58b}$$

$$Q_{21} = K_2 m \sin^2\phi - K_1(C_2\sin^2\phi + \cos^2\phi), \tag{58c}$$

$$P(c, \phi) = [F_1 K_1\cos^2\phi + (m_F - 1)K_2\sin^2\phi]c^2(\phi) + [F_1 P_1\cos^2\phi + (m_F - 1)P_2\sin^2\phi]\sin\phi\cos\phi, \tag{59a}$$

$$P_1 = K_1(m\cos 2\phi + 1 + 2C_2\sin^2\phi) - K_2(C_1\cos 2\phi + 2m\sin^2\phi), \tag{59b}$$

$$P_2 = K_2(m\cos 2\phi - 1 - 2C_1\cos^2\phi) - K_1(C_2\cos 2\phi - 2m\cos^2\phi). \tag{59c}$$

The function S is defined by

$$
\begin{aligned}
&S(c, \phi) \\
&= [c^2(\phi) - 1](\varepsilon_1\cos^2\phi + \varepsilon_2\sin^2\phi) + [(C_2 - 1)\sqrt{\varepsilon_1} - m\sqrt{\varepsilon_2}]\sqrt{\varepsilon_1} + [(C_1 - 1)\sqrt{\varepsilon_2} - m\sqrt{\varepsilon_1}]\sqrt{\varepsilon_2}. \tag{60}
\end{aligned}
$$

Here only the condition that the half-space surface remains traction-free is imposed. In accordance with (28) and (29a), reflected plane waves travel in positive x-directions that form angles $\phi_\pm$ with the positive x_1-axis at speed $c_{F\pm}(\phi_\pm)v_6$. The temperature steps are $\Theta_\pm$, Equations (58)–(60) hold with ϕ replaced by $\phi_\pm$, and for $x_1 = 0$ and $c_{F\pm}(\phi_\pm)\tau - x_2\sin\phi_\pm \geq 0$ we have

$$\theta = \Theta_\pm, \quad T_{12} = -c_{66}\frac{Q(c_{F\pm}, \phi_\pm)}{S(c_{F\pm}, \phi_\pm)}\Theta_\pm, \quad T_{11} = -c_{66}\frac{P(c_{F\pm}, \phi_\pm)}{S(c_{F\pm}, \phi_\pm)}\Theta_\pm. \tag{61}$$

Reflection requires that

$$c_{F+}(\phi)\sin\phi_+ + c_{F+}(\phi_+)\sin\phi = 0, \quad c_{F+}(\phi)\sin\phi_- + c_{F-}(\phi_-)\sin\phi = 0. \tag{62}$$

This equation is satisfied when

$$\phi_+ = -\phi, \quad \phi_- = -\sin^{-1}\frac{\sin\phi}{c_{12}^F(\phi)}, \tag{63a}$$

$$c_{12}^F(\phi) = \frac{1}{\sqrt{F_1}}\sqrt{(1 + F_1)c_{F+}^2(\phi) - F_1\cos^2\phi - (F_2 + \Omega_F)\sin^2\phi}. \tag{63b}$$

Imposing a traction-free surface gives

$$\Theta_+ = -\Theta_i \frac{Q(c_{F+}, \phi)P(c_{F-}, \phi_-) - P(c_{F+}, \phi)Q(c_{F-}, \phi_-)}{Q(c_{F+}, -\phi)P(c_{F-}, \phi_-) - P(c_{F+}, -\phi)Q(c_{F-}, \phi_-)}, \tag{64a}$$

$$\Theta_- = -\Theta_i \frac{S(c_{F-}, \phi_-)}{S(c_{F+}, \phi)} \frac{P(c_{F+}, \phi)Q(c_{F+}, -\phi) - Q(c_{F+}, \phi)P(c_{F+}, -\phi)}{Q(c_{F+}, -\phi)P(c_{F-}, \phi_-) - P(c_{F+}, -\phi)Q(c_{F-}, \phi_-)}. \tag{64b}$$

It is noted that $\Theta_i + \Theta_+ + \Theta_- \neq 0$, i.e., the Fourier model predicts surface temperature change. Moreover a surface heat flux spike occurs in $\partial_1\theta$ unless

$$(\Theta_i - \Theta_+)\cos\phi - \Theta_-\cos\phi_- = 0. \tag{65}$$

The nature of c_{12}^F depends on material categorization that differs from the counterpart to (27) mentioned in Section 5, e.g., $\gamma_F < 0$ implies $\Omega_F < 0$, and $\Omega_F - m_F^2 > 0$ implies $\Omega_F > 0$. Therefore (63b) gives the positive real results

$$c_{12}^F(\Phi_{A\pm}) = \frac{1}{\sqrt{F_1}}\sqrt{1 + F_1\cos^2\Phi_{A\pm} + F_2\sin^2\Phi_{A\pm}}, \tag{66a}$$

$$c_{12}^F(\Phi_{B\pm}) = \sqrt{F_1\cos^2\Phi_{B\pm} + (F_2 + \Omega_F\cos^2\Phi_{B\pm})\sin^2\Phi_{B\pm}}. \tag{66b}$$

However, when $(F_1 - F_2)^2 - 4\Omega_F < 0$ and $(1 + F_1)(F_1 - F_2) - 2(F_2 + \Omega_F) > 0$, the value of c_{12}^F vanishes when $|\phi| = \Phi_\pm$ and is imaginary for $\Phi_- < |\phi| < \Phi_+$, where

$$\Phi_\pm = \tan^{-1}\sqrt{\frac{F_1}{2}}\sqrt{(1 + F_1)[F_1 - F_2 \pm \sqrt{(F_1 - F_2)^2 - 4\Omega_F}] - 2(F_2 + \Omega_F)} \tag{67}$$

This behavior implies that (63a) is subject to the restriction $\sin^2\phi < [c_{12}^F(\phi)]^2$. It can be shown that the restriction is satisfied except in the following cases:

$$B1: \quad Q_B > 0, \quad \Phi_{B1} < |\phi| < \pi/2 \tag{68a}$$

$$B2: \quad P_B^2 + Q_B > 0, \quad P_B < 0, \quad Q_B < 0, \quad \Phi_{2B}^- < |\phi| < \Phi_{2B}^+, \tag{68b}$$

with

$$\Phi_{1B} = \tan^{-1}\frac{1}{\sqrt{Q_B}}\sqrt{P_B + \sqrt{P_B^2 + Q_B}}, \qquad \Phi_{2B}^\pm = \tan^{-1}\frac{1}{\sqrt{-Q_B}}\sqrt{-P_B \mp \sqrt{P_B^2 + Q_B}}, \tag{68c}$$

$$P_B = 1 - \frac{F_1}{2}(1 + F_1)(1 + F_2), \qquad\qquad Q_B = F_1 F_2 + \Omega_F(1 + F_2). \tag{68d}$$

For angles of incidence ϕ that lie outside of the ranges prescribed by B1 and B2, a reflected wave travels in the negative x_2-direction at speed $c_{F+}v_6$. In view of (68) and the thermoelastic counterpart to (27) in Section 4, (61) and (64) are governed by two cases. A study of the various limits $|\phi| = \Phi_{A\pm}$, Φ_{1B}, $\Phi_{2B}^\pm$, as well as study of possible situations for which (61) and (64) vanish or become unbounded, is beyond the scope of this single paper. Such efforts are planned for a longer format.

As observed above the reflections uncouple as a thermal and a shear wave in the isotropic limit. Thus for $x_1 = 0$ the incident and reflected fields give

for $c_F\tau + x_2 \sin\phi \geq 0$: $\theta = \Theta_i,$ $T_{12} = -\dfrac{\mu K Q(\phi)}{m_F\varepsilon}\Theta_i,$ $T_{11} = -\dfrac{\mu K P(\phi)}{m_F\varepsilon}\Theta_i;$ (69a)

for $c_F\tau - x_2 \sin\phi_F \geq 0$: $\theta = \Theta_F,$ $T_{12} = -\dfrac{\mu K Q(\phi_F)}{m_F\varepsilon}\Theta_F,$ $T_{11} = -\dfrac{\mu K P(\phi_F)}{m_F\varepsilon}\Theta_F;$ (69b)

for $\tau - x_2 \sin\phi_S \geq 0$: $\theta = 0,$ $T_{12} = T_S \cos 2\phi_S,$ $T_{11} = T_S \sin 2\phi_S.$ (69c)

Here T_S is the unknown stress T_{xy} due to the reflected shear wave, and

$$Q(\phi) = c_F^2 \sin 2\phi - 1 + \cos^2 2\phi, \tag{70a}$$

$$P(\phi) = c_F^2(c_F^2 - 2\sin^2\phi) + 8\sin^3\phi\,\cos^3\phi. \tag{70b}$$

A traction-free surface now requires that

$$\phi_F = -\phi, \quad \phi_S = -\sin^{-1}\frac{\sin\phi}{c_F}, \tag{71a}$$

$$T_S = \frac{2\mu K\Theta_i}{m_F D_0\varepsilon}[4(1 - 2\cos^2 2\phi) + c_F^2(c_F^2 - 2\sin^2\phi)]\sin 2\phi, \tag{71b}$$

$$\Theta_F = \Theta_i\Big[-1 + \frac{4}{D_0}(c_F^2\sin\phi_S\cos\phi_S - \sin^2 2\phi\cos 2\phi_S)\sin 2\phi\Big], \tag{71c}$$

$$D_0 = Q(-\phi)\sin 2\phi_S - P(-\phi)\cos 2\phi_S. \tag{71d}$$

On the surface we have $\Theta_F + \Theta_i \neq 0$, and a heat flux spike occurs unless

$$(\Theta_i - \Theta_F)\cos\phi = 0. \tag{72}$$

The parameter $c_{12}(\phi)$ equals c_F when $F_1 = F_2$ and $\Omega_F = 0$, and for the weakly orthotropic case

$$c_{12}^F(\phi) \approx c_F + (1 + \cos^2\phi)\frac{\delta F_1}{2c_F} - \frac{c_F}{m_F}\Omega_F\sin^2\phi. \tag{73}$$

Here δF_1, Ω_F are governed by (41).

10. Orthotropic isothermal case: stress wave reflection

Consider the situation in Section 9, except that material response is isothermal, and an incident plane wave moves with speed $c_+(\phi)v_6$ toward the half-space surface. In terms of the traction T_i, the wave is defined in accordance with (20)–(22) as step-stresses

$T_{xx} = T_i c_+^2(\phi)[\cos^2 2\phi + (C_1 + C_2 - 2C_{12})\sin^2\phi\cos^2\phi - c_+^2(\phi)],$ (74a)

$T_{xy} = -T_i c_+^2(\phi)[C_2\sin^2\phi - C_1\cos^2\phi + (2 + C_{12})\cos 2\phi]\sin\phi\,\cos\phi,$ (74b)

$T_{yy} = T_i c_+^2(\phi)[C_{12}\cos^2 2\phi + (C_1 + C_2 + 2C_{12} - 4)\sin^2\phi\cos^2\phi] - T_i(C_{12} + \Omega\sin^2\phi\cos^2\phi).$ (74c)

For $x_1 = 0$ and $c_+(\phi)\tau + x_2\sin\phi \geq 0$, the step-stresses T_{xx}, T_{xy}, T_{yy} generate traction

$$T_{11} = T_i P(c_+, \phi), \quad T_{12} = T_i Q(c_+, \phi), \tag{75}$$

where P, Q are defined by

$$P(c, \phi) = c^2(\phi)[C_{12}\cos^2 2\phi + 2C_2\sin^2\phi - C_1\cos 2\phi - 8\sin^2\phi]\sin^2\phi$$
$$+ (C_1\cos^2\phi + C_2\sin^2\phi)\cos^2\phi + \Omega\cos 2\phi \sin^2\phi\cos^2\phi - C_{12}\sin^2\phi, \quad (76\text{a})$$

$$Q(c, \phi) = -c^2(\phi)[(C_1 + C_2)\sin^2\phi\cos^2\phi + C_{12}(\cos^4\phi + \sin^4\phi) + \cos^2 2\phi]\sin 2\phi$$
$$+ [C_{12} + C_1\cos^2\phi + C_2\sin^2\phi + 2\Omega\sin^2\phi\cos^2\phi]\sin\phi\cos\phi. \quad (76\text{b})$$

(Thus P is symmetric and Q is antisymmetric in ϕ.) Reflection generates plane waves T_{xx}, T_{xy}, T_{yy} that travel away from the surface with speed $c_\pm(\phi_\pm)v_6$, so that for $x_1 = 0$ and $c_\pm(\phi)\tau - x_2\sin\phi_\pm \geq 0$ we have

$$T_{11} = T_\pm P(c_\pm, \phi_\pm), \quad T_{12} = T_\pm Q(c_\pm, \phi_\pm). \quad (77)$$

In this case a traction-free surface requires that

$$\phi_+ = -\phi, \quad \phi_- = -\sin^{-1}\frac{\sin\phi}{c_{12}(\phi)}, \quad (78\text{a})$$

$$T_+ = T_i \frac{Q(c_+, \phi)P(c_-, \phi_-) - Q(c_-, \phi_-)P(c_+, \phi)}{Q(c_+, \phi)P(c_-, \phi_-) + Q(c_-, \phi_-)P(c_+, \phi)}, \quad (78\text{b})$$

$$T_- = -2T_i \frac{Q(c_+, \phi)P(c_+, \phi)}{Q(c_+, \phi)P(c_-, \phi_-) + Q(c_-, \phi_-)P(c_+, \phi)}, \quad (78\text{c})$$

$$c_{12}(\phi) = \frac{1}{\sqrt{C_1}}\sqrt{(1 + C_1)c_+^2(\phi) - C_1\cos^2\phi - (C_2 + \Omega)\sin^2\phi}. \quad (78\text{d})$$

The behavior of (78d) is analogous to that of (63b), so that (78a) must be subject to the restriction $\sin^2\phi < c_{12}^2(\phi)$. Condition (68) in Section 9 again holds, but with $(c_{12}^F, \Omega_F, F_1, F_2)$ replaced by $(c_{12}, \Omega, C_1, C_2)$. Consistent with the observations in Sections 7, 8 and 9, study of (78a) in light of (27), (68) and their analogues, and consideration of cases for which (78b) and (78c) vanish or become unbounded, is reserved for future work.

However, for some insight into both (27) in Section 4 and the isothermal analogue to (68) given in Section 9, Table 2 in the Appendix provides calculations for four orthotropic wood materials (see Table 1) under isothermal conditions. The table entries show that $c_\pm$ exists for these materials, i.e., there are no angles $\Phi_{A\pm}$ that restrict angle of incidence ϕ. Similarly, the entries show that only the isothermal counterpart of restriction B1 governs reflection; that is, Φ_{B1} exists, but $\Phi_{2B}^\pm$ does not.

The isotropic limit case is a standard problem [Achenbach 1973]. However, for completeness some results are presented here:

$$\phi_D = -\phi, \quad \phi_S = -\sin^{-1}\frac{\sin\phi}{c_D}, \quad T_D = T_i\frac{N}{R}, \quad T_S = -\frac{T_i}{R}\sin 2\phi(c_D^2 - 2\sin^2\phi), \quad (79\text{a})$$

$$N = 2\sin 2\phi\sin\phi\sqrt{c_D^2 - \sin^2\phi} - (c_D^2 - 2\sin^2\phi)^2, \quad (79\text{b})$$

$$R = 2\sin 2\phi\sin\phi\sqrt{c_D^2 - \sin^2\phi} + (c_D^2 - 2\sin^2\phi)^2. \quad (79\text{c})$$

The isothermal counterpart to (73) is

$$c_{12}(\phi) \approx c_D + (1 + \cos^2 \phi) \frac{\delta C_1}{2c_D} - \frac{c_D}{m} \Omega \sin^2 \phi. \tag{80}$$

Here δC_1, Ω are governed by (41).

11. Some observations

The results of Sections 7 and 8 illustrate the alteration in half-space temperature that occurs when an isothermal (shear) step-stress wave is reflected from its surface, and only the wave class studied in [Brock 2010] is considered. In Section 7 thermal relaxation with a single relaxation time [Lord and Shulman 1960] governs. The [Green and Lindsay 1972] model, with an additional thermal relaxation time, governs in Section 8, and this time is coupled with the convection parameter β in the solution. A minimum grazing angle of incidence arises in both Sections, but distinctive changes in solution behavior do not seem to occur at this angle.

Section 9 treats an incident temperature step-wave that propagates without dispersion or attenuation [Brock 2010] in an orthotropic half-space governed by the Fourier law [Chadwick 1960]. Only two corresponding reflection waves arise, so that only the requirement of a traction-free surface is met. Formulas for reflection angles are also presented. In the isotropic limit one reflection becomes an isothermal shear wave, and an asymptotic reflection angle formula valid for weak orthotropy is given.

Section 10 involves the commonly studied isothermal wave reflection process in an orthotropic half-space. For comparison with Section 9 the incident wave is a step-stress. In the isotropic limit the two waves generated by reflection reduce to the standard dilatational/shear wave pair. As in Section 9, distinctive behavior does not seem to occur at the minimum grazing angle. The reflection angle, its isotropic limit, and an asymptotic form for weak orthotropy are also given.

These results can in general be predicted by work in Sections 4 and 5. There governing equations and associated (dimensionless) speeds for plane wave propagation in a principal plane of isothermal and thermoelastic orthotropic solids are examined. Only the wave class discussed in [Brock 2010] is treated in the latter instance. Isotropic limits for the dimensionless wave speeds are also given, as well as asymptotic formulas for weakly orthotropic solids.

Sections 7, 8 and 9 demonstrate the limited applicability of thermoelastic plane wave ensembles that travel without dispersion or attenuation [Brock 2010], both for isotropic and orthotropic solids. Because the Fourier model allows only two signals, stress-free surfaces can result only if a prescribed uniform temperature and associated heat flux spite form the thermal boundary conditions. General time-harmonic [Chadwick 1960] or transient [Brock 2005; Brock and Hanson 2006] studies admit three, so the "missing" signal corresponds to the Fourier paradox of infinite speed.

However, Sections 7, 8 and 9 also demonstrate that there are combinations of speed, wave profile and angle of incidence that do not couple prescribed stress and thermal boundary conditions. As implied at the outset, moreover, problems that do couple conditions represent in effect special cases of nonconventional thermoelastic processes [Ignaczak and Ostoja-Starzewski 2010]. Such boundary conditions are artificial. Nevertheless, the problems discussed here do illustrate that the transient response of thermoelastic solids subject to surface reflection can differ from that described by analyses based on time-harmonic waves, dispersion and attenuation.

Finally, it is noted again that the restrictions imposed on wave travel by orthotropic elastic solids, whether thermoelastic or isothermal, need to be examined in more detail. The work by Kraut [1963] and Payton [1983] in isothermal transversely isotropic cases, and by Ignaczak and Ostoja-Starzewski [2010] in isotropic thermoelasticity with relaxation, are models in this regard. At present work is also proceeding on thermoelastic plane waves that exhibit a particular form of attenuation, but without dispersion.

Appendix

Data for four orthotropic wood materials — balsa (B), yellow birch (YB), Douglas fir (DF) and Sitka spruce (SS) — are taken from [Crandall and Dahl 1959, pp. 224–228] and summarized here:

	S_{22}/S_{11}	S_{33}/S_{11}	S_{12}/S_{11}	S_{23}/S_{11}	S_{13}/S_{11}	S_{44}/S_{11}	S_{55}/S_{11}	S_{66}/S_{11}
B	20	70	−0.3	−15	−0.5	18	200	27
DF	13	20	−0.5	−9	−0.5	14	60	15
YB	15	20	−0.4	−7	−0.5	16	140	13
SS	13	23	−0.4	−6	−0.5	16	20	16

Table 1. Compliance ratios.

The dimensionless constants C_1, C_2, C_{12} can be obtained from [Jones 1999]:

$$C_1 = \frac{S_{66}}{S_{11}s}\left[\frac{S_{22}}{S_{11}}\frac{S_{33}}{S_{11}} - \left(\frac{S_{23}}{S_{11}}\right)^2\right], \quad C_2 = \frac{S_{66}}{S_{11}s}\left[\frac{S_{33}}{S_{11}} - \left(\frac{S_{13}}{S_{11}}\right)^2\right], \quad C_{12} = \frac{S_{66}}{S_{11}s}\left[\frac{S_{13}}{S_{11}}\frac{S_{23}}{S_{11}} - \frac{S_{33}}{S_{11}}\frac{S_{12}}{S_{11}}\right],$$

$$s = \frac{S_{22}}{S_{11}}\frac{S_{33}}{S_{11}} - \left(\frac{S_{23}}{S_{11}}\right)^2 - \frac{S_{22}}{S_{11}}\left(\frac{S_{13}}{S_{11}}\right)^2 - \frac{S_{33}}{S_{11}}\left(\frac{S_{12}}{S_{11}}\right)^2 + 2\frac{S_{12}}{S_{11}}\frac{S_{13}}{S_{11}}\frac{S_{23}}{S_{11}}.$$

Equations (25), (27) and the isothermal analogue of (68) then give results in Table 2 for the various dimensionless parameters.

References

[Abramowitz and Stegun 1972] M. A. Abramowitz and I. A. Stegun, *Handbook of mathematical functions*, Dover, New York, 1972.

[Achenbach 1973] J. D. Achenbach, *Wave propagation in elastic solids*, North-Holland, Amsterdam, 1973.

[Boley and Weiner 1985] B. A. Boley and J. H. Weiner, *Theory of thermal stresses*, Krieger, Malabar, FL, 1985.

	C_1	C_2	m	Ω	$c_\pm$ ($\phi = 15°$)		$c_\pm$ ($\phi = 45°$)		$c_\pm$ ($\phi = 75°$)		Φ_{1B}
B	27.37	1.63	4.67	−5.32	5.07	0.99	3.82	0.95	1.87	0.93	1.27°
DF	16.15	1.78	2.31	6.49	3.898	0.996	3.004	0.97	1.42	0.94	2.07°
YB	13.53	1.06	1.62	−1.82	3.56	1.02	2.65	1.13	1.57	1.13	2.85°
SS	16.59	1.43	2.16	−2.64	3.95	0.99	3.05	0.84	1.598	0.95	2.15°

Table 2. Dimensionless parameters. The conditions for existence of $\Phi_{A\pm}$ and $\Phi_{2B}^{\pm}$ are not met.

[Brekhovskikh 1957] L. M. Brekhovskikh, Волны в слоистых средах, Izd. Akad. Nauk SSSR, Moscow, 1957. Translated as *Waves in layered media*, Academic Press, New York, 1960.

[Brock 2005] L. M. Brock, "The coupled transversely isotropic bimaterial: interface crack extension", *J. Appl. Mech* (*ASME*) **72** (2005), 68–75.

[Brock 2009] L. M. Brock, "Basic problems of coupled thermoelasticity with thermal relaxation and pre-stress: aspects observed in exact and asymptotic solutions", *J. Therm. Stresses* **32** (2009), 593–622.

[Brock 2010] L. M. Brock, "Reflection and diffraction of plane temperature-step waves in orthotropic thermoelastic solids", *J. Therm. Stresses* **33** (2010), 879–904.

[Brock and Georgiadis 2007] L. M. Brock and H. G. Georgiadis, "Multiple-zone sliding contact with friction on an anisotropic half-space", *Int. J. Solids Struct.* **44** (2007), 2820–2836.

[Brock and Hanson 2006] L. M. Brock and M. T. Hanson, "Transient analysis of a suddenly-opening crack in a coupled thermoelastic solid with thermal relaxation", *J. Mech. Mater. Struct.* **1**:7 (2006), 1257–1268.

[Cagniard 1962] L. Cagniard, *Reflection and refraction of progressive seismic waves*, McGraw-Hill, New York, 1962. (E. A. Flinn and C. H. Dix, translators).

[Chadwick 1960] P. Chadwick, "Thermoelasticity. The dynamical theory", pp. 263–328 in *Progress in solid mechanics*, vol. 1, edited by I. N. Sneddon and R. Hill, North-Holland, Amsterdam, 1960.

[Crandall and Dahl 1959] S. H. Crandall and N. C. Dahl, *An introduction to the mechanics of solids*, McGraw-Hill, New York, 1959.

[Green and Lindsay 1972] A. E. Green and K. A. Lindsay, "Thermoelasticity", *J. Elasticity* **2** (1972), 1–7.

[Ignaczak and Ostoja-Starzewski 2010] J. Ignaczak and M. Ostoja-Starzewski, *Thermoelasticity with finite wave speeds*, Oxford, New York, 2010.

[Jones 1999] R. M. Jones, *Mechanics of composite materials*, 2nd ed., Brunner-Routledge, New York, 1999.

[Kraut 1963] E. A. Kraut, "Advances in the theory of anisotropic elastic wave propagation", *Rev. Geophys.* **1** (1963), 401–448.

[Lindsay 1960] R. B. Lindsay, *Mechanical radiation*, McGraw-Hill, New York, 1960.

[Lord and Shulman 1960] H. W. Lord and Y. Shulman, "Generalized dynamical theory of thermoelasticity", *J. Mech. Phys. Solids* **17** (1960), 297–309.

[Miklowitz 1978] J. Miklowitz, *The theory of elastic waves and waveguides*, North-Holland, Amsterdam, 1978.

[Payton 1983] R. G. Payton, *Elastic wave propagation in transversely isotropic materials*, Martinus Nijhoff, The Hague, 1983.

[Scott and Miklowitz 1967] R. A. Scott and J. Miklowitz, "Transient waves in anisotropic plates", *J. Appl. Mech* (*ASME*) **34** (1967), 104–110.

[Sharma and Sidhu 1986] J. N. Sharma and R. S. Sidhu, "On the propagation of plane harmonic waves in anisotropic generalized thermoelasticity", *Int. J. Eng. Sci.* **24** (1986), 1511–1516.

[Sokolnikoff 1956] I. S. Sokolnikoff, *Mathematical theory of elasticity*, 2nd ed., McGraw-Hill, New York, 1956.

LOUIS MILTON BROCK: brock@engr.uky.edu
Department of Mechanical Engineering, University of Kentucky, 265 RGAN, Lexington, KY 40506-0503, United States

■■
■msp

ON THE MYSTERY OF CALDERÓN'S FORMULA
FOR THE GEOMETRY OF AN INCLUSION IN ELASTIC MATERIALS

HUY DUONG BUI

I write this paper to pay homage to Marie-Louise Steele and in honor of Charles R. Steele. I have had the pleasure and the honor to serve their journals IJSS and JoMMS, with George Herrmann. They have made Solids & Structures and now Material Sciences a subject of nobility to all of us.

We consider the nonlinear inverse problem of determining an inclusion in an elastic body, in antiplane shear loading. The perturbation of the shear modulus due to the inclusion was determined by Calderón (1980) in the case of a small amplitude of perturbation. For the general nonlinear case, the problem is decomposed into two linear problems: a source inverse problem, which determines the geometry of the inclusion, and a Volterra integral equation of the first kind for determining the amplitude. In this paper, we deal only with the determination of the inclusion geometry in the two-dimensional problem. We derive a simple formula for determining the inclusion geometry. This formula enables us to investigate the mystery of Calderón's solution for the linearized perturbation h^0, raised by Isaacson and Isaacson (1986), in the case of axisymmetry. By using a series method for numerical analysis, they found that the supports of the perturbation, in the linearized theory and the nonlinear theory in the axisymmetric case, are practically the same. We elucidate the mystery by discovering that both theories give exactly the same support of the perturbation, $\operatorname{supp}(h^0) \equiv \operatorname{supp}(h)$, for the general case of geometry and loadings. Then, we discuss an application of the geometry method to locate an inclusion and solve the source inverse problem, which gives an indication of the amplitude of the perturbation.

1. Introduction

Inverse problems for defect and crack identification in elasticity have many applications in medicine and the mechanics of materials. In medicine, tomography techniques using mechanical loads such as an antiplane shear loading on animal tissue, are worked out in [Catheline et al. 2004]. Cancer tumors are expected to have higher density and higher stiffness or shear modulus than sound tissues, so the difference in material properties between sound and malicious tissues is detected by mechanical loads and responses. Auscultation by a doctor is nothing but a rudimentary method of endoscopy relying on the same principle.

In the mechanics of materials, damage is known to result from microcracks which lower locally the elastic constants. New topics in mechanical tomography have recently been the subjects of several works. For example, solutions to crack inverse problems in two and three dimensions are known in elasticity [Andrieux and Ben Abda 1992; Andrieux et al. 1999; Bui et al. 2005] and in viscoelasticity [Bui et al. 2009], in statics as well as in dynamics, under the assumption of small frequencies. In elastodynamics, solutions of inverse crack problems are obtained in [Bui et al. 2005] where the solution to an earthquake

---

inverse problem to recover the faulting process was proposed. A review of several exact solutions to inverse problems is found in [Bui 2006].

The case of distributed defects was first studied in [Calderón 1980] for the scalar elastic equation, then in [Bui and Chaillat 2009] for the case of dynamic viscoelasticity in the low frequency domain. Most of the works mentioned used reciprocity functional techniques for solving the following inverse problem. Find functions h and u satisfying the field equation and two superabundant boundary data f and g:

$$\operatorname{div}\big[(1 + h(\boldsymbol{x}))\operatorname{grad} u\big] = 0 \quad \text{in } \Omega, \tag{1-1}$$

$$u = f \quad \text{and} \quad \frac{\partial u}{\partial n} = g \quad \text{in } \partial\Omega. \tag{1-2}$$

The pair of data (f, g) is compatible with the perturbation h. We consider normalized constants so that the shear modulus of an homogeneous body is 1 and the inhomogeneity is characterised by the material constant $1 + h(\boldsymbol{x})$. We assume nonzero measure of support of h which excludes the hair line inclusion case. For small amplitude of the perturbation h, Calderón [1980] derived explicitly the solution, denoted hereafter by $h^0(\boldsymbol{x})$.

In this paper, we shall reconsider the nonlinear Calderón inverse problem for identifying distributed defects in elasticity of arbitrary geometry. First we give a formula for the geometry inclusion. This nonlinear problem has been solved numerically in [Isaacson and Isaacson 1989] for data corresponding to a circular geometry of the inclusion and body, under axisymmetric radial loads. They used a series method to solve the equation in the axisymmetric case, with data (f, g) corresponding to $h(r)$ not necessary of small amplitude. The data for the numerical test is obtaining first by solving a Dirichlet boundary value with condition $u = f$ on the boundary and with a given h on the inclusion. Then the boundary value of the gradient $g = \partial u/\partial n$ is calculated. They have shown by numerical experiments (with a series method) that the support of the inclusion function $h^0(\boldsymbol{x})$ from Calderón's formula is almost indiscernible from the circle introduced for (f, g). This is the mystery of Calderón's solution, which seems to work for the nonlinear case as well. It is incredible that a formula derived for a small amplitude perturbation still works for the general case. We shall clarify the mystery by comparing the geometry inclusions in both theories. Then we discuss an application to the source inverse problem.

2. The nonlinear Calderón equation for *h* and the linearized solution

We consider the identification of internal defects in the antiplane problem of elasticity, or stationary heat/mass transfer phenomena, or electricity conduction. In the antiplane problem the shear modulus is of the form $\mu(\boldsymbol{x}) = 1 + h(\boldsymbol{x})$. The shear modulus in the absence of defect ($h = 0$) is normalized to unity.

We assume that the defect is characterized by the function $h(\boldsymbol{x})$ with compact support $C \subset \Omega$ and $h = 0$ on the boundary $\partial\Omega$, see Figure 1. Defects may not be necessarily unique.

Consider an adjoint problem for the sound solid $h = 0$ to determine the function φ:

$$\operatorname{div} \operatorname{grad} \varphi = 0 \quad \text{in } \Omega, \tag{2-1a}$$

$$\frac{\partial\varphi}{\partial n} = a(\boldsymbol{x}) \quad \text{on } \partial\Omega, \tag{2-1b}$$

with the equilibrium condition $\int_{\partial\Omega} a \, dS = 0$, and let $b(\boldsymbol{x})$ be the boundary value of φ.
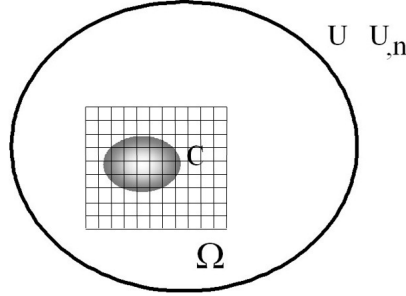
Figure 1. Identification of an inclusion C represented by a perturbation of the material constant. A mesh for numerical analysis with data U and $\partial U/\partial n =: U_{,n}$.

Combining (1-1) and (2-1a), after multiplying (1-1) by φ and (2-1a) by u, integrating by parts and taking account of $h = 0$ on $\partial\Omega$, we obtain the Calderón equation for h for any φ

$$\int_C h(\boldsymbol{x}) \operatorname{grad} u(\boldsymbol{x}; h) \cdot \operatorname{grad} \varphi(\boldsymbol{x}) d^2 x = \int_{\partial\Omega} (gb - fa) dS =: R(f, g; \varphi), \quad \forall\varphi. \qquad (2\text{-}2)$$

The right-hand side of (2-2) is known from the boundary data (f, g) of the current field u and the data (a, b) of the adjoint function φ. In the following, it will be denoted by $R(f, g; \varphi)$ or simply $R(\varphi)$. We remark that (2-2) is nonlinear in h, firstly because function $u(\boldsymbol{x}; h)$ is yet unknown and secondly because the integration domain C is unknown. Equation (2-2) constitutes the nonlinear variational equation for h. The arbitrariness of φ is the key point to solve the nonlinear problem, as shown in several examples given in [Bui 2006]. An important case is when $u(\boldsymbol{x}; h)$ as well as the geometry of C are known, for which (2-2) becomes linear.

Following [Calderón 1980] we can linearize (2-2) by replacing the unknown function u by the function u^0 which satisfies the harmonic equation

$$\operatorname{div} \operatorname{grad} u^0 = 0 \quad \text{in } \Omega. \qquad (2\text{-}3)$$

We obtain

$$\int_C h(\boldsymbol{x}) \operatorname{grad} u^0(\boldsymbol{x}) \cdot \operatorname{grad} \varphi(\boldsymbol{x}) d^2 x = R(f, g; \varphi). \qquad (2\text{-}4)$$

Whatever the extended function $\operatorname{grad} u$ outside the domain may be, (2-2) can also be written with the integral over the whole plane because the extended function $\tilde{h}$ vanishes outside the domain:

$$\int_{\mathbb{R}^2} \tilde{h}(\boldsymbol{x}) \operatorname{grad} u(\boldsymbol{x}; h) \cdot \operatorname{grad} \varphi(\boldsymbol{x}) d^2 x = R(f, g; \varphi), \quad \forall\varphi. \qquad (2\text{-}5)$$

In this case boundary data a, b of the harmonic function φ may not be specified. We consider the adjoint harmonic function, which depends on parameters $\boldsymbol{\xi}$ of the ξ-plane:

$$\varphi(\boldsymbol{x}, \xi) = \exp\big(-i(x_1\xi_1 + x_2\xi_2)\big) \exp(-x_1\xi_2 + x_2\xi_1). \qquad (2\text{-}6)$$

Calderón assumed a small amplitude of h and expanded the unknown function in the form $u(\boldsymbol{x}; h) = u(\boldsymbol{x}; 0) + O(h)$. By taking $u(\boldsymbol{x}; 0) = \varphi(\boldsymbol{x}, \xi)$, given in (2-6), he obtained explicitly the solution with

$R(\boldsymbol{\xi}) =: R(f, g; \varphi(\boldsymbol{x}; \boldsymbol{\xi}))$:

$$\tilde{h}^0(\boldsymbol{x}) = -\frac{1}{4\pi^2} \int_{\mathbb{R}^2} \frac{2}{\|\boldsymbol{\xi}\|^2} R(\boldsymbol{\xi}) \exp(i\boldsymbol{x} \cdot \boldsymbol{\xi}) d^2\xi. \tag{2-7}$$

Strictly speaking, (2-7) is derived for (f, g) compatible with small perturbation. What happens when one applies this formula to the general case of arbitrary h? It seems, from [Isaacson and Isaacson 1989] with their example of a circular inclusion in a circular domain, that the support of the solution given by (2-7) is indiscernible from the circle introduced for obtaining data f, g. To clarify this mystery, we need to derive the solution for the inclusion geometry, in the general nonlinear theory.

3. Reduction to two successive simpler problems

Let us rewrite (1-1) in the following form:

$$\mathrm{div}\,\mathrm{grad}\,u + S(\boldsymbol{x}) = 0 \quad \text{in } \Omega, \tag{3-1}$$

where $S(\boldsymbol{x}) = \mathrm{div}(h(\boldsymbol{x})\,\mathrm{grad}\,u(\boldsymbol{x}; h))$; see also [Bui and Chaillat 2009].

Determining S with boundary data $u = f$ and $\partial u/\partial n = g$ is a classical *source* inverse problem; see [Isakov 1990; Alves and Ha-Duong 1997]. Here, we consider the source inverse problem already solved, to obtain the source $S(\boldsymbol{x})$, its support C, and the true displacement field denoted by $U(\boldsymbol{x}) := u(\boldsymbol{x}; h)$, even if $h(\boldsymbol{x})$ is unknown. We obtain also the strain, but not the stress because h is yet unknown. It is interesting to remark that the solution of the source inverse problem (3-1) provides the geometry or support of the unknown $h(\boldsymbol{x})$. As a matter of fact, since the true displacement is known as U, we have

$$S(\boldsymbol{x}) = \mathrm{div}\big(h(\boldsymbol{x})\,\mathrm{grad}\,U(\boldsymbol{x})\big). \tag{3-2}$$

The source term (3-2) is a linear combination of derivatives of h in the distributional sense. Its support is the inclusion itself, $\mathrm{supp}(S) \equiv \mathrm{supp}(h) = C$. Therefore, we have a first method to derive only the geometry by solving a source inverse problem. We will discuss next a simpler method. It is based on the nonlinear Calderón equation rewritten as

$$\int_C h(\boldsymbol{x})\,\mathrm{grad}\,U(\boldsymbol{x}) \cdot \mathrm{grad}\,\varphi(\boldsymbol{x}; \boldsymbol{\xi}) d^2x = R\big(\varphi(\boldsymbol{\xi})\big). \tag{3-3}$$

With appropriate choice of φ, for example given by (2-6), we obtain a Volterra equation of the first kind for h, with kernel $K(\boldsymbol{x}, \boldsymbol{\xi}) := \mathrm{grad}\,U(\boldsymbol{x}) \cdot \mathrm{grad}\,\varphi(\boldsymbol{x}, \boldsymbol{\xi})$ which determines both geometry and amplitude of the perturbation. However, we do not solve the Volterra integral equation, but show how we derive the support of h directly.

4. The mystery of Calderón's solution

The adjoint function (2-6) as well as its gradient $\mathrm{grad}\,\varphi(\boldsymbol{x}; \boldsymbol{\xi})$ are analytic in the whole x-space and ξ-space (except at infinity), and thus can be expanded into infinite series of x_r and ξ_h. We expand $\mathrm{grad}\,\varphi(\boldsymbol{x}; \boldsymbol{\xi}) = \exp(-i(\boldsymbol{x} \cdot \boldsymbol{\xi})) \exp(-x_1\xi_2 + x_2\xi_1)\big((-i\xi_1 - \xi_2)\boldsymbol{e}^1 + (-i\xi_2 + \xi_1)\boldsymbol{e}^2\big)$ as

$$\mathrm{grad}\,\varphi(\boldsymbol{x}; \boldsymbol{\xi}) = \left(\sum_{h,k,r,s=1}^{2} \sum_{n,m,p,q=0}^{\infty} \boldsymbol{a}_{nmpq}^{hkrs}(i\xi_h)^n(i\xi_k)^m x_r^p x_s^q\right) \exp\big(-i(\boldsymbol{x} \cdot \boldsymbol{\xi})\big), \tag{4-1}$$

with constant complex vectors a^{hkrs}_{nmpq}. We extend $h(\boldsymbol{x})$ to the infinite plane $\mathbb{R}^2$ by putting $h = 0$ outside C and denote its extension by $\tilde{h}$ and obtain the nonlinear Calderón equation in the form (the dot means a scalar product between vectors)

$$\int_{\mathbb{R}^2} \tilde{h}(\boldsymbol{x}) \operatorname{grad} U(\boldsymbol{x}) \cdot \left(\sum_{h,k,r,s=1}^{2} \sum_{n,m,p,q=0}^{\infty} a^{hkrs}_{nmpq} (i\xi_h)^n (i\xi_k)^m x_r^p x_s^q \right) \exp(-i\boldsymbol{x} \cdot \boldsymbol{\xi}) d^2x = R\big(\varphi(\boldsymbol{\xi})\big), \quad (4\text{-}2)$$

or equivalently (using the properties of the Fourier transform)

$$\int_{\mathbb{R}^2} \sum_{h,k,r,s=1}^{2} \sum_{n,m,p,q=0}^{\infty} a^{hkrs}_{nmpq} \cdot \frac{\partial^n}{\partial x_h^n} \frac{\partial^m}{\partial x_k^m} \big(x_r^p x_s^q \tilde{h}(\boldsymbol{x}) \operatorname{grad} U(\boldsymbol{x}) \big) \exp(-i\boldsymbol{x} \cdot \boldsymbol{\xi}) d^2x = R\big(\varphi(\boldsymbol{\xi})\big). \quad (4\text{-}3)$$

Let us define the function appearing in the above series by $F(\boldsymbol{x})$:

$$F(\boldsymbol{x}) = \sum_{h,k,r,s=1}^{2} \sum_{n,m,p,q=0}^{\infty} a^{hkrs}_{nmpq} \cdot \frac{\partial^n}{\partial x_h^n} \frac{\partial^m}{\partial x_k^m} \big(x_r^p x_s^q \tilde{h}(\boldsymbol{x}) \operatorname{grad} U(\boldsymbol{x}) \big), \quad (4\text{-}4)$$

$$\int_{\mathbb{R}^2} F(\boldsymbol{x}) \exp(-i\boldsymbol{x} \cdot \boldsymbol{\xi}) d^2x = R\big(\varphi(\boldsymbol{\xi})\big). \quad (4\text{-}5)$$

Because the function $F(\boldsymbol{x})$ is a linear combination of $\tilde{h}(\boldsymbol{x})$ and its partial derivatives, it has the same support $C = \operatorname{supp}(F) = \operatorname{supp}(h)$. Therefore we get the inclusion geometry C by the support of F which satisfies (4-5). It follows that function $F(\boldsymbol{x})$ is the inverse Fourier transform of $R(\varphi(\boldsymbol{\xi}))$. The inclusion geometry for the nonlinear theory is solved by

$$F(\boldsymbol{x}) = \frac{1}{4\pi^2} \int_{\mathbb{R}^2} R\big(\varphi(\boldsymbol{\xi})\big) \exp(+i\boldsymbol{x} \cdot \boldsymbol{\xi}) d^2\xi. \quad (4\text{-}6)$$

This exact solution for the geometry of the inclusion enables us to investigate the mystery of Calderón's linearized solution which seems to work well for any perturbation. Let us make first a remark about function F given in terms of $R(\varphi(\boldsymbol{\xi}))$. From the expression of $R(\varphi(\boldsymbol{\xi}))$ we can easily check, for a bounded solid, that R is of the "exponential type", that is, the complex extension $R(z)$ of $R(\boldsymbol{\xi})$, obtained by the substitution $\boldsymbol{\xi} = (\xi_1, \xi_2) \to (\xi_1 + i\eta_1, \xi_2 + i\eta_2) =: z$ in $R(\boldsymbol{\xi})$, has the bound $\|R(z)\| \le C \exp(a\|z\|)$ with $C > 0$ and $a > 0$. According to the Paley–Wiener theorem [Schwartz 1966], the function $R(\boldsymbol{\xi})$ is the Fourier transform of a function with *compact* support. Therefore $F(\boldsymbol{x})$ given by (4-6) is a compactly supported function.

Consider now the linearized solution (2-7) for $\tilde{h}^0(\boldsymbol{x})$ which we recall below:

$$\tilde{h}^0(\boldsymbol{x}) = -\frac{1}{4\pi^2} \int_{\mathbb{R}^2} \frac{2}{\|\boldsymbol{\xi}\|^2} R(\boldsymbol{\xi}) \exp(i\boldsymbol{x} \cdot \boldsymbol{\xi}) d^2\xi. \quad (4\text{-}7)$$

Using the properties of the Fourier transform, we can rewrite (4-7) in the form

$$-\frac{1}{2} \left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} \right) \tilde{h}^0(\boldsymbol{x}) = \frac{1}{4\pi^2} \int_{\mathbb{R}^2} R(\boldsymbol{\xi}) \exp(i\boldsymbol{x} \cdot \boldsymbol{\xi}) d^2\xi \equiv F(\boldsymbol{x}). \quad (4\text{-}8)$$

The left-hand side of this equation is a function having the same support as $\tilde{h}^0(\boldsymbol{x})$. Therefore we conclude that both functions $\tilde{h}^0(\boldsymbol{x})$ of the linearized theory and $h(\boldsymbol{x})$ of the nonlinear theory have the same geometry as supp(F). The mystery revealed by [Isaacson and Isaacson 1989] is finally elucidated. It is amazing that Calderón's linearized theory gives the exact solution for the geometry of the inclusion, whatever the perturbation may be. Of course, the amplitude of function $h(\boldsymbol{x})$ has to be determined by the Volterra integral equation, by a standard numerical approach. In the next section, we shall discuss a numerical procedure to determine not the function $h(\boldsymbol{x})$ itself, which can be found by standard techniques for the Volterra integral equation, but the source term $S(\boldsymbol{x})$, which is directly related to the amplitude of the perturbation.

5. A numerical approach to the source inverse problem by moving windows

It is known that the Volterra integral equation (3-3) is an ill-posed problem. Numerically it is difficult to recover exactly a function h which is strictly equal to zero outside the inclusion, which is assumed here to be unique. Therefore, it is of interest to consider a small *moving window* which is discretized in regular meshes and to solve numerically the source inverse problem for N point sources $S(\boldsymbol{x}) = \sum_{i=1}^{N} \lambda_i \delta(\boldsymbol{x} - \boldsymbol{a}_i)$, with source points at the centers of finite elements, and unknown amplitudes λ_i. With a chosen window, we enforce the condition $h = 0$ outside it. For a large window enclosing the solid, it is shown in [El Badia and Ha-Duong 2000] that the solution for a finite number N of sources approaching the source $S(\boldsymbol{x})$ exists and is unique. By choosing a particular window inside the solid domain, we search a solution which vanishes outside it. If the window does not contain entirely the source, we get a wrong solution and the corresponding image of the numerical solution is then blurred. Only in the case where the window contains the inclusion is a sharp image obtained. This procedure resembles echography imaging of a body. For example, by trial and error, one moves the echography device on the body of an expectant mother in order to search the right location to reveal a sharp image of her fetus. In our examples of the source problem, studying a tumor in live tissue or a damaged zone in materials, the moving window is a 4×5 mesh. For example, in Figure 2, the image on the left corresponds to the wrong solution, while the one on the right is correct.

Therefore it is of great interest to know the right location of the initial mesh. This is precisely provided by the support of the function F given in (4-6).
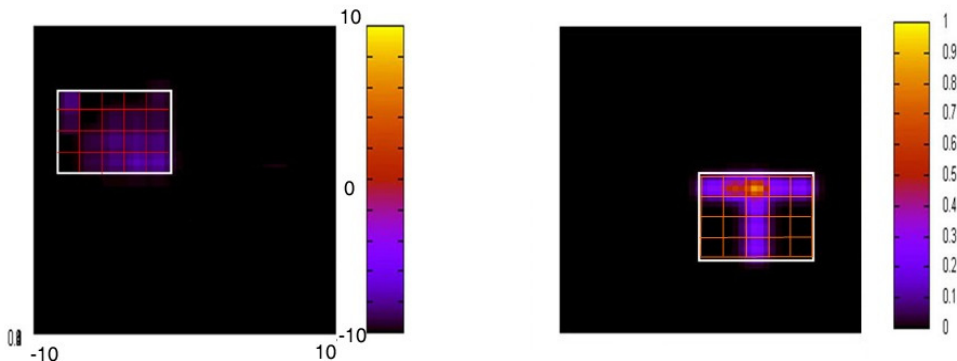

Figure 2. Imaging of a defect. Left: bad window, wrong solution. Right: right window, good solution.

6. Conclusions

In this paper, we consider the problem of finding the perturbation of a material constant in elastic solids which satisfies the nonlinear Calderón equation. The nonlinear problem reduces to two successive ones: a source inverse problem and a Volterra integral equation of the first kind. The first problem provides the inclusion geometry supp(h) explicitly. The second provides the magnitude of h. We make a comparison between the geometry of an inclusion in the small perturbation case and the geometry in the nonlinear case and find that both inclusion geometries are identical for arbitrary loading and geometry of the solid. Our result elucidates the mystery of the linearized Calderón solution for geometry which works well for the nonlinear case, as revealed numerically in the axisymmetric example given in [Isaacson and Isaacson 1989].

References

[Alves and Ha-Duong 1997] C. J. S. Alves and T. Ha-Duong, "On inverse scattering by screens", *Inverse Problems* **13**:5 (1997), 1161–1176.

[Andrieux and Ben Abda 1992] S. Andrieux and A. Ben Abda, "Identification de fissures planes par une donnée de bord unique: un procédé direct de localisation et d'identification", *C. R. Acad. Sci. Paris Sér. I Math.* **315**:12 (1992), 1323–1328.

[Andrieux et al. 1999] S. Andrieux, A. Ben Abda, and H. D. Bui, "Reciprocity principle and crack identification", *Inverse Problems* **15**:1 (1999), 59–65.

[Bui 2006] H. D. Bui, *Fracture mechanics: inverse problems and solutions*, Springer, Dordrecht, 2006.

[Bui and Chaillat 2009] H. D. Bui and S. Chaillat, "On a nonlinear inverse problem in viscoelasticity", *Vietnam J. Mech.* **31**:3-4 (2009), 211–219.

[Bui et al. 2005] H. D. Bui, A. Constantinescu, and H. Maigre, "An exact inversion formula from determining a planar fault from boundary measurements", *J. Inverse Ill-Posed Probl.* **13**:6 (2005), 553–565.

[Bui et al. 2009] H. D. Bui, S. Chaillat, A. Constantinescu, and E. Grasso, "Identification of a planar crack in Zener type viscoelasticity", *Ann. Solid Struct. Mech.* **1**:1 (2009), 3–8.

[Calderón 1980] A.-P. Calderón, "On an inverse boundary value problem", pp. 65–73 in *Seminar on Numerical Analysis and its Applications to Continuum Physics* (Rio de Janeiro, 1980), Coleção Atlas **12**, Soc. Brasil. Mat., Rio de Janeiro, 1980.

[Catheline et al. 2004] S. Catheline, J. L. Gennisson, G. Delon, M. Fink, R. Sinkus, S. Abouelkaram, and J. Culioli, "Measurement of viscoelastic properties of homogeneous soft solid using transient elastography: an inverse problem approach", *J. Acoust. Soc. Am.* **116**:6 (2004), 3734–3741.

[El Badia and Ha-Duong 2000] A. El Badia and T. Ha-Duong, "An inverse source problem in potential analysis", *Inverse Problems* **16**:3 (2000), 651–663.

[Isaacson and Isaacson 1989] D. Isaacson and E. L. Isaacson, "Comment on A.-P. Calderón's paper: 'On an inverse boundary value problem'", *Math. Comp.* **52**:186 (1989), 553–559.

[Isakov 1990] V. Isakov, *Inverse source problems*, Mathematical surveys and monographs **34**, American Mathematical Society, Providence, RI, 1990.

[Schwartz 1966] L. Schwartz, *Théorie des distributions*, Hermann, Paris, 1966.

HUY DUONG BUI: hdbui37@yahoo.fr
Laboratoire de Mécanique des Solides, École Polytechnique, Route de Saclay, 91128 Palaiseau CEDEX, France

msp

SINGULAR HARMONIC PROBLEMS AT A WEDGE VERTEX: MATHEMATICAL ANALOGIES BETWEEN ELASTICITY, DIFFUSION, ELECTROMAGNETISM, AND FLUID DYNAMICS

ALBERTO CARPINTERI AND MARCO PAGGI

Dedicated to the memory of Marie-Louise Steele.

Multimaterial wedges are frequently observed in composite materials. They consist of two or more sectors of dissimilar materials joined together, whose interfaces converge at the same vertex. Due to the mismatch in material properties such as Young's modulus, thermal conductivity, dielectric permittivity, or magnetic permeability, these geometrical configurations can lead to singular fields at the junction vertex. This paper discusses mathematical analogies, focused on singular harmonic problems, between antiplane shear problem in elasticity due to mode III loading or torsion, the steady-state heat transfer problem, and the diffraction of waves in electromagnetism. In the case of a single material wedge, a mathematical analogy between elasticity and fluid dynamics is also outlined. The proposed unified mathematical formulation is particularly convenient for the identification of common types of singularities (power-law or logarithmic type), the definition of a standardized method to solve nonlinear eigenvalue problems, and the determination of common geometrical and material configurations allowing the relief or removal of different singularities.

1. Introduction

Singular stress states occur in many boundary value problems of linear elasticity where different materials are present (see [England 1971; Paggi and Carpinteri 2008; Sinclair 2004a; 2004b] for a broad overview). In this context, problems involving multimaterial wedges or junctions have received a great attention from the scientific community, since they are commonly observed in composite materials. In linear elasticity, most research has been directed toward the characterization of stress singularities for in-plane loading, where the problem is governed by a biharmonic equation. Out-of-plane loading, also referred to as the antiplane shear problem, is governed by a simpler harmonic equation. Stress singularities due to antiplane loading were firstly addressed by Rao [1971]. Afterwards, Fenner [1976] examined the mode III loading problem of a crack meeting a bimaterial interface using the eigenfunction expansion method proposed by Williams [1952]. More recently, Ma and Hour [1989; 1990] analyzed bimaterial wedges using the Mellin transform technique, whereas Pageau et al. [1995] investigated the singular stress field of bonded and debonded tri-material junctions according to the eigenfunction expansion method.

Mathematical analogies among elasticity, electromagnetism, and conductivity have been known and exploited for a long time (see [Hashin and Shtrikman 1962; Duan et al. 2006], for instance). Sinclair [1980] pointed out the mathematical analogy between the singular steady-state heat transfer and the

singular antiplane loading of composite regions (see also [Paggi and Carpinteri 2008] for a detailed discussion of the boundary conditions). In [Paggi et al. 2009; 2010] we have established an analogy between elasticity and electromagnetism in the case of singular fields. In the solution of diffraction problems, in fact, Bouwkamp [1946] and Meixner [1972] found that the electromagnetic field vectors may become infinite at the sharp edges of a diffracting obstacle. For in-plane problems, a mathematical analogy between elasticity and dynamics of viscous fluids also exists; see [Dean and Montagnon 1949; Paggi and Carpinteri 2008; Carpinteri and Paggi 2009] for more details.

This paper presents mathematical analogies between antiplane shear problem in elasticity due to mode III loading or torsion, the steady-state heat transfer problem, and the diffraction of waves in electromagnetism. In the case of a single material wedge, a mathematical analogy between antiplane elasticity and fluid dynamics of incompressible fluids characterized by a potential flow will also be outlined. The proposed unified mathematical formulation will be based on the eigenfunction expansion method, which has been proven in [Paggi and Carpinteri 2008] to be mathematically equivalent to the Muskhelishvili complex function representation and to the Mellin transform technique for the characterization of elastic singularities at multimaterial junctions. As a main outcome, the order of the stress singularities of various geometrical and mechanical configurations already determined in the literature can be adopted for the analogous diffusion, electromagnetic, and fluid dynamics problems, without the need of performing new calculations. Finally, the possibility to extend the dimensionless numbers used in elasticity to the other analogous physical problems is discussed. In particular, as the *brittleness number* related to the stress-intensity factor rules the competition between brittle crack propagation and plastic flow collapse, a *turbulence number* that rules the competition between laminar and turbulent flow is proposed for fluid dynamics. The use of this new dimensionless number, in addition to the classical Reynolds number, is expected to be of paramount importance.

2. Stress singularities in antiplane elasticity

The geometry of a plane elastostatic problem consisting of $n-1$ dissimilar isotropic, homogeneous sectors of arbitrary angles perfectly bonded along their interfaces converging to the same vertex O is
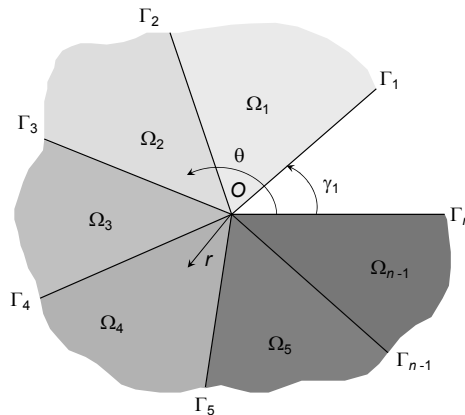

Figure 1. Geometry of a multimaterial wedge.

shown in Figure 1. Each of the material regions is denoted by Ω_i with $i = 1, \ldots, n-1$, and it is comprised between the interfaces Γ_i and Γ_{i+1}.

Antiplane shear (mode III) due to out-of-plane loading on composite wedges can lead to stresses that can be unbounded at the junction vertex O. When out-of-plane deformations only exist, the following displacements in cylindrical coordinates can be considered with the origin at the vertex O:

$$u_r = 0, \quad u_\theta = 0, \quad u_z = u_z(r, \theta), \tag{2-1}$$

where u_z is the out-of-plane displacement, which depends on r and θ. For such a system of displacements, the strain field components become

$$\varepsilon_r = \varepsilon_\theta = \varepsilon_z = \gamma_{r\theta} = 0, \quad \gamma_{rz} = \frac{\partial u_z}{\partial r}, \quad \gamma_{\theta z} = \frac{1}{r}\frac{\partial u_z}{\partial \theta}. \tag{2-2}$$

From the application of the Hooke's law, the stress field components are given by:

$$\sigma_r^i = \sigma_\theta^i = \sigma_z^i = \tau_{r\theta}^i = 0, \quad \tau_{rz}^i = G_i\gamma_{rz}^i = G_i\frac{\partial u_z^i}{\partial r}, \quad \tau_{\theta z}^i = G_i\gamma_{\theta z}^i = \frac{G_i}{r}\frac{\partial u_z^i}{\partial \theta}, \tag{2-3}$$

where G_i is the shear modulus of the i-th material region. The equilibrium equations in absence of body forces reduce to a single relationship between the tangential stresses:

$$\frac{\partial \tau_{rz}^i}{\partial r} + \frac{1}{r}\frac{\partial \tau_{\theta z}^i}{\partial \theta} + \frac{1}{r}\tau_{rz}^i = 0 \quad \text{for all } (r, \theta) \in \Omega_i. \tag{2-4}$$

Introducing (2-3) into (2-4), the Laplace condition upon u_z is derived:

$$\frac{\partial^2 u_z^i}{\partial r^2} + \frac{1}{r}\frac{\partial u_z^i}{\partial r} + \frac{1}{r^2}\frac{\partial^2 u_z^i}{\partial \theta^2} = \nabla^2 u_z^i = 0 \quad \text{for all } (r, \theta) \in \Omega_i. \tag{2-5}$$

In the framework of the eigenfunction expansion method [Paggi and Carpinteri 2008], the following separable variable form for the longitudinal displacement u_z^i can be adopted, for all $(r, \theta) \in \Omega_i$:

$$u_z^i(r, \theta) = \sum_j r^{\lambda_j} f_{i,j}(\theta, \lambda_j), \tag{2-6}$$

where the λ_j are the eigenvalues and $f_{i,j}$ the eigenfunctions of the problem. Summation over j is introduced in (2-6) since it is possible to have more than one eigenvalue and the superposition principle can be applied.

Introducing (2-6) into (2-5), we find the following relation, holding for each eigenvalue λ_j:

$$r^{\lambda_j - 2}\left(\frac{d^2 f_{i,j}}{d\theta^2} + \lambda_j^2 f_{i,j}\right) = 0. \tag{2-7}$$

Hence, the coefficients of the term in $r^{\lambda_j - 2}$ must vanish, implying that the eigenfunctions $f_{i,j}$ are a linear combination of trigonometric functions:

$$f_{i,j}(\theta, \lambda_j) = A_{i,j}\sin(\lambda_j\theta) + B_{i,j}\cos(\lambda_j\theta). \tag{2-8}$$

If we introduce the series expansion (2-6) into (2-3), the longitudinal displacement and the tangential stresses can be expressed in terms of the eigenfunction and its first derivative:

$$u_z^i = r^{\lambda_j} f_{i,j} = r^{\lambda_j} \left[A_{i,j} \sin(\lambda_j \theta) + B_{i,j} \cos(\lambda_j \theta) \right], \tag{2-9a}$$

$$\tau_{rz}^i = G_i \lambda_j r^{\lambda_j - 1} f_{i,j} = G_i \lambda_j r^{\lambda_j - 1} \left[A_{i,j} \sin(\lambda_j \theta) + B_{i,j} \cos(\lambda_j \theta) \right], \tag{2-9b}$$

$$\tau_{\theta z}^i = G_i r^{\lambda_j - 1} f_{i,j}' = G_i \lambda_j r^{\lambda_j - 1} \left[A_{i,j} \cos(\lambda_j \theta) - B_{i,j} \sin(\lambda_j \theta) \right]. \tag{2-9c}$$

The determination of the power of the stress singularity, $\lambda_j - 1$, can be performed by imposing the boundary conditions (BCs) along the edges Γ_1 and Γ_n and at the bimaterial interfaces Γ_i, with $i = 2, \ldots, n-1$. Along the edges Γ_1 and Γ_n, defined by the angles γ_1 and γ_n, we consider two possibilities: one corresponding to unrestrained stress-free edges

$$\tau_{\theta z}^i(r, \gamma_1) = 0, \quad \tau_{\theta z}^i(r, \gamma_n) = 0, \tag{2-10}$$

and the other for fully restrained (clamped) edges

$$u_z^i(r, \gamma_1) = 0, \quad u_z^i(r, \gamma_n) = 0. \tag{2-11}$$

At the interfaces, the following continuity conditions of displacements and stresses have to be imposed ($i = 1, \ldots, n-2$):

$$u_z^i(r, \gamma_{i+1}) = u_z^{i+1}(r, \gamma_{i+1}), \quad \tau_{\theta z}^i(r, \gamma_{i+1}) = \tau_{\theta z}^{i+1}(r, \gamma_{i+1}). \tag{2-12}$$

In this way, a set of $2n - 2$ homogeneous equations in the $2n - 1$ unknowns $A_{i,j}$, $B_{i,j}$, and λ_j can be symbolically written as:

$$\boldsymbol{\Lambda} \, \boldsymbol{v} = \boldsymbol{0}, \tag{2-13}$$

where $\boldsymbol{\Lambda}$ denotes the coefficient matrix which depends on the eigenvalue, and $\boldsymbol{v}$ represents the vector that collects the unknowns $A_{i,j}$ and $B_{i,j}$. More specifically, the coefficient matrix appearing in (2-13) is characterized by a sparse structure:

$$\boldsymbol{\Lambda} = \begin{bmatrix} \boldsymbol{N}_{\gamma_1}^1 & & & & & & \\ \boldsymbol{M}_{\gamma_2}^1 & -\boldsymbol{M}_{\gamma_2}^2 & & & & & \\ & \boldsymbol{M}_{\gamma_3}^2 & -\boldsymbol{M}_{\gamma_3}^3 & & & & \\ & & \ddots & \ddots & & & \\ & & & \boldsymbol{M}_{\gamma_i}^{i-1} & -\boldsymbol{M}_{\gamma_i}^i & & \\ & & & & \ddots & \ddots & \\ & & & & & \boldsymbol{M}_{\gamma_{n-1}}^{n-2} & -\boldsymbol{M}_{\gamma_{n-1}}^{n-1} \\ & & & & & & \boldsymbol{N}_{\gamma_n}^{n-1} \end{bmatrix} \tag{2-14}$$

where the nonnull elementary matrix $\boldsymbol{M}_\theta^i$ related to the interface BCs is given by

$$\boldsymbol{M}_\theta^i = \begin{bmatrix} \sin(\lambda_j \theta) & \cos(\lambda_j \theta) \\ G_i \cos(\lambda_j \theta) & -G_i \sin(\lambda_j \theta) \end{bmatrix} \tag{2-15}$$

and the components of the vector $\boldsymbol{v}$ are

$$\boldsymbol{v} = \{\boldsymbol{v}^1, \boldsymbol{v}^2, \ldots, \boldsymbol{v}^i, \ldots, \boldsymbol{v}^{n-2}, \boldsymbol{v}^{n-1}\}, \tag{2-16}$$

with $\boldsymbol{v}^i = \{A_{i,j}, B_{i,j}\}^T$. The two remaining terms N_θ^i depend on the BCs along the edges Γ_1 and Γ_n. For stress-free edges we have

$$N_\theta^i = \{\cos(\lambda_j\theta), -\sin(\lambda_j\theta)\}, \tag{2-17}$$

whereas for clamped edges it is given by

$$N_\theta^i = \{\sin(\lambda_j\theta), \cos(\lambda_j\theta)\}. \tag{2-18}$$

A nontrivial solution of the equation system (2-13) exists if and only if the determinant of the coefficient matrix vanishes. This condition yields an eigenequation which has to be solved for the eigenvalues λ_j that, in the most general case, do depend on the elastic properties of the materials.

3. Heat flux singularities in diffusion problems

The analogy between steady-state heat transfer and antiplane shear in composite regions was discovered by Sinclair [1980]. In both problems, the field equations for the longitudinal displacement, u_z^i, and for the temperature, T^i, are harmonic. As a result, the following correspondences between these two problems can be set down:

$$\begin{aligned}
\nabla^2 T^i = 0 &\iff \nabla^2 u_z^i = 0, \\
q_r^i = -k_i \frac{\partial T^i}{\partial r} &\iff \tau_{rz}^i = G_i \frac{\partial u_z^i}{\partial r}, \\
q_\theta^i = \frac{k_i}{r} \frac{\partial T^i}{\partial \theta} &\iff \tau_{\theta z}^i = \frac{G_i}{r} \frac{\partial u_z^i}{\partial \theta},
\end{aligned} \tag{3-1}$$

where q_r^i and q_θ^i are the heat flux in the radial and circumferential directions and k_i is the thermal conductivity in the i-th material region. Therefore, the analogy is straightforward: the temperature field is analogous to the out-of-plane displacement field, whereas the heat flux components are the analogous counterparts of the stress field components, diverging to infinity as $r \to 0$.

As far as the BCs are concerned, the free-edge conditions (2-10) correspond to insulated edges in diffusion problems, provided that the elastic variables are replaced by the steady-state heat transfer variables according to (3-1). Similarly, the clamped BCs (2-11) in elasticity correspond to zero temperature prescribed along the edges. Finally, the continuity of the longitudinal displacement u_z and of the tangential stress $\tau_{\theta z}$ in (2-12) at the interfaces corresponds to the continuity of temperature, T, and heat-flux, q_θ. The eigenvalue problem for the diffusion problem has therefore the same coefficient matrix as in (2-13).

4. Singularities in the electromagnetic fields

Consider the multimaterial wedge shown in Figure 2. Each material is isotropic and has a dielectric permittivity ϵ_i and a magnetic permeability μ_i. We also admit the presence of a perfect electric conductor (PEC) in the region labeled 1 and defined by the interfaces Γ_1 and Γ_n.
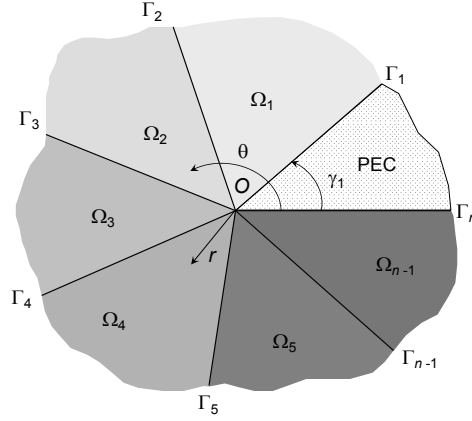
Figure 2. Geometry of a multimaterial wedge including one made of a perfect electric conductor.

For periodic fields with circular frequency ω, the Maxwell's equations for each homogeneous angular domain read as follows [van Bladel 1991]:

$$\mathrm{j}\omega\epsilon_i \boldsymbol{E}^i = \nabla \times \boldsymbol{H}^i, \qquad -\mathrm{j}\omega\mu_i \boldsymbol{H}^i = \nabla \times \boldsymbol{E}^i, \tag{4-1}$$

where $\boldsymbol{E}^i$ and $\boldsymbol{H}^i$ are, respectively, the electric and magnetic fields, and the symbol j stands for the imaginary unit.

In cylindrical coordinates r, θ, z, with the z axis perpendicular to the plane of the wedge, and considering electromagnetic fields independent of z, Maxwell's equations reduce to the following conditions upon the components of the electric and magnetic fields:

$$\begin{aligned}
\mathrm{j}\omega\epsilon_i E_r^i &= \frac{1}{r}\frac{\partial H_z^i}{\partial \theta}, & \mathrm{j}\omega\epsilon_i E_\theta^i &= -\frac{\partial H_z^i}{\partial r}, & \mathrm{j}\omega\epsilon_i E_z^i &= \frac{1}{r}\frac{\partial}{\partial r}(r H_\theta^i) - \frac{1}{r}\frac{\partial H_r^i}{\partial \theta}, \\
-\mathrm{j}\omega\mu_i H_r^i &= \frac{1}{r}\frac{\partial E_z^i}{\partial \theta}, & -\mathrm{j}\omega\mu_i H_\theta^i &= -\frac{\partial E_z^i}{\partial r}, & -\mathrm{j}\omega\mu_i H_z^i &= \frac{1}{r}\frac{\partial}{\partial r}(r E_\theta^i) - \frac{1}{r}\frac{\partial E_r^i}{\partial \theta}.
\end{aligned} \tag{4-2}$$

It is easy to verify that the E_z^i and H_z^i components satisfy the Helmholtz equation [van Bladel 1991]:

$$\begin{aligned}
\frac{\partial^2 E_z^i}{\partial r^2} + \frac{1}{r}\frac{\partial E_z^i}{\partial r} + \frac{1}{r^2}\frac{\partial^2 E_z^i}{\partial \theta^2} + k_i^2 E_z^i &= \nabla^2 E_z^i + k_i^2 E_z^i = 0, \\
\frac{\partial^2 H_z^i}{\partial r^2} + \frac{1}{r}\frac{\partial H_z^i}{\partial r} + \frac{1}{r^2}\frac{\partial^2 H_z^i}{\partial \theta^2} + k_i^2 H_z^i &= \nabla^2 H_z^i + k_i^2 H_z^i = 0,
\end{aligned} \tag{4-3}$$

where $k_i = \omega^2 \epsilon_i \mu_i$.

In close analogy with the antiplane problem in linear elasticity, the following separable form for E_z^i and H_z^i can be postulated for all $(r, \theta) \in \Omega_i$ [Meixner 1972]:

$$E_z^i(r, \theta) = \sum_j r^{\lambda_j} f_{i,j}(\theta, \lambda_j), \quad H_z^i(r, \theta) = \sum_j r^{\lambda_j} F_{i,j}(\theta, \lambda_j) \tag{4-4}$$

where λ_j are the eigenvalues, and $f_{i,j}$, and $F_{i,j}$ are the eigenfunctions.

We can introduce (4-4) into (4-3), obtaining the equalities

$$r^{\lambda_j - 2}\left(\frac{\mathrm{d}^2 f_{i,j}}{\mathrm{d}\theta^2} + \lambda_j^2 f_{i,j}\right) = 0, \quad r^{\lambda_j - 2}\left(\frac{\mathrm{d}^2 F_{i,j}}{\mathrm{d}\theta^2} + \lambda_j^2 F_{i,j}\right) = 0. \tag{4-5}$$

Hence, we find that the eigenfunctions $f_{i,j}$ and $F_{i,j}$ are linear combinations of trigonometric functions, in perfect analogy with the eigenfunction $f_{i,j}$ in antiplane elasticity (see (2-8)):

$$f_{i,j}(\theta, \lambda_j) = A_i \sin(\lambda_j \theta) + B_i \cos(\lambda_j \theta), \quad F_{i,j}(\theta, \lambda_j) = C_i \sin(\lambda_j \theta) + D_i \cos(\lambda_j \theta). \tag{4-6}$$

These eigenfunctions are responsible for the singular behavior of the components E_r^i, E_θ^i, H_r^i and H_θ^i of the electric and magnetic fields near the wedge apex. In particular, from (4-2), we observe that

$$E_r^i = \frac{1}{r\mathrm{j}\omega\epsilon_i}\frac{\partial H_z^i}{\partial \theta} = \frac{1}{\mathrm{j}\omega\epsilon_i}\sum_j \lambda_j r^{\lambda_j - 1} F'_{i,j} \sim O(r^{\lambda_j - 1}),$$

$$E_\theta^i = -\frac{1}{\mathrm{j}\omega\epsilon_i}\frac{\partial H_z^i}{\partial r} = -\frac{1}{\mathrm{j}\omega\epsilon_i}\sum_j \lambda_j r^{\lambda_j - 1} F_{i,j} \sim O(r^{\lambda_j - 1}),$$

$$H_r^i = -\frac{1}{r\mathrm{j}\omega\mu_i}\frac{\partial E_z^i}{\partial \theta} = -\frac{1}{\mathrm{j}\omega\mu_i}\sum_j \lambda_j r^{\lambda_j - 1} f'_{i,j} \sim O(r^{\lambda_j - 1}),$$

$$H_\theta^i = \frac{1}{\mathrm{j}\omega\mu_i}\frac{\partial E_z^i}{\partial r} = \frac{1}{\mathrm{j}\omega\mu_i}\sum_j \lambda_j r^{\lambda_j - 1} f'_{i,j} \sim O(r^{\lambda_j - 1}). \tag{4-7}$$

Hence, $E_z^i \sim O(r^{\lambda_j})$ and $H_z^i \sim O(r^{\lambda_j})$ are the analogous counterparts of u_z^i and remain finite as $r \to 0$. Moreover, the radial components of the electric and magnetic fields, E_r^i and H_r^i, are analogous to $\tau_{\theta z}^i$ and the circumferential components, E_θ^i and H_θ^i, are analogous to τ_{rz}^i. More specifically, we have $E_r^i = \tau_{\theta z}^i/(\mathrm{j}\omega\epsilon_i G_i)$, $H_r^i = -\tau_{\theta z}^i/(\mathrm{j}\omega\mu_i G_i)$, $E_\theta^i = -\tau_{rz}^i/(\mathrm{j}\omega\epsilon_i G_i)$ and $H_\theta^i = \tau_{rz}^i/(\mathrm{j}\omega\mu_i G_i)$. All of these components diverge when $r \to 0$ with a power-law singularity of order $-1 < (\lambda_j - 1) < 0$.

Regarding the BCs, the tangential components of the electric field vanish along the edges Γ_1 and Γ_n of the PEC:

$$E_z^1(r, \gamma_1) = 0, \quad E_r^1(r, \gamma_1) = 0, \quad E_z^{n-1}(r, \gamma_n) = 0, \quad E_r^{n-1}(r, \gamma_n) = 0, \tag{4-8}$$

On the PEC surface also $H_\theta = 0$, but this condition does not need be enforced, since it is a consequence of the previous ones. Along each bimaterial interface ($i = 1, \ldots, n-2$), the tangential components of the electric and magnetic fields are continuous:

$$E_z^i(r, \gamma_{i+1}) = E_z^{i+1}(r, \gamma_{i+1}), \quad E_r^i(r, \gamma_{i+1}) = E_r^{i+1}(r, \gamma_{i+1}),$$
$$H_z^i(r, \gamma_{i+1}) = H_z^{i+1}(r, \gamma_{i+1}), \quad H_r^i(r, \gamma_{i+1}) = H_r^{i+1}(r, \gamma_{i+1}). \tag{4-9}$$

Using the equations (4-7), the BCs (4-8) become

$$E_z^1(r, \gamma_1) = 0, \quad E_z^{n-1}(r, \gamma_n) = 0, \quad \frac{\partial H_z^1}{\partial \theta}(r, \gamma_1) = 0, \quad \frac{\partial H_z^{n-1}}{\partial \theta}(r, \gamma_n) = 0, \tag{4-10}$$

whereas those defined by (4-9) become ($i = 1, \ldots, n-2$)

$$E_z^i(r, \gamma_{i+1}) = E_z^{i+1}(r, \gamma_{i+1}), \quad \frac{1}{\epsilon_i} \frac{\partial H_z^i}{\partial \theta}(r, \gamma_{i+1}) = \frac{1}{\epsilon_{i+1}} \frac{\partial H_z^{i+1}}{\partial \theta}(r, \gamma_{i+1})$$

$$H_z^i(r, \gamma_{i+1}) = H_z^{i+1}(r, \gamma_{i+1}), \quad \frac{1}{\mu_i} \frac{\partial E_z^i}{\partial \theta}(r, \gamma_{i+1}) = \frac{1}{\mu_{i+1}} \frac{\partial E_z^{i+1}}{\partial \theta}(r, \gamma_{i+1}).$$

(4-11)

It is interesting that (4-3), (4-10) and (4-11) can be separated into two independent sets of equations, one involving only H_z and another involving only E_z. Hence, the electromagnetic field for this problem can be decomposed into two distinct independently evolving fields, the so-called transverse electric (TE) and transverse magnetic (TM) fields. The TE (resp. TM) field has vanishing electric (resp. magnetic) but nonzero magnetic (resp. electric) components parallel to the cylinder axis z.

Considering the series expansion for E_z and H_z, along with the expressions for the eigenfunctions $f_{i,j}$ and $F_{i,j}$, the boundary value problem consists of two sets of $2n - 2$ equations in $2n - 1$ unknowns, one for E_z and another for H_z. The former equation set (TM case) involves the coefficients $A_{i,j}$, $B_{i,j}$ and λ_j and can be symbolically written as

$$\boldsymbol{\Lambda} \boldsymbol{v} = \boldsymbol{0},$$

(4-12)

where $\boldsymbol{\Lambda}$ denotes the coefficient matrix which depends on the eigenvalue and $\boldsymbol{v}$ represents the vector which collects the unknowns $A_{i,j}$ and $B_{i,j}$. The coefficient matrix in (4-12) has exactly the same structure as that for the elasticity problem in (2-13), provided that we consider $N_\theta^i = \{\sin(\lambda_j \theta), \cos(\lambda_j \theta)\}$ and we set $G_i = 1/\mu_i$.

The latter equation set (TE case) involves the coefficients $C_{i,j}$, $D_{i,j}$ and λ_j and can be symbolically written as:

$$\boldsymbol{\Lambda} \boldsymbol{w} = \boldsymbol{0},$$

(4-13)

where $\boldsymbol{\Lambda}$ is the coefficient matrix which depends on the eigenvalue and $\boldsymbol{w}$ represents the vector which collects the unknowns $C_{i,j}$ and $D_{i,j}$. Again, the coefficient matrix in (4-13) has exactly the same structure as that for the elasticity problem in (2-13), provided that we consider $N_\theta^i = \{\cos(\lambda_j \theta), -\sin(\lambda_j \theta)\}$ and we set $G_i = 1/\epsilon_i$.

For the existence of nontrivial solutions, the determinants of the coefficient matrices must vanish, yielding two eigenequations that, for given values of ϵ_i and μ_i, determine the eigenvalues λ_j^{TE} and λ_j^{TM}. Hence, this proves that the analysis of the singularities of the electromagnetic field is mathematically analogous to that for the elastic field due to antiplane loading.

5. Singularities in fluid dynamics

In fluid dynamics, a large class of problems can be described by a potential flow. In such cases, a stream function, $\boldsymbol{\Psi}$, can be introduced such that the flow velocity $\boldsymbol{v}$ can be determined from its curl:

$$\boldsymbol{v} = \nabla \times \boldsymbol{\Psi},$$

(5-1)

where, in polar coordinates and for 2D problems, we have $\boldsymbol{v} = (v_r, v_\theta, 0)^T$ and $\boldsymbol{\Psi} = (0, 0, \Psi)^T$. Moreover, if the flow is irrotational, the curl of the velocity is zero [Batchelor 1973]:

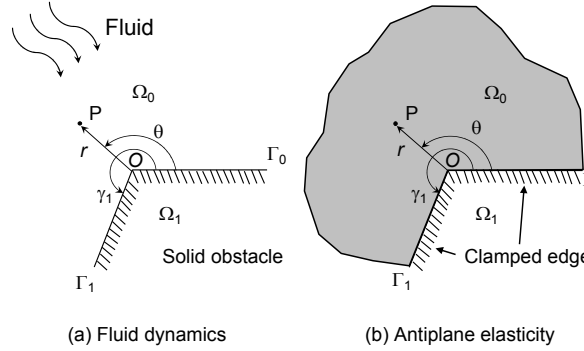$$\nabla \times \boldsymbol{v} = \boldsymbol{0}.$$

(5-2)

Figure 3. Geometry of a flow meeting a sharp obstacle: analogy between fluid dynamics and antiplane elasticity.

As a result, introducing (5-1) into (5-2), after some algebra we obtain the Laplace condition upon the stream function Ψ:

$$\nabla \times (\nabla \times \mathbf{\Psi}) = \nabla^2 \Psi = 0. \tag{5-3}$$

For a problem where a fluid flow meets a sharp obstacle, as shown in Figure 3, the flow may have singularities in the velocities. The power of the singularity can be determined according to an asymptotic analysis, adopting the eigenfunction expansion method for the stream function Ψ:

$$\Psi(r, \theta) = \sum_j r^{\lambda_j} f_j(\theta, \lambda_j), \tag{5-4}$$

where, again, λ_j are the eigenvalues of the problem and $f_j(\theta, \lambda_j)$ are the eigenfunctions. The Laplace condition upon Ψ implies that the functions $f_j(\theta, \lambda_j)$ are a combination of trigonometric functions, as in (2-8):

$$f_j(\theta, \lambda_j) = A_j \sin(\lambda_j \theta) + B_j \cos(\lambda_j \theta). \tag{5-5}$$

The velocity field components in polar coordinates, v_r and v_θ, can be obtained by differentiating the stream function Ψ:

$$v_r = \frac{1}{r}\frac{\partial \Psi}{\partial \theta} = \sum_j r^{\lambda_j - 1} f_j'(\theta, \lambda_j), \quad v_\theta = -\frac{\partial \Psi}{\partial r} = -\sum_j \lambda_j r^{\lambda_j - 1} f_j(\theta, \lambda_j), \tag{5-6}$$

where $f_j'(\theta, \lambda_j)$ are the angular derivatives of the eigenfunctions $f_j(\theta, \lambda_j)$.

Hence, the following correspondences between fluid dynamics and antiplane elasticity can be set down:

$$\begin{aligned}
\nabla^2 \Psi = 0 &\iff \nabla^2 u_z = 0, \\
v_\theta = -\frac{\partial \Psi}{\partial r} &\iff \tau_{rz} = G\frac{\partial u_z}{\partial r}, \\
v_r = \frac{1}{r}\frac{\partial \Psi}{\partial \theta} &\iff \tau_{\theta z} = \frac{G}{r}\frac{\partial u_z}{\partial \theta}.
\end{aligned} \tag{5-7}$$

The boundary conditions along the edges of the obstacle correspond to a vanishing velocity in the direction normal to the interface. According to Figure 3, this corresponds to the following conditions:

$$v_\theta(\theta = 0) = 0, \quad v_\theta(\theta = \gamma_1) = 0. \tag{5-8}$$

Introducing $(5\text{-}6)_2$ into (5-8), as well as the expression of the eigenfunction in (5-5), the boundary conditions (5-8) reduce to

$$f_j(0, \lambda_j) = B_j = 0, \quad f_j(\gamma_1, \lambda_j) = A_j \sin(\lambda_j \gamma_1) + B_j \cos(\lambda_j \gamma_1) = 0. \tag{5-9}$$

Note that the analogous of these boundary conditions in antiplane elasticity corresponds to the clamped boundary conditions (2-11) along Γ_0 and Γ_1 (see Figure 3). Therefore, it is possible to state that the singularities in fluid dynamics correspond to those in antiplane elasticity, provided that clamped-clamped BCs are considered along the edges of the elastic wedge. More specifically, since $(5\text{-}9)_1$ leads to $B_j = 0$ and we are looking for nontrivial solutions, the eigenequation of the problems is

$$\sin(\lambda_j \gamma_1) = 0, \tag{5-10}$$

from which we determine the lowest eigenvalue

$$\lambda_1 = \frac{\pi}{\gamma_1}. \tag{5-11}$$

Hence, a singularity exists ($\lambda_1 < 1$) for $\gamma_1 > \pi$, i.e., when the fluid domain Ω_1 presents a reentrant corner. Some notable eigenvalues can be found in [Batchelor 1973] and match exactly those provided in [Sinclair 1980] for the antiplane problem of a single material wedge. For instance, a flow around a right corner ($\gamma_1 = 3\pi/2$) has $\lambda_1 = \frac{2}{3}$, whereas a uniform flow ($\gamma_1 = \pi$) is nonsingular. The case of a flow around a thin obstacle ($\gamma_1 = 2\pi$) leads to $\lambda_1 = \frac{1}{2}$, as for an anticrack (rigid line inclusion) in antiplane elasticity [dal Corso et al. 2008; Bigoni et al. 2008].

6. Discussion and conclusion

In the present paper, we have compared and unified the mathematical formulations for the asymptotic characterization of the singular fields at multimaterial wedges in antiplane elasticity, diffusion problems, and electromagnetic diffraction. For a single and homogeneous material sector, we have also established the analogy between fluid dynamics and antiplane elasticity.

The asymptotic analysis of the stress singularities at the vertex of multimaterial wedges and junctions in antiplane elasticity is perfectly analogous to the corresponding diffusion problem. The temperature field plays the same role as the out-of-plane displacement field and the heat fluxes correspond to the tangential stresses. On the other hand, the analogy with electromagnetism is more complex. In particular, when an isotropic multimaterial wedge with PEC boundaries is considered, we have shown that two independent problems can be defined, one for TE fields, associated to an eigenequation for H_z, and one for TM fields, associated to an eigenequation for E_z. The eigenequation for E_z corresponds exactly to that obtained for the same geometrical configuration in antiplane elasticity by setting $G_i = 1/\mu_i$ and replacing the PEC region with an infinitely stiff material leading to clamped edge BCs along Γ_1 and Γ_n. Similarly, the other eigenequation for H_z can be obtained in antiplane elasticity for the same geometrical

configuration by setting $G_i = 1/\epsilon_i$ and replacing the PEC region with an infinitely soft material leading to stress-free BCs along Γ_1 and Γ_n.

As far as the analogy for fluid dynamics is concerned, we have found that the stream function Ψ plays the same role as u_z and that the singularities in fluid dynamics correspond to those in antiplane elasticity, provided that clamped-clamped BCs are considered along the edges of the elastic wedge. These analogical results are also important for stress concentration problems. For instance, in the case of a flux around a circular cylinder, the velocity concentration factor is equal to 2, that is the velocity of the fluid at the border of the cylinder is twice higher than the velocity at infinity. Analogously, the stress concentration factor for a plate in uniaxial tension with an infinitely stiff round inclusion tends also to 2 [Duan et al. 2005].

Finally, the presence of singularities in diffusion, electromagnetism, and fluid dynamics suggests extending to these fields the dimensional analysis considerations we have proposed for the scaling of structural strength in [Carpinteri 1981; 1982a; 1982b; 1983; 1987; Carpinteri and Paggi 2006; 2009]. In solid mechanics, the presence of a stress singularity of power $\lambda - 1$ requires the use of a generalized stress-intensity factor, K^*, which has anomalous physical dimensions:

$$\sigma_{ij} = K^* r^{\lambda-1} f_{ij}(\theta, \lambda) \Rightarrow K^* = \sigma b^{1-\lambda} g, \tag{6-1}$$

where σ is a nominal applied stress, b is the characteristic structural size, and g is a shape factor depending on the geometry of the structure and the topology of the junction. The same reasoning can therefore be applied to the other analogous fields, defining generalized heat-intensity, electromagnetic-intensity, and velocity-intensity factors. In structural mechanics, the anomalous physical dimensions of K^* lead to size-scale effects on the nominal strength, i.e., the material strength becomes a function of the structural size [Carpinteri 1987; Carpinteri and Paggi 2006]. Therefore, size-scale effects in diffusion, electromagnetism, and fluid dynamics are also expected and can be analyzed using the same mathematical formalism as in structural mechanics.

Finally, it is well-known that elastic singularities are usually a mathematical artifact and that plasticity relieves the singularities in elastoplastic materials. Similarly to plasticity, a saturation of the electromagnetic fields at the sharp tip of an antenna is experimentally found in electromagnetism. In fluid dynamics, a possible analogous mechanism could be the occurrence of turbulence. In structural mechanics, the competition between crack propagation and plastic flow collapse is ruled by the brittleness number s (see [Carpinteri 1981; 1982a; 1982b] for more details):

$$s = \frac{K_{IC}}{\sigma_c b^{1/2}}, \tag{6-2}$$

where K_{IC}, σ_c and b are, respectively, the fracture toughness, the material strength, and the structural size. Brittle failure, characterized by *small scale yielding*, takes place when the brittleness number is lower than a threshold value. Above that, *large scale yielding* takes place and plastic flow collapse prevails over brittle failure. Generalizing these concepts to fluid dynamics, it is possible to define a similar dimensionless number in the case of a fluid against a thin obstacle. We shall call it *turbulence number*, s_t, and it is given by:

$$s_t = \frac{K_C}{v_c b^{1/2}}, \tag{6-3}$$

where K_C and v_c are, respectively, a critical velocity-intensity factor and a critical velocity for the appearance of turbulence. This resembles the Reynolds number, $\mathrm{Re} = v_c b / \nu$, where ν is the kinematic viscosity of the fluid, although the proposed turbulence number comes directly from the presence of singularities. The use of s_t in addition to the Reynolds number has never been explored so far. Therefore, further developments of this work will regard the assessment of the applicability of these concepts to interpret the laminar-turbulent flow transition in fluid mechanics, a still open problem nowadays.

References

[Batchelor 1973] G. K. Batchelor, *An introduction to fluid dynamics*, Cambridge University Press, Cambridge, 1973.

[Bigoni et al. 2008] D. Bigoni, F. dal Corso, and M. Gei, "The stress concentration near a rigid line inclusion in a prestressed, elastic material, II: Implications on shear band nucleation, growth and energy release rate", *J. Mech. Phys. Solids* **56**:3 (2008), 839–857.

[van Bladel 1991] J. van Bladel, *Singular electromagnetic fields and sources*, Clarendon, Oxford, 1991.

[Bouwkamp 1946] C. J. Bouwkamp, "A note on singularities occurring at sharp edges in electromagnetic diffraction theory", *Physica* **12**:7 (1946), 467–474.

[Carpinteri 1981] A. Carpinteri, "Size effect in fracture toughness testing: a dimensional analysis approach", pp. 785–797 in *Analytical and experimental fracture mechanics* (Rome, 1980), edited by G. C. Sih and M. Mirabile, Sijthoff & Noordhoff, Alphen an den Rijn, 1981.

[Carpinteri 1982a] A. Carpinteri, "Experimental determination of fracture toughness parameters K_{IC} and J_{IC} for aggregative materials", pp. 1491–1498 in *Advances in fracture research: proceedings of ICF5* (Cannes, 1981), edited by D. François, Pergamon, Oxford, 1982.

[Carpinteri 1982b] A. Carpinteri, "Notch sensitivity in fracture testing of aggregative materials", *Eng. Fract. Mech.* **16**:4 (1982), 467–481.

[Carpinteri 1983] A. Carpinteri, "Plastic flow collapse vs. separation collapse (fracture) in elastic-plastic strain-hardening structures", *Mater. Struct.* **16**:2 (1983), 85–96.

[Carpinteri 1987] A. Carpinteri, "Stress-singularity and generalized fracture toughness at the vertex of re-entrant corners", *Eng. Fract. Mech.* **26**:1 (1987), 143–155.

[Carpinteri and Paggi 2006] A. Carpinteri and M. Paggi, "Influence of the intermediate material on the singular stress field in trimaterial junctions", *Mater. Sci.* (*Russia*) **42**:1 (2006), 95–101.

[Carpinteri and Paggi 2009] A. Carpinteri and M. Paggi, "Asymptotic analysis in linear elasticity: from the pioneering studies by Wieghardt and Irwin until today", *Eng. Fract. Mech.* **76**:12 (2009), 1771–1784.

[dal Corso et al. 2008] F. dal Corso, D. Bigoni, and M. Gei, "The stress concentration near a rigid line inclusion in a prestressed, elastic material, I: Full-field solution and asymptotics", *J. Mech. Phys. Solids* **56**:3 (2008), 815–838.

[Dean and Montagnon 1949] W. R. Dean and P. E. Montagnon, "On the steady motion of viscous liquid in a corner", *Math. Proc. Cambridge Philos. Soc.* **45**:3 (1949), 389–395.

[Duan et al. 2005] H. L. Duan, J. Wang, Z. P. Huang, and Y. Zhong, "Stress fields of a spheroidal inhomogeneity with an interphase in an infinite medium under remote loadings", *Proc. R. Soc. Lond. A* **461**:2056 (2005), 1055–1080.

[Duan et al. 2006] H. L. Duan, B. L. Karihaloo, J. Wang, and X. Yi, "Effective conductivities of heterogeneous media containing multiple inclusions with various spatial distributions", *Phys. Rev. B* **73**:17 (2006), 174203.

[England 1971] A. H. England, "On stress singularities in linear elasticity", *Int. J. Eng. Sci.* **9**:6 (1971), 571–585.

[Fenner 1976] D. N. Fenner, "Stress singularities in composite materials with an arbitrarily oriented crack meeting an interface", *Int. J. Fract.* **12**:5 (1976), 705–712.

[Hashin and Shtrikman 1962] Z. Hashin and S. Shtrikman, "A variational approach to the theory of the effective magnetic permeability of multiphase materials", *J. Appl. Phys.* **33**:10 (1962), 3125–3131.

[Ma and Hour 1989] C.-C. Ma and B.-L. Hour, "Analysis of dissimilar anisotropic wedges subjected to antiplane shear deformation", *Int. J. Solids Struct.* **25**:11 (1989), 1295–1308.

[Ma and Hour 1990] C.-C. Ma and B.-L. Hour, "Antiplane problems in composite anisotropic materials with an inclined crack terminating at a bimaterial interface", *Int. J. Solids Struct.* **26**:12 (1990), 1387–1400.

[Meixner 1972] J. Meixner, "The behavior of electromagnetic fields at edges", *IEEE Trans. Antenn. Propag.* **20**:4 (1972), 442–446.

[Pageau et al. 1995] S. S. Pageau, P. F. Joseph, and S. B. Biggers, Jr., "Singular antiplane stress fields for bonded and disbonded three-material junctions", *Eng. Fract. Mech.* **52**:5 (1995), 821–832.

[Paggi and Carpinteri 2008] M. Paggi and A. Carpinteri, "On the stress singularities at multimaterial interfaces and related analogies with fluid dynamics and diffusion", *Appl. Mech. Rev. (ASME)* **61**:2 (2008), 020801.

[Paggi et al. 2009] M. Paggi, A. Carpinteri, and R. Orta, "A unified mathematical formulation for the asymptotic analysis of singular elastic and electromagnetic fields", paper # 141 in *AIMETA 2009: atti del XIX Congresso dell'associazione italiana di meccanica teorica e applicata* (Ancona, 2009), edited by S. Lenci, Università Politecnica delle Marche, Facoltà di Ingegneria, Ancona, 2009.

[Paggi et al. 2010] M. Paggi, A. Carpinteri, and R. Orta, "A mathematical analogy and a unified asymptotic formulation for singular elastic and electromagnetic fields at multimaterial wedges", *J. Elasticity* **99**:2 (2010), 131–146.

[Rao 1971] A. K. Rao, "Stress concentrations and singularities at interface corners", *Z. Angew. Math. Mech.* **51**:5 (1971), 395–406.

[Sinclair 1980] G. B. Sinclair, "On the singular eigenfunctions for plane harmonic problems in composite regions", *J. Appl. Mech. (ASME)* **47**:1 (1980), 87–92.

[Sinclair 2004a] G. B. Sinclair, "Stress singularities in classical elasticity, I: Removal, interpretation, and analysis", *Appl. Mech. Rev. (ASME)* **57**:4 (2004), 251–297.

[Sinclair 2004b] G. B. Sinclair, "Stress singularities in classical elasticity, II: Asymptotic identification", *Appl. Mech. Rev. (ASME)* **57**:5 (2004), 385–439.

[Williams 1952] M. L. Williams, "Stress singularities resulting from various boundary conditions in angular corners of plates in extension", *J. Appl. Mech. (ASME)* **19** (1952), 526–528.

ALBERTO CARPINTERI: alberto.carpinteri@polito.it
Department of Structural and Geotechnical Engineering, Politecnico di Torino, Corso Duca degli Abruzzi, 24, I-10129 Torino, Italy
http://staff.polito.it/alberto.carpinteri

MARCO PAGGI: marco.paggi@polito.it
Department of Structural and Geotechnical Engineering, Politecnico di Torino, Corso Duca degli Abruzzi, 24, I-10129 Torino, Italy
http://staff.polito.it/marco.paggi

■msp

DEEP PENETRATION AND LIQUID INJECTION INTO ADIPOSE TISSUE

Kerstyn Comley and Norman Fleck

The subcutaneous injection of porcine adipose tissue by a hypodermic needle involves two stages: tissue penetration followed by the delivery of liquid into the tissue. The force required to penetrate adipose tissue by a series of conically tipped and flat-bottomed circular punches has been measured. Scanning electron microscopy and light microscopy are used to observe the mechanism of crack formation during penetration. The experiments reveal that penetration by either a flat bottomed or 45° conically tipped punch involves the formation of a mode II ring crack. The predicted penetration pressure according to the Shergold–Fleck model (Proc. R. Soc. Lond. A **460** (2004), 3037–3058) is in good agreement with the measured pressure on the punch. The subsequent delivery of liquid into adipose tissue by the hypodermic needle has also been examined: the injection pressure for phosphate buffered saline has been measured for a range of flow rates. X-ray images of the injected liquid suggest that micro-cracks are formed by the fluid pressure within the tissue and this leads to an increase in permeability. A seepage model is developed, based on the Darcy flow law, to relate the volumetric flow rate to the injection delivery pressure. Finally, a model of hydraulic fracture is used to assess the toughness associated with the formation of the micro-cracks during injection.

1. Introduction

A wide range of medical devices exist for the injection of liquid-based drugs into the subcutaneous layer of adipose tissue. These include a hypodermic needle and a needle-free syringe. Injection involves two stages: (i) a needle or liquid jet penetrates the tissue to a desired depth (intradermal, subcutaneous, intramuscular and so on); and (ii) the liquid drug is delivered from the syringe into the tissue. Needle-free devices achieve control of tissue penetration and liquid delivery through control of the pressure versus time characteristic of the jet. For example, when subcutaneous injection occurs by a needle-free device, the dermal and subcutaneous layers are first penetrated by a high pressure jet (static head of order 25 MPa for 5 ms). Subsequent delivery of the bulk of the liquid into the subcutaneous layer follows at a lower jet pressure (static head of approximately 5 MPa) and typically takes 100 ms [Shergold et al. 2006].

Consider first the penetration stage. Shergold and Fleck [2004] have recently developed two penetration models for soft solids; a model based on the formation of a planar crack and a model based on the formation of a ring-crack. They demonstrated that the force to penetrate dermis with a sharp metal punch depends upon the Young's modulus and toughness of the tissue, and upon the diameter and tip geometry of the punch. They argue that both a sharp tipped needle and a liquid jet penetrate a soft solid by the formation and opening of a planar crack, whereas penetration by a flat-bottomed punch involves the formation of a ring crack. The predicted pressure for penetration of the dermis by a liquid jet is of similar value to the pressures employed in needle-free injectors [Shergold et al. 2006]. The

---

Keywords: adipose tissue, soft solids, deep penetration, injection, fracture mechanics, toughness.

Shergold-Fleck model gives the penetration pressure but no explicit information on the penetration depth. The depth of penetration achieved by high speed liquid jet injections has been measured by Baxter and Mitragotri [2005]: The depth of penetration into dermis and subcutaneous adipose tissue was a function of both the diameter and velocity of the jet.

Second, consider the liquid-delivery stage. To date, research into the mechanisms governing the delivery of liquid injected into soft tissue has been confined to an examination of the depth of the injection and the degree of dispersion of the liquid within the tissue. Differences in the dispersion behaviour of liquid injected at low and high speed may affect the rate at which a drug is absorbed and may result in the formation of different types of wound within the tissue. Cooke et al. [1980] have compared the bioavailability from a needle-free injection and from a hypodermic needle and syringe for subcutaneous and intramuscular injections of lignocane into humans. For both injection methods, the injectate was dispersed to a similar degree within the tissue. More recently, Schramm-Baxter and Mitragotri [2004] have performed liquid jet injections into human dermis and observed that the shape and dispersion of the liquid depends upon nozzle diameter and upon the jet velocity. Barry et al. [1995] have developed a poroelastic model for the seepage of a liquid into a spherical cavity, and have included the effect of the fluid pressure upon the small-strain elastic deformation of the porous medium. The relevance of this model to fluid delivery must await an experimental observation of the injection mechanism.

Scope of the paper.   The aim of the current study is to investigate the penetration and delivery of liquid into subcutaneous adipose tissue. We begin by reviewing the microstructure and mechanical properties of adipose tissue. The study is split into two parts: penetration of adipose tissue by a solid punch, and liquid injection by hypodermic syringe. In Section 2, penetration tests by conically tipped and flat-bottom punches are reported, and scanning electron microscopy, light microscopy are used to reveal the mechanisms of penetration. The results are compared with existing models in Section 3. The response of adipose tissue to hypodermic injection is described in Section 4. X-ray techniques are used to reveal the mechanisms of liquid delivery and are combined with the results from the injection tests to motivate a seepage model in Section 5. Concluding remarks are stated in Section 6.

Review of the microstructure and mechanical properties of adipose tissue.   Subcutaneous adipose tissue is a connective tissue comprising lipid-filled cells called *adipocytes* and resides directly under the dermal layer of human skin. The lipid is a triacylglyceride of molecular weight on the order of $900 \, \text{g mol}^{-1}$ and viscosity of $40 \, \text{mPa s}$ at $37° \text{C}$ [Comley and Fleck 2010a]. The adipocytes are of diameter $80 \, \mu\text{m}$ and are supported by two interpenetrating collagen-based structures: (i) a *reinforced basement membrane* in the form of a collagen mesh, containing primarily type I and IV collagen, surrounds each cell, and (ii) a type I collagen fibre network, termed the *interlobular septa* [Björntorp and Martinsson 1966; Nakajima et al. 1998]. The reinforced basement membrane resembles a closed cell foam with a unit cell dimension of $80 \, \mu\text{m}$ and walls of thickness $2 \, \mu\text{m}$. In contrast, the interlobular septa take the form of an open cell foam with cell edge length $1 \, \text{mm}$ and a cell wall thickness of $10 \, \mu\text{m}$ [Comley and Fleck 2010a]. The macroscopic stiffness and toughness derive primarily from the reinforced basement membrane.

Additional structures such as blood vessels exist within the tissue, but have negligible influence on the overall mechanical properties. The intervening space is filled with *ground substance*. Overall, 60–80% (by mass) of adipose tissue is lipid, 5–30% is water and the remaining 2–3% is protein [Greenwood and Johnson 1983]. The large liquid content enforces material incompressibility [Samani et al. 2003].

Histology of adipose tissue suggests that it is approximately isotropic in structure and is thereby isotropic in mechanical properties, to a first approximation [Samani et al. 2003; Comley and Fleck 2010a].

Uniaxial compression tests of adipose tissue suggest that at low strain rates (on the order of $10^{-3}\,\mathrm{s}^{-1}$) the tissue has a Young's modulus of approximately $E = 1\,\mathrm{kPa}$, whereas at strain rates of order $1000\,\mathrm{s}^{-1}$ the modulus increases by more than three orders of magnitude to $E = 3\,\mathrm{MPa}$ [Miller-Young et al. 2002; Nightingale et al. 2003; Gefen and Haberman 2007; Comley and Fleck 2009]. Comley and Fleck [2011] demonstrates that the stress versus strain behaviour of adipose tissue can be adequately described by a one term Ogden strain energy density function [1972] with a shear modulus $\mu = E/3$ and strain hardening exponent $\alpha = 20$ of the form

$$\phi = \frac{2\mu}{\alpha^2}(\lambda_1^\alpha + \lambda_2^\alpha + \lambda_3^\alpha - 3) \tag{1}$$

Here, ϕ is the strain energy density per undeformed unit volume and λ_i are the three principal stretch ratios.

2. Penetration tests using a solid punch

Test method. Deep penetration tests were performed on samples of adipose tissue using circular, cylindrical punches of conical and flat bottomed headshape. Fresh porcine adipose tissue from the jowl of a pig was obtained (Dalehead Foods, Linton, Cambridgeshire, UK), cut into rectilinear blocks of approximate dimension $30\,\mathrm{mm} \times 30\,\mathrm{mm}$ by $10\,\mathrm{mm}$ thick and stored in phosphate buffered saline (PBS) prior to testing. Testing always commenced within four hours of slaughter.

Cylindrical steel punches of diameter in the range $2R = 0.4\,\mathrm{mm}$ to $2R = 3.9\,\mathrm{mm}$ were machined with either a flat bottom or a conical tip of semi-included apex angle $45°$. The punches were then mounted on the cross-head of a screw-driven tensile test machine. A block of adipose tissue was placed on a PMMA plate (thickness $10\,\mathrm{mm}$), placed in turn on the test-machine platen. The penetration force F versus displacement u was measured at selected displacement rates in the range $0.05\,\mathrm{mm\,s}^{-1}$ to $10\,\mathrm{mm\,s}^{-1}$. The cross-head was displaced until the punch made contact with the PMMA plate. Three tests were performed for each configuration of punch and displacement speed. Measurements of the friction force between the side wall of the punch and the tissue following full penetration of the specimen were additionally made as follows. A specimen was placed on a PMMA block containing a through hole of diameter three times that of the punch. The centre-line of the hole was in line with that of the punch and the punch was depressed until its tip penetrated the lower surface of the tissue and passed into the hole in the support block. Continued pushing of the punch through the tissue required a friction force F_{fric}.

Force versus displacement results. A representative set of F versus u results for punches with a diameter $2R = 0.6\,\mathrm{mm}$ is shown in Figure 1. The response is qualitatively the same for the flat-tipped punch and the punch with the $45°$ conical tip. For both headshapes, the force increases steadily with punch displacement until the front surface of the tissue is penetrated. After penetration of the front face, the punch enters the tissue and the force oscillates about a mean (plateau) value F_p. These oscillations in force are associated with the propagation of the crack ahead of the punch tip. Similar profiles of force versus displacement were seen for all punch diameters tested, according to their tip type. Displacement rate was not found to have a significant effect on the measured force. These additional curves are omitted for the sake of brevity.
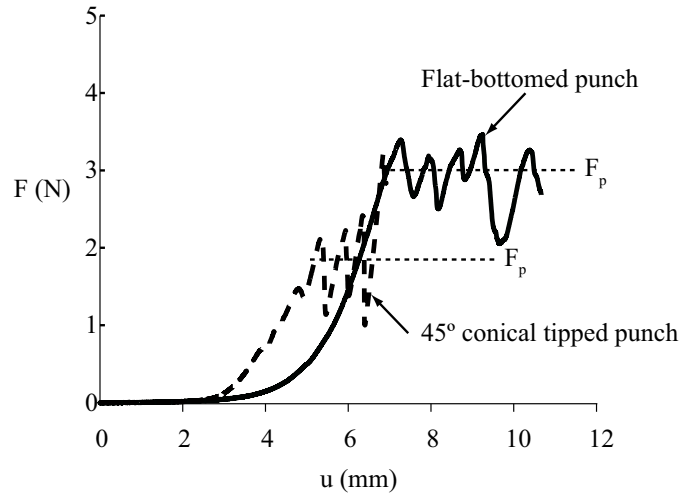
Figure 1. Force F versus displacement u for the penetration of porcine adipose tissue
with 0.6 mm diameter metal punches with either a flat-bottom (solid line) or a 45° conical
tip (dashed line). The displacement rate was 0.1 mm s^{-1}.

The frictional force F_{fric} between the shaft of the punch and the tissue, following full penetration of
the specimen, was less than 5% of the average plateau force F_p. This is consistent with the observation
that the penetration force oscillates about a mean value F_p, with no observable rise in F_p due to frictional
effects.

Observations of hole formation. Observations of the hole formation in the tissue were made as follows.
Upon removal of the punch the surfaces of a selection of samples were examined with an optical micro-
scope. A separate set of samples were fixed in glutaraldehyde, sectioned along the diametral plane of the
hole and prepared for viewing in a scanning electron microscope (SEM). See [Comley and Fleck 2010a]
for full details of sample preparation.

The hole generated by each punch was observed by both light microscopy and a scanning electron
microscope (SEM). The residual hole at the front face of the tissue is shown in Figure 2, for a range
of punch diameters and headshapes. In all cases, the residual hole was almost circular in shape and of
diameter about 20% less that that of the punch.

SEM images of the fracture surface following penetration are shown in Figure 3 for both the flat-
bottomed punch and the 45° conical tipped punch. Adipocyte cells have ruptured and septa fibres have
pulled out and torn, as labelled in the figure. Reinforced membrane and septa fibres are compacted at
the bottom of each hole.

3. Discussion of the penetration behaviour

The nominal penetration pressure p is now examined for the full range of punch diameters and headshapes
that have been tested. Penetration pressure is a useful quantity for comparison with the penetration models
of [Shergold and Fleck 2004] and for comparison with the pressures required for liquid injection. It is
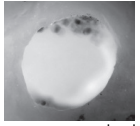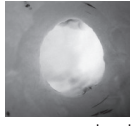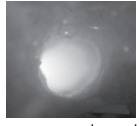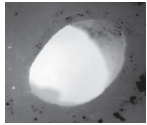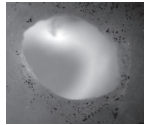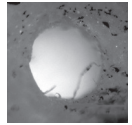
Punch diameter

Punch tip	3.9 mm	2.4 mm	0.6 mm
Flat	1 mm	1 mm	0.5 mm
45º Conical	1 mm	1 mm	0.5 mm

Figure 2. Images of the puncture site at the tissue surface following puncture by a range of metal punches of varying headshape at $0.1\,\mathrm{mm\,s^{-1}}$.
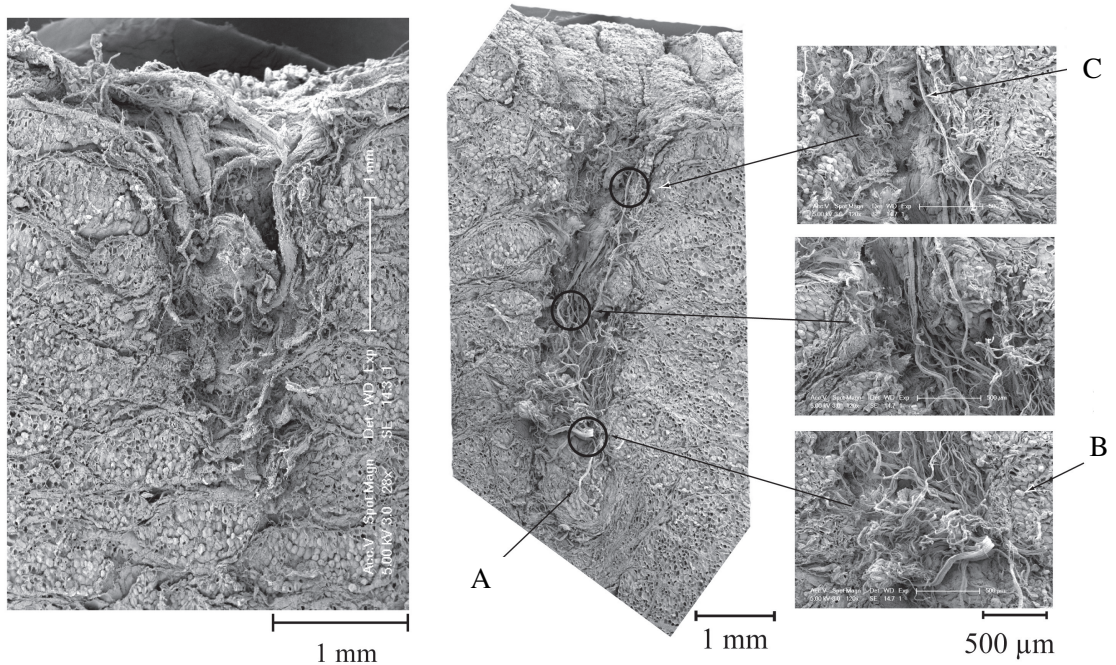

Figure 3. SEM images of the fracture surface of adipose tissue punctured by two punches of diameter 2.4 mm and tip geometry. Left: 45° conical tipped. Middle and right: flat-bottomed (A: compacted interlobular septa and reinforced basement membrane; B: interlobular septa fibres; C: reinforced basement membrane surrounding an adipocyte.
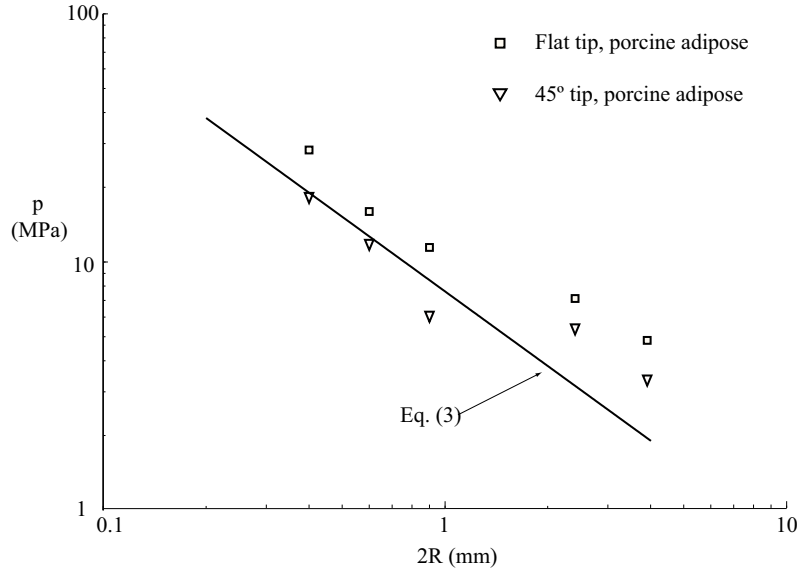
Figure 4. Comparison of the average pressure p required to penetrate adipose tissue with a range of solid punches of varying headshape. A prediction is shown for the pressure required to puncture adipose tissue with a flat tipped punch calculated from (3). The following parameter values were used in the calculations: $J_C = 4.1\,\text{kJ m}^{-2}$, $b/R = 0.5$, $a/R = 0.8$, $\mu = E/3 = 0.3\,\text{kPa}$.

defined by

$$p = \frac{F_p}{\pi R^2} \tag{2}$$

in terms of punch force F_p and punch radius R. For each headshape, the average penetration pressure p (taken from the pooled set of pressure results across all the displacement rates tested) is plotted as a function of punch diameter $2R$ (see Figure 4). For a given punch diameter $2R$ the penetration pressure for the flat-bottomed punch exceeds that for the $45°$ tipped punch.

We note in passing that the pressure to penetrate adipose tissue with a flat-bottomed punch is half the magnitude of the pressure to penetrate dermal tissue (a related collagen based connective tissue) with a similar punch. Pressures of 60 MPa were recorded in [Shergold and Fleck 2005] during tests *in vivo* on human dermis with a 0.3 mm diameter flat-bottomed punch. This compares to 30 MPa needed to penetrate adipose with a 0.4 mm diameter flat-tipped punch.

Shergold and Fleck [2004] have developed two models for the deep penetration of a soft solids (see Section 1). In order to select the appropriate Shergold-Fleck penetration model for comparison with the measured penetration pressures it is necessary to determine the mode of hole formation, during penetration. The light microscopy and SEM images of Figures 2 and 3 suggest that both the flat-bottomed and the $45°$ conical tipped punches compact a circular cylindrical slug of tissue beneath the punch (see Figures 3 and 5a). This indicates that a mode II ring crack is generated in adipose tissue during penetration by either a flat-bottomed or $45°$ conical tipped punch. The characteristics of the puncture force response, shown in Figure 1, are consistent with force response observed in [Shergold et al. 2006] for punches that
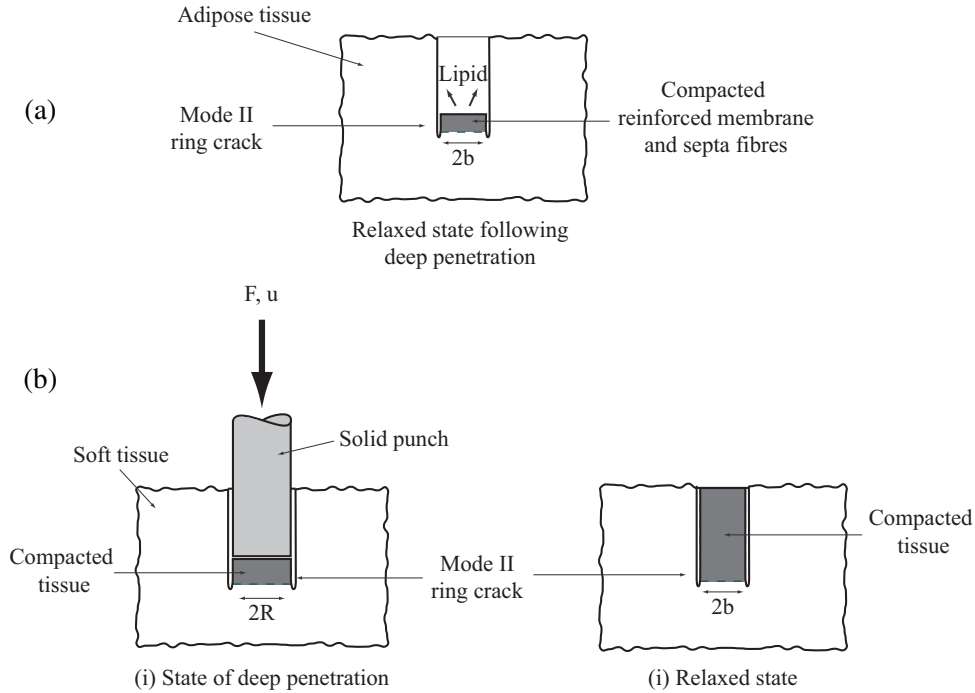
Figure 5. Sketches of (a) the hole formation in adipose tissue following deep penetration by a flat tipped or 45° conical tipped punch, and (b) the hole formation of the Shergold-Fleck ring-crack model.

generate ring cracks. Also the lack of a frictional component to the measured penetration force provides further evidence of ring-cracking. It now remains to briefly review the model and then compare the predictions from the model with the observed penetration pressures reported in Figure 4.

Application of the Shergold-Fleck ring-cracking model. Shergold and Fleck [2004] observed that a flat-bottomed punch generated a mode II ring crack below the punch. Their model is sketched in Figure 5b. They predicted that the pressure p_f scaled with the toughness J_C and shear modulus μ according to

$$\frac{p_F}{\mu} = \left[1 - \left(\frac{b}{R}\right)^2\right]^{-1}\left[\frac{2b}{R}\frac{J_C}{\mu R} + f\left(\frac{b}{R}\right)\right] \tag{3}$$

where b is the radius of the ring crack in the relaxed configuration (Figure 5b). The function $f(b/R)$ does not exist in closed form but is defined in Equation (3.17) of [Shergold and Fleck 2004]; the details are omitted here. The pressure given by (3) attains a minimum for a particular value of b/R and, following that paper, we take the minimum value to be the penetration pressure for ring cracking. Upon assuming that $J_C = 4.1\,\text{kJ}\,\text{m}^{-2}$, from the trouser-tear tests of [Comley and Fleck 2010b], and taking $\mu = E/3 = 0.3\,\text{kPa}$ and $\alpha = 20$ (as discussed in the introduction) we find that as $2R$ increases from 0.4 mm to 3.9 mm, b/R increases from approximately 0.48 to 0.53, whereas p_f decreases from 17 MPa to 3 MPa. For simplicity a single value of $b/R = 0.5$ is used for the remainder of the study.

The Shergold-Fleck model is now compared to the measured data. Using the values for toughness J_C, shear modulus μ and b/R given above predictions for the penetration pressure calculated from (3) are compared to the measured data in Figure 4. The ring-cracking model is in good agreement with the experimental data for punches of diameter $2R$ less than 1 mm. The pressure to penetrate adipose tissue with punches of diameter greater than 1 mm is underestimated by the ring-cracking model. Further measurements from punches of diameter greater than 10 mm are required in order to establish whether this discrepancy is a statistical artefact or a physical effect.

Finally the prediction of $b/R = 0.5$ can be compared to the measured diameter of the holes, created by punch penetration, divided by the corresponding punch radius. Recall that the diameters of the holes were on average 80% smaller than the corresponding punch radius, suggesting an observed b/R of 0.8. The model assumes non-linear elastic behaviour whereas viscoelastic effects would lead to stress relaxation within the adipose tissue in the vicinity of the penetrator. It is unclear, however, whether viscoelastic effects are significant in these tests, as no time dependence of residual hole diameter was observed over a period of minutes following the penetration test. In a parallel study, Comley and Fleck [2009] have explored the sensitivity of uniaxial response of adipose tissue to strain rate. They found that the response is independent of strain rate in the range 10^{-4} to $1\,\mathrm{s}^{-1}$, implying that viscoelastic effects are minor over this regime.

4. Injection tests

Test methods. Measurements were made of the force to inject PBS into blocks of adipose tissue (dimension $20\,\mathrm{mm} \times 20\,\mathrm{mm} \times 10\,\mathrm{mm}$) using a hypodermic syringe. A plastic 5 ml syringe was clamped directly beneath the cross-head of a screw driven tensile test machine (see Figure 6). The syringe body was clamped at the top, above the rubber seal of the plunger, to prevent distortion of the portion of the syringe barrel containing the PBS. The syringe was fitted with either a gauge 27 hypodermic needle (internal orifice diameter $D = 0.21\,\mathrm{mm}$, needle length $L = 20\,\mathrm{mm}$) or a gauge 21 hypodermic needle (internal orifice of $D = 0.51\,\mathrm{mm}$, needle length $L = 25\,\mathrm{mm}$). The internal cross-sectional area of the
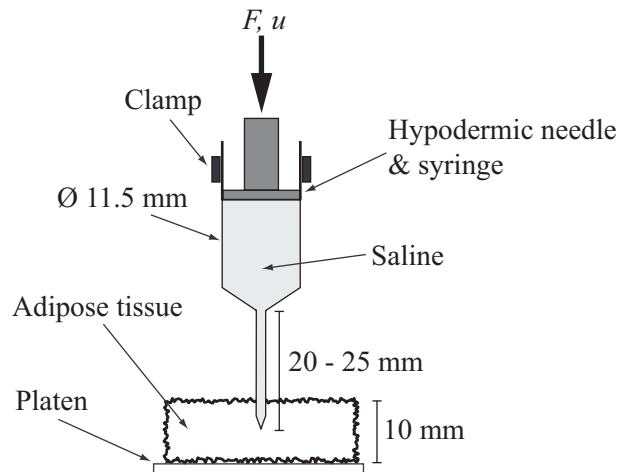

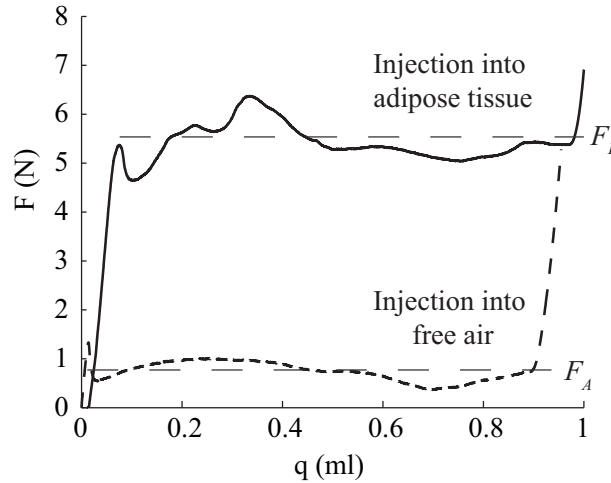**Figure 6.** Experimental setup for injection tests into adipose tissue.

Figure 7. Force F versus the volume of saline injected q using a 0.4 mm diameter hypodermic syringe with a cross head displacement rate of $0.1\,\mathrm{mm\,s^{-1}}$.

syringe was $A_S = 103\,\mathrm{mm^2}$. In each injection test, the plunger was depressed into the syringe by the cross-head and the force F versus cross head displacement u were recorded in two stages, as follows:

<u>Stage 1:</u> The force to inject 1 ml PBS into free air was measured.

<u>Stage 2:</u> The syringe was re-filled with PBS. A fresh piece of adipose tissue was placed on the platen of the test-machine. The clamped syringe was manually lowered to insert the needle into the tissue to a depth of 4 mm (see Figure 6). The force to inject 1 ml PBS into the tissue was then measured.

Tests were conducted at a cross-head displacement speed in the range of $0.05\,\mathrm{mm\,s^{-1}}$ to $1\,\mathrm{mm\,s^{-1}}$. A fresh needle and syringe was used for each test.

Test results. Representative results for both stage 1 and stage 2 of an injection test are shown in Figure 7, for a 27 gauge needle and a cross-head displacement rate of $0.1\,\mathrm{mm\,s^{-1}}$. Injection into air requires a level of force F_A, equal to 1 N for the data shown. In contrast, the force to inject PBS into the tissue increases with injected volume to an approximately constant level F_I, equal to 5.5 N for the data shown. The pressure to inject into adipose tissue p_D is calculated as

$$p_D = \frac{(F_I - F_A)}{A_S} \tag{4}$$

where A_S is the cross-sectional area of the syringe. The dependence of p_D upon the volumetric flow rate $\dot{q}$ is shown in Figure 8. A least squares regression reveals a linear relation between p_D and $\dot{q}$ of the form

$$p_D = c\dot{q} + p_0 \tag{5}$$

where c and p_0 are independent of the needle diameter. Values for the coefficients were found to be $c = 740\,\mathrm{GPa\,s\,m^{-3}}$ and $p_0 = 23\,\mathrm{kPa}$, with a measure of fit $R^2 = 0.9$. An explanation for the origins of (5) is given below.
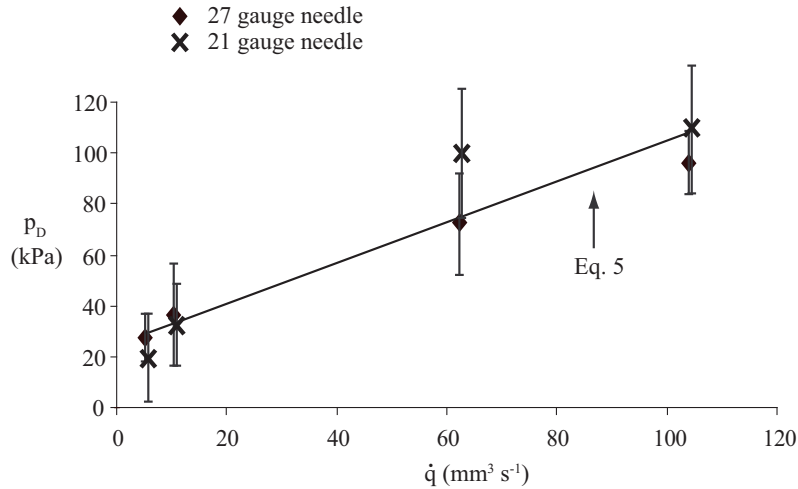
Figure 8. The mean pressure $p_d \pm$ s.d. versus volumetric flow rate $\dot{q}$ for the injection of PBS into adipose tissue using a 27 gauge needle and a 21 gauge needle. A least squares regression fit overlays the data.

X-ray imaging of injected liquid. Adipose tissue was injected with 0.5 ml of Urografin 150 (Bayer Plc, UK) radio-opaque fluid (viscosity similar to water) using a hypodermic syringe with a 21 gauge needle. Following injection the specimen was immediately placed in an x-ray inspection scanner (HMX160, X-Tech, Metris, Tring, UK) and 720 images were taken at $0.5°$ intervals about a $360°$ rotation. Image reconstruction software was used to generate a 3D representation from the 2D images.

X-ray images of the injection sites are shown in Figure 9. Examination of cross sections from the 3D reconstruction of the x-ray images show an ellipsoid shaped bolus of liquid embedded within the tissue. The injectate is distributed within micro-channels of maximum length 1 mm. The volume of the bolus is approximately $2300\,mm^3$. Recall that 0.5 ml Urografin fluid was injected. Consequently the average volume fraction of Urografin fluid in the overall bolus equals 0.23.

5. Discussion of injection response

The delivery pressure required to injected a liquid at low injection velocity is 10^3 times less than the pressure required for deep penetration by a solid punch. This suggests that different fracture mechanisms are involved. A seepage model for the delivery pressure and a hydraulic fracture model of the micro-cracks formed during liquid delivery are now assessed.

A model for delivery pressure. A model is now developed which describes the flow of a fluid from a spherical reservoir at the tip of the needle into micro-cracks that are created within the tissue during an injection. It is suggested that the dependence of delivery pressure p_D on volumetric flow rate $\dot{q}$ can be modelled by Darcy's law [Wang 2000]. Measurements of delivery pressure p_D are used to estimate the permeability k of the tissue during an injection, and possible mechanisms for the formation of the micro-cracks are assessed.
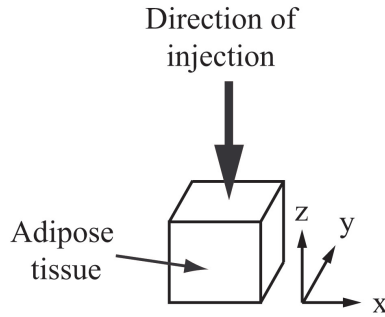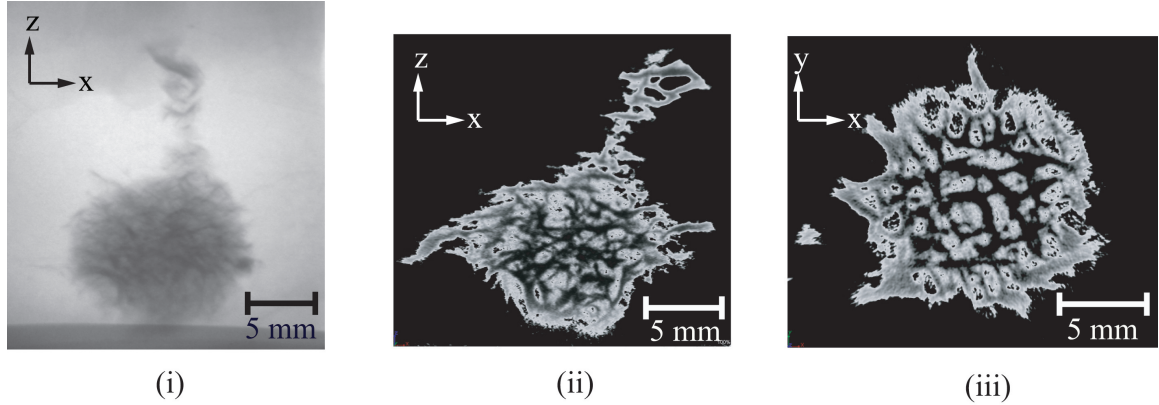
Figure 9. X-ray images of injection of 0.5 ml radio opaque dye (150 Urografin) into porcine adipose tissue by a hypodermic needle and syringe. Image (i) is a composite x-ray image. Images (ii) and (iii) are cross sections from a 3D reconstruction of 720 images taken at 0.5° around 360° of the sample.

First we examine the use of Darcy's law to model delivery pressure p_D. Darcy's law describes the nominal flow rate per unit area $\dot{u}$ as a function of the pressure gradient ∇_p and the permeability of the solid k, where

$$\dot{u} = -k\nabla_p \tag{6}$$

A model of the permeation of the injected fluid can be established as follows, see Figure 10. A fluid at a volumetric flow rate $\dot{q}$ is injected into adipose tissue. Assume that a fluid-filled reservoir of radius r_D and pressure p_D is formed at the tip of the needle, and the surrounding solid is of permeability k (see Figure 10). The volumetric flow rate $\dot{q}$ is related to the flow per unit area $\dot{u}(r)$ according to

$$\dot{u}(r) = \frac{\dot{q}}{4\pi r^2} \tag{7}$$

Substitution of (7) into (6) gives

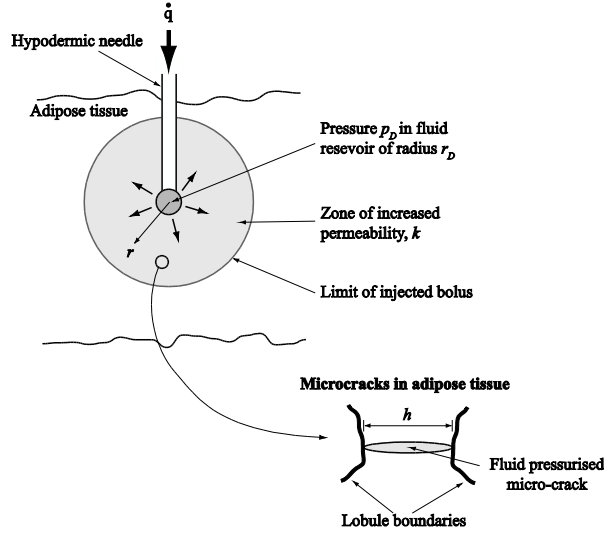$$-k\frac{\partial p}{\partial r} = \frac{\dot{q}}{4\pi r^2} \tag{8}$$

Figure 10. A sketch of the injection of a liquid into adipose tissue.

and thereby leads to an injection p_D at $r = r_D$ of

$$p_D = \frac{\dot{q}}{4\pi k r_D} \tag{9}$$

In order to perform an order of magnitude calculation for the permeability k it is assumed that $2r_D = 800\,\mu$m. This value matches the width of the fluid-filled channels observed in the x-ray images of Figure 9. Upon matching the regression coefficient $c = 740\,\text{GPa s m}^{-3}$ of (5) to the constant of proportionality in (9) one obtains

$$c = \left(4\pi k r_D\right)^{-1} \tag{10}$$

and consequently $k = 2.7 \times 10^{-10}\,\text{m}^4\,\text{N}^{-1}\,\text{s}^{-1}$.

How does this value compare with data reported in the literature? The permeability of adipose tissue for quasi-static flow is reported elsewhere to be between 1.8×10^{-12} and $1 \times 10^{-13}\,\text{m}^4\,\text{N}^{-1}\text{s}^{-1}$ [Guyton et al. 1966; Reddy et al. 1981]. These values are 10^2 and 10^3 times smaller than the permeability observed during the current injection tests. This indicates that injection of a pressurised fluid increases the permeability of the tissue by the creation of a connecting network of micro-cracks as observed in Figure 9. This phenomenon is analogous to hydraulic fracture of rocks where pumped fluid leads to the formation of a connecting network of cracks [Wang 2000].

Hydraulic fracture model. It is suggested that during injection the fluid tunnels through pathways of low resistance, possibly between lobules of adipocytes The mode I toughness J_C associated with the tunnelling fracture of a micro-crack of width h (see Figure 10), within a tissue of Young's modulus E and under a pressure p_0 is given in [Hutchinson and Suo 1991] as

$$J_C = \frac{p_0^2 h}{1.27E} \tag{11}$$

A measure of the pressure required to generate micro-cracks within the tissue is given by the value of $p_0 = 23$ kPa, taken from the least squares regression fit of the pressure data; see (5). Let the Young's modulus of adipose tissue be equal to $E = 1$ kPa [Comley and Fleck 2009]. Then for a crack of width $h = 800 \, \mu$m (matching the width of the observed micro-cracks in Figure 9) the corresponding fracture toughness is of the order $J_C = 330$ Jm^{-2}. This is an order of magnitude lower than the fracture toughness of adipose tissue $J_C = 4.1$ kJ m^{-2}, as measured via trouser tear test [Comley and Fleck 2010b]. This result indicates that the pressurised fluid seeks out relatively brittle interfaces within the adipose tissue. Further work is required to examine the injection site within the tissue via microscopy, in order to confirm (or refute) our conjecture of hydraulic fracture along weak paths.

6. Concluding remarks

Measurements have been made of the pressure to penetrate adipose tissue by a solid punch and the pressure to inject adipose tissue with a low viscosity liquid. The punch pressure is about three orders of magnitude greater than the injection pressure, indicating that different fracture mechanisms are involved. It is shown that the deep penetration pressure can be adequately modelled by the Shergold-Fleck model for ring-cracks. A seepage model based on Darcy flow is presented to account for the linear dependence of delivery pressure on volumetric flow rate. The seepage model reveals that the permeability of the tissue is significantly increased during the injection event. X-ray images of the injected bolus within the tissue indicate the presence of micro-cracks, formed during injection. A hydraulic fracture model reveals a low value of toughness, $J_C = 330$ J m^{-2} associated with the formation of these micro-cracks.

Acknowledgements

References

[Barry et al. 1995] S. I. Barry, G. K. Aldis, and G. N. Mercer, "Injection of fluid into a layer of deformable porous medium", *Appl. Mech. Rev. (ASME)* **48**:10 (1995), 722–726.

[Baxter and Mitragotri 2005] J. Baxter and S. Mitragotri, "Jet-induced skin puncture and its impact on needle-free jet injections: experimental studies and a predictive model", *J. Control. Release* **106**:3 (2005), 361–373.

[Björntorp and Martinsson 1966] P. Björntorp and A. Martinsson, "The composition of human subcutaneous adipose tissue in relation to its morphology", *Acta Med. Scand.* **179**:4 (1966), 475–481.

[Comley and Fleck 2009] K. S. C. Comley and N. A. Fleck, "The high strain rate response of adipose tissue", pp. 27–33 in *IUTAM Symposium on Mechanical Properties of Cellular Materials* (Cachan, 2007), edited by H. Zhao and N. A. Fleck, IUTAM Bookseries **12**, Springer, Dordrecht, 2009.

[Comley and Fleck 2010a] K. S. C. Comley and N. A. Fleck, "A micromechanical model for the Young's modulus of adipose tissue", *Int. J. Solids Struct.* **47**:21 (2010), 2982–2990.

[Comley and Fleck 2010b] K. S. C. Comley and N. A. Fleck, "The toughness of adipose tissue: measurements and physical basis", *J. Biomech.* **43**:9 (2010), 1823–1826.

[Comley and Fleck 2011] K. S. C. Comley and N. A. Fleck, "The constitutive response of porcine adipose tissue", 2011. Submitted to *Int. J. Impact Eng.*

[Cooke et al. 1980] E. D. Cooke, S. A. Bowcock, A. Johnston, and A. T. Elliott, "Gas-powered jet injection compared with conventional methods of injection using lignocaine and technetium-99m", *Br. Med. J.* **281**:6241 (1980), 643–644.

[Gefen and Haberman 2007] A. Gefen and E. Haberman, "Viscoelastic properties of ovine adipose tissue covering the gluteus muscles", *J. Biomech. Eng.* (*ASME*) **129**:6 (2007), 924–930.

[Greenwood and Johnson 1983] M. R. C. Greenwood and P. R. Johnson, "The adipose tissue", pp. 178–199 in *Histology: cell and tissue biology*, 5th ed., edited by L. Weiss, Elsevier, New York, 1983.

[Guyton et al. 1966] A. C. Guyton, K. Scheel, and D. Murphree, "Interstitial fluid pressure, 3: Its effect on resistance to tissue fluid mobility", *Circ. Res.* **19**:2 (1966), 412–419.

[Hutchinson and Suo 1991] J. W. Hutchinson and Z. Suo, "Mixed mode cracking in layered materials", *Adv. Appl. Mech.* **29** (1991), 63–191.

[Miller-Young et al. 2002] J. E. Miller-Young, N. A. Duncan, and G. Baroud, "Material properties of the human calcaneal fat pad in compression: experiment and theory", *J. Biomech.* **35**:12 (2002), 1523–1531.

[Nakajima et al. 1998] I. Nakajima, T. Yamaguchi, K. Ozutsumi, and H. Aso, "Adipose tissue extracellular matrix: newly organized by adipocytes during differentiation", *Differ.* **63**:4 (1998), 193–200.

[Nightingale et al. 2003] K. Nightingale, S. McAleavey, and G. Trahey, "Shear-wave generation using acoustic radiation force: *in vivo* and *ex vivo* results", *Ultrasound Med. Biol.* **29**:12 (2003), 1715–1723.

[Ogden 1972] R. W. Ogden, "Large deformation isotropic elasticity: on the correlation of theory and experiment for incompressible rubberlike solids", *Proc. R. Soc. Lond. A* **326**:1567 (1972), 565–584.

[Reddy et al. 1981] N. P. Reddy, G. V. B. Cochran, and T. A. Krouskop, "Interstitial fluid flow as a factor in decubitus ulcer formation", *J. Biomech.* **14**:12 (1981), 879–881.

[Samani et al. 2003] A. Samani, J. Bishop, C. Luginbuhl, and D. B. Plewes, "Measuring the elastic modulus of *ex vivo* small tissue samples", *Phys. Med. Biol.* **48**:14 (2003), 2183–2198.

[Schramm-Baxter and Mitragotri 2004] J. Schramm-Baxter and S. Mitragotri, "Needle-free jet injections: dependence of jet penetration and dispersion in the skin on jet power", *J. Control. Release* **97**:3 (2004), 527–535.

[Shergold and Fleck 2004] O. A. Shergold and N. A. Fleck, "Mechanisms of deep penetration of soft solids, with application to the injection and wounding of skin", *Proc. R. Soc. Lond. A* **460** (2004), 3037–3058.

[Shergold and Fleck 2005] O. A. Shergold and N. A. Fleck, "Experimental investigation into the deep penetration of soft solids by sharp and blunt punches, with application to the piercing of skin", *J. Biomech. Eng.* (*ASME*) **127**:5 (2005), 838–848.

[Shergold et al. 2006] O. A. Shergold, N. A. Fleck, and T. S. King, "The penetration of a soft solid by a liquid jet, with application to the administration of a needle-free injection", *J. Biomech.* **39**:14 (2006), 2593–2602.

[Wang 2000] H. F. Wang, *Theory of linear poroelasticity with applications to geomechanics and hydrogeology*, Princeton Univ. Press, Princeton, NJ, 2000.

KERSTYN COMLEY: kerst22@gmail.com
Cambridge University, Department of Engineering, Trumpington Road, Cambridge CB2 1PZ, United Kingdom

NORMAN FLECK: naf1@cam.ac.uk
Cambridge University, Department of Engineering, Trumpington Road, Cambridge CB2 1PZ, United Kingdom

msp

msp

ON SMALL AZIMUTHAL SHEAR DEFORMATION OF FIBRE-REINFORCED CYLINDRICAL TUBES

MOHAMED A. DAGHER AND KOSTAS P. SOLDATOS

The problem of azimuthal shear deformation of a transversely isotropic elastic circular cylindrical tube is considered and studied in the small deformation regime. The preferred direction of the transverse isotropy is assumed to lie on the plane of the tube cross-section and is due to the existence of a single family of plane spiral fibres. Consideration of the manner that either the tube material or the fibres may be constrained gives rise to four different versions of the problem which are all susceptible to an exact closed form solution when fibres are perfectly flexible. Particular attention is paid to the special case of straight fibres aligned along the radial direction of the tube cross-section, where comparisons are made between the aforementioned solution obtained when fibres are perfectly flexible and a corresponding solution obtained when fibres posses bending stiffness. It is found that the conventional linear elasticity considerations associated with the perfectly flexible fibre assumption cannot adequately account for the effects of material anisotropy. In contrast, effects of material anisotropy can be accounted for when fibres posses bending stiffness, by taking into consideration the action of couple-stress and therefore asymmetric stress. Moreover, an intrinsic material length parameter which appears naturally in the associated governing equations may be chosen as a representative of the fibre thickness in this case. It is also seen that deformation patterns of fibres possessing bending stiffness as well as corresponding stress distributions developed within the tube cross-section fit physical expectation much closer than their perfectly flexible fibre counterparts.

1. Introduction

The classical version of the problem of azimuthal shear deformation of an elastic circular cylindrical tube of infinite extent is due to Rivlin [1949] and, in several forms and variations, has been considered and studied afterwards by several investigators. This refers to a particular, plane-strain type of finite strain which is applied on the cross-section of an incompressible isotropic hyper-elastic circular cylindrical tube of infinite extent. Accordingly, under the action of an appropriate set of boundary conditions, the tube cross-section is subjected to pure azimuthal shear strain during which it remains circular while its inner and outer radii do not change. A comprehensive review of the relevant literature was presented recently in [Kassianidis et al. 2008], which introduced further and dealt with a new version of this problem; namely the case in which the incompressible material of the tube exhibits some kind of anisotropy. In some detail, the tube cross-section was considered to be reinforced by a single family of unidirectional extensible fibres; this consideration furnished the tube material with properties of transverse isotropy.

Most recently, a second family of plane extensible fibres was placed on the tube cross-section [Dorf-mann et al. 2010], thus assuming that the anisotropy of the material of interest proceeds beyond the relatively simple symmetries of transverse isotropy. As already implied, the directions of preference considered in either this latter paper or [Kassianidis et al. 2008] were assumed capable to extend or contract considerably, allowing thus the incompressible material of the tube to withstand the imposed conditions of pure azimuthal shear deformation.

A relevant problem was also considered recently in [Soldatos 2010], as an application of a study having as principal purpose to investigate the influence that some new, second gradient effects have on finite plane deformations of ideal fibre-reinforced hyper-elastic solids. That problem dealt with azimuthal shear strain of an incompressible hyper-elastic circular cylindrical tube having its cross-section reinforced by a single family of inextensible fibres (see also [Soldatos 2009a]); recall that an incompressible material which is further reinforced by one or more families of inextensible fibres is known as ideal fibre-reinforced material (see [Spencer 1972], for example). The new development in [Soldatos 2009a; 2010] made clear that pure azimuthal shear strain is not possible when the incompressible material of the tube contains an inextensible direction of transverse isotropy. Unlike [Kassianidis et al. 2008] where extension or contraction of fibres is assumed possible, a single family of inextensible fibres causes change of both the inner and outer tube radii in a manner that preserves the area of the tube cross-section. It is noted in passing that, since the cross-sectional area remains also unchanged under conditions of pure azimuthal shear strain, the latter kind of deformation [Rivlin 1949; Kassianidis et al. 2008; Dorfmann et al. 2010] becomes essentially a particular case of the outlined "area-preserving" azimuthal shear strain of a circular cylindrical tube.

The new, second gradient deformation effects that [Soldatos 2010] is mainly interested on are relevant with the ability of fibres to resist bending. However, the described "area-preserving" azimuthal shear deformation was found attainable by the ideal fibre-reinforced material considered in [Soldatos 2010] regardless of whether the inextensible fibres involved possess bending stiffness or not (in the latter case fibres are assumed perfectly flexible). The analysis in [Soldatos 2009a; 2010] revealed further that, if the inextensible fibres involved are initially straight and aligned along the radial direction of the tube cross-section, they remain straight during deformation and force the tube to undergo area-preserving azimuthal shear strain by changing their slope only. Some link was therefore observed between the strength of fibres in extension or contraction and their ability to resist bending. It was accordingly concluded that, if the direction of transverse isotropy is due to the existence of strong fibres, the tube should be expected to resist the conditions of pure azimuthal shear deformation. Instead, tendency will be observed for creation of a deformation pattern that couples azimuthal shear strain and radial stretching. Moreover, fibre bending stiffness should be dominant in the formation of such a pattern.

It is instructive at this point to mention that the principal problem met in nature is essentially the problem in which the transversely isotropic material of the tube is completely unconstrained. Hence, by employing the concept of the ideal fibre-reinforced material, the references [Soldatos 2009a; 2010] dealt essentially with a first approximation to the solution of the finite azimuthal shear strain problem of a fibre-reinforced cylindrical tube. Many materials are of course nearly incompressible and, similarly, many kinds of natural or structural fibres are nearly inextensible. Hence, in many cases of interest, either the incompressible material considered in [Kassianidis et al. 2008] or a compressible material reinforced by inextensible fibres yields a realistic and plausible simplification of the principal problem. Either case

is regarded as an intermediate step between the latter problem and that considered in [Soldatos 2009a; 2010] for a corresponding doubly constrained material. Further progress in the subject should therefore consider to dismiss one or both of the material constraints involved in these two references.

Dismissal of the fibre inextensibility constraint leads to the comprehensive relevant study presented already in [Kassianidis et al. 2008] where attention was focused on the particular case of pure azimuthal shear deformation only. In this context, dismissal of either the material incompressibility constraint only or both of the constraints involved in [Soldatos 2009a; 2010] produces two additional versions of the problem. Solution to either of those two versions in the finite strain regime seems to be a more difficult task as compared with the solutions achieved in all three of these references. Nevertheless, the outlined hierarchical manner of approaching difficult problems met in finite elasticity assists enormously the effort of achieving basic understanding of associated complicated issues.

Another plausible way for achieving basic understanding of some of those issues is by restricting initially attention to the small deformation regime within which the material is regarded as linearly elastic. There exists in fact an extensive literature of linear anisotropic elasticity solutions, most of which are associated or can become relevant to the mechanics of fibre-reinforced solids. In this context, the present study adds a new contribution to that literature by focusing attention to the linear elasticity counterpart of each one of the aforementioned four versions of the azimuthal shear strain problem of a transversely isotropic circular cylindrical tube. Based on the outlined history of the problem considered, this investigation aims therefore to identify which of the four versions of the problem anticipate that within the small strain regime (i) existence of possible coupling between azimuthal shear strain and radial stretching can cause change of the inner and outer radii of the tube; and/or (ii) the initial deformation pattern is or may still be interpreted as that of pure azimuthal shear strain, in the sense that the tube inner and outer radii do not tend to changed during deformation. Moreover, (iii) the particular case of straight fibres aligned along the radial direction of the tube cross-section is treated separately and, in the light of the relevant studies initiated in [Soldatos 2009a; 2010], comparisons are made between corresponding solutions and results obtained when fibres are either perfectly flexible or possess bending stiffness.

Under these considerations, section 2 formulates the problem of axially symmetric plane strain of a transversely isotropic, linearly elastic, annular disc (the tube cross-section) subjected to external boundary conditions that may cause pure azimuthal shear strain. It is noted that the formulation detailed in section 2 is based on symmetric elasticity considerations which concur with the assumption that fibres are perfectly flexible. For the case that the direction of transverse isotropy is due to fibres of a certain spiral shape, section 3 outlines next the exact, closed form solution obtained for each one of the aforementioned four versions of the azimuthal shear strain problem; namely, the case in which (i) the material of the tube is completely unconstrained and therefore compressible, (ii) the material is assumed incompressible but the fibres can extent or contract, (iii) the material of the tube is compressible but the fibres are assumed inextensible, and (iv) the material of the tube is incompressible and the fibres are inextensible (ideal fibre-reinforced material). Section 4 deals separately with the particular case of perfectly flexible straight fibres aligned along the radial direction of the tube cross-section. This case, along with its counterpart that considers fibres resistant in bending, was discussed also separately in [Soldatos 2010] for the purposes of the area-preserving finite azimuthal shear strain problem introduced and studied there.

When the perfectly flexible radial fibres considered in section 4 are replaced with radial fibres that posses bending stiffness the linear theory of elasticity is required to account further for possible effects of

couple-stress and therefore asymmetric stress. Moreover, micromechanics considerations reveal that, due to the natural appearance of an intrinsic material length parameter which is of the fibre thickness scale, the manner in which fibres are supported on the tube boundaries can also be accounted for, with use of appropriate boundary conditions. This is the case discussed and resolved completely in section 5, where the principal governing differential equation of the problem is solved exactly with use of the power-series method as well as the successive approximate method introduced in [Soldatos and Hadjigeorgiou 1990]; see also [Shuvalov and Soldatos 2003]. Relevant numerical results are presented in section 6, where the differences between conventional linear elasticity and the new developments introduced in section 5 are also discussed in detail. Finally, section 7 summarises the main results, observations and conclusions drawn in this investigation.

2. Problem formulation for perfectly flexible fibres

Consider a circular cylindrical hollow tube defined by

$$B_0 \leq r \leq B_1, \qquad 0 \leq \theta \leq 2\pi, \qquad -\infty \leq z \leq \infty, \tag{2-1}$$

where r, θ and z are appropriate cylindrical polar coordinate parameters and the nonnegative constants B_0 and B_1 represent the inner and outer radii of the tube, respectively. It is assumed that the tube is made of a transversely isotropic linearly elastic material and that the preferred direction of transverse isotropy lies on the plane of the tube cross-section. Accordingly, the preferred material direction is described as follows (e.g., [Kassianidis et al. 2008]):

$$\theta = G(r) + \theta_0, \qquad G(B_0) = 0, \qquad G(B_1) = \theta_1 - \theta_0 \geq 0, \tag{2-2}$$

where $0 \leq \theta_0 \leq 2\pi$ and $\theta_1 - \theta_0$ is fixed regardless of the value of θ_0. Here, the scalar function $G(r)$ defines the direction of transverse isotropy which is due to the existence of a single family of plane fibres making an angle $\alpha(r)$ with the radial direction. It is convenient to assume that $0 \leq \alpha(r) < \pi/2$ and, hence, that the family of fibres (the a-curves) have the form shown in Figure 1.
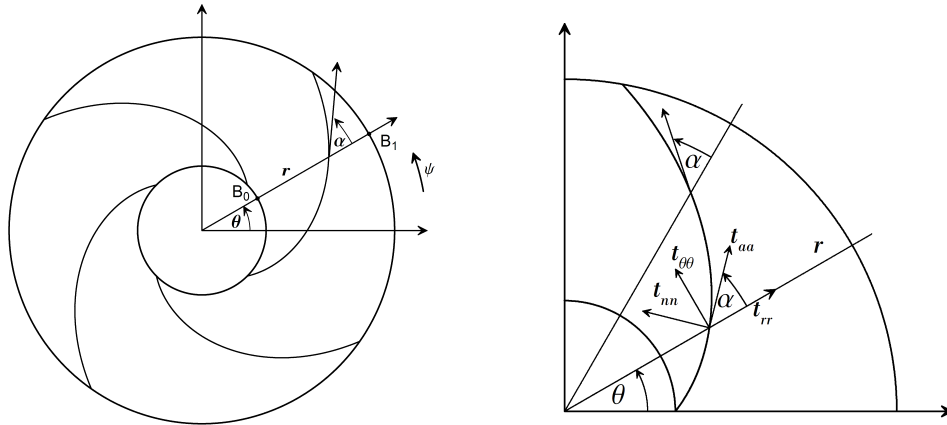

Figure 1. Left: schematic representation of the tube cross-section and associated notation. Right: illustration of normal in-plane stress components in both the local and polar coordinate systems.

The unit tangent $\boldsymbol{a}$ and the unit normal $\boldsymbol{n}$ of this family of material curves form the base of a local rectangular curvilinear coordinate system. In components, these unit vectors are represented as

$$\boldsymbol{a} = (a_r, a_\theta)^T, \qquad \boldsymbol{n} = (-a_\theta, a_r)^T, \tag{2-3}$$

where

$$a_r = \left[(rG'(r))^2 + 1\right]^{-1/2} \qquad a_\theta = rG'(r)\left[(rG'(r))^2 + 1\right]^{-1/2}, \tag{2-4}$$

and a prime denotes differentiation with respect to r. It follows that

$$\tan\alpha = rG'(r), \tag{2-5}$$

and, therefore, $G'(r) \geq 0$.

The material of the tube is assumed to be linearly elastic and, hence, the plane strain version of Hooke's law has the form

$$\begin{pmatrix} t_{aa} \\ t_{nn} \\ t_{an} \end{pmatrix} = \begin{pmatrix} C_{11} & C_{12} & 0 \\ C_{12} & C_{22} & 0 \\ 0 & 0 & C_{66} \end{pmatrix} \begin{pmatrix} e_{aa} \\ e_{nn} \\ 2e_{an} \end{pmatrix}, \qquad t_{zz} = C_{12}e_{aa} + C_{23}e_{nn}, \tag{2-6}$$

where $t_{aa}, t_{nn}, t_{an}, t_{zz}$ and e_{aa}, e_{nn}, e_{an} represent the nonzero components of the stress and the strain tensors respectively, in the aforementioned curvilinear local coordinate system (see also Figure 1, right); $C_{11}, C_{12}, C_{22}, C_{23}$ and C_{66} are appropriate nonzero elastic moduli (e.g., [Jones 1998]) which are assumed constant in what follows. In polar coordinates, $(2\text{-}6)_1$ takes the form

$$\begin{pmatrix} t_{rr} \\ t_{\theta\theta} \\ t_{r\theta} \end{pmatrix} = \begin{pmatrix} \bar{C}_{11} & \bar{C}_{12} & \bar{C}_{16} \\ \bar{C}_{12} & \bar{C}_{22} & \bar{C}_{26} \\ \bar{C}_{16} & \bar{C}_{26} & \bar{C}_{66} \end{pmatrix} \begin{pmatrix} e_{rr} \\ e_{\theta\theta} \\ 2e_{r\theta} \end{pmatrix}, \tag{2-7}$$

where $t_{rr}, t_{\theta\theta}, t_{r\theta}$ and $e_{rr}, e_{\theta\theta}, e_{r\theta}$ are the corresponding polar components of the stress and strain tensors, respectively. In general, it is $\alpha \neq 0$ and, hence, the stiffness matrices $[C]$ and $[\bar{C}]$ appearing in $(2\text{-}6)_1$ and (2-7) are related according to

$$[\bar{C}] = [T]^{-1}[C][T]^{-T}, \tag{2-8}$$

where (e.g., [Jones 1998])

$$[T] = \begin{pmatrix} \cos^2\alpha & \sin^2\alpha & \sin 2\alpha \\ \sin^2\alpha & \cos^2\alpha & -\sin 2\alpha \\ \frac{-1}{2}\sin 2\alpha & \frac{1}{2}\sin 2\alpha & \cos 2\alpha \end{pmatrix}, \tag{2-9}$$

and a superscript $^{-T}$ denotes the inverse of a transposed matrix. It is noted for later use, that both matrices $[C]$ and $[\bar{C}]$ are required to be positive definite, in order for the strain-energy of the system to be positive (e.g., [Ting 1996]).

For the axially symmetric plane deformations of interest, the strain components appearing in (2-7) are

$$e_{rr} = u'(r), \qquad e_{\theta\theta} = \frac{u(r)}{r}, \qquad 2e_{r\theta} = v'(r) - \frac{v(r)}{r}, \tag{2-10}$$

where $u(r)$ and $v(r)$ are the cross-sectional radial and the azimuthal displacement components, respectively. The set of available equations is completed with the equations of equilibrium. In the present case,

these take the form

$$rt'_{rr} + t_{rr} - t_{\theta\theta} = 0, \qquad (r^2 t_{r\theta})' = 0, \qquad \frac{\partial t_{zz}}{\partial z} = 0, \tag{2-11}$$

and, since plane strain assumes that stresses are independent of z, equation $(2\text{-}11)_3$ is satisfied identically.

The description of the azimuthal strain problem considered is completed by associating to it appropriate sets of boundary conditions. Accordingly, the boundary conditions imposed in the azimuthal direction of the tube inner and outer boundaries are

$$v(B_0) = 0, \qquad v(B_1) = \psi. \tag{2-12}$$

Since the inner boundary is assumed restrained from rotation, the azimuthal displacement ψ, which is imposed on the tube outer boundary, is assumed to be the cause of the axially symmetric plane-strain deformation of interest. If positive (negative), the known displacement ψ causes an anticlockwise (clockwise) rotation on the outer boundary of the tube cross-section; the azimuthal boundary traction is considered unknown on both boundaries and should therefore be determined from the analysis.

Anticipation of possible coupling between azimuthal shear strain and radial stretching is associated with the ability of the inner and outer tube radii to change during deformation, thus leading to the additional boundary conditions

$$t_{rr}(B_0) = t_{rr}(B_1) = 0. \tag{2-13}$$

However, when pure azimuthal shear strain becomes the principal deformation of interest, the tube radii remain unchanged during deformation and, hence, the natural boundary conditions (2-13) are replaced by the geometrical boundary conditions

$$u(B_0) = u(B_1) = 0. \tag{2-14}$$

3. Spiral fibres

For simplicity, it is now considered that α is constant and, hence, the fibres have the form of a logarithmic spiral; namely, a curve described by the function

$$G(r) = \tan\alpha \ln\frac{r}{B_0}. \tag{3-1}$$

With this relatively simple choice of $G(r)$, the components of the stiffness matrix $[\bar{C}]$ become independent of the polar distance, r. It is seen next that, as a consequence, an exact closed form solution of the problem is possible regardless of whether the material of the tube is unconstrained or is subjected to any combination of the aforementioned incompressibility and inextensibility constrains.

In what follows, equations are made nondimensional with use of the main nondimensional quantities

$$r^* = \frac{r}{B_0}, \qquad \beta = \frac{B_1}{B_0}, \qquad u^* = \frac{u}{\psi}, \qquad v^* = \frac{v}{\psi},$$

$$t_{ij}^* = \frac{B_0 t_{ij}}{\psi \bar{C}_{66}} \quad (i, j = r, \theta), \qquad \bar{C}_{ij}^* = \frac{\bar{C}_{ij}}{\bar{C}_{66}} \quad (i, j = 1, 2, 6). \tag{3-2}$$

It is also noted that the additional nondimensional quantities

$$p^* = \frac{B_0 p}{\psi \bar{C}_{66}}, \qquad T^* = \frac{B_0 T}{\psi \bar{C}_{66}} \qquad (3\text{-}3)$$

will be employed latter in sections 3.2, 3.3 and 3.4 in order to represent nondimensional forms of the arbitrary pressure and tension, respectively, introduced there. It is however noted that, for convenience, asterisks are dropped in all equations met next in sections 3 and 4. Hence, the form of the in-plane equilibrium equations $(2\text{-}11)_{1,2}$ remains unchanged under the implied nondimensional analysis while the corresponding nondimensional form of the boundary conditions (2-12), (2-13) and (2-14) is, respectively,

$$v(1) = 0, \ v(\beta) = 1, \qquad (3\text{-}4)$$

$$t_{rr}(1) = t_{rr}(\beta) = 0, \qquad (3\text{-}5)$$

$$u(1) = u(\beta) = 0. \qquad (3\text{-}6)$$

3.1. *Unconstrained material.* Use of the kinematic equations (2-10) and the Hooke's law (2-7) yields the nondimensional Navier-type form of the equilibrium equations $(2\text{-}11)_{1,2}$ as follows:

$$\bar{C}_{11} r(ru')' - \bar{C}_{22} u + \bar{C}_{16} r^2 v'' - \bar{C}_{26}(rv' - v) = 0,$$

$$\bar{C}_{16}(r^2 u')' + \bar{C}_{26}(ru)' + r^2 v'' + rv' - v = 0. \qquad (3\text{-}7)$$

This is a system of two second-order simultaneous, Euler-type ordinary differential equations (ODEs) which admit solutions of the form $u(r) = c_1 r^n$, $v(r) = c_2 r^n$.

Accordingly, the general solution of equations (3-7) is found to be

$$u(r) = \gamma_1 \bar{A}_1 r^{-1} + \bar{A}_2 r^{\eta} + \bar{A}_3 r^{-\eta},$$

$$v(r) = \bar{A}_4 r + \bar{A}_1 r^{-1} + \gamma_2 \bar{A}_2 r^{\eta} + \gamma_3 \bar{A}_3 r^{-\eta}, \qquad (3\text{-}8)$$

where $\bar{A}_1$, $\bar{A}_2$, $\bar{A}_3$ and $\bar{A}_4$ are arbitrary constants of integration, the constants γ_1, γ_2 and γ_3 are given in the Appendix and

$$\eta^2 = \frac{\bar{C}_{22} - \bar{C}_{26}^2}{\bar{C}_{11} - \bar{C}_{16}^2}. \qquad (3\text{-}9)$$

It can be shown that, due to the positive definiteness of $[\bar{C}]$, both the numerator and the denominator appearing in the right-hand side of (3-9) are positive and, therefore, η is always a real constant. Use of Hooke's law yields further the associated in-plane nondimensional stress components as follows:

$$t_{rr} = \bar{A}_1 F_1 r^{-2} + \bar{A}_2 F_2 r^{\eta-1} + \bar{A}_3 F_3 r^{-\eta-1},$$

$$t_{\theta\theta} = \bar{A}_1 F_4 r^{-2} + \bar{A}_2 F_5 r^{\eta-1} + \bar{A}_3 F_6 r^{-\eta-1},$$

$$t_{r\theta} = \bar{A}_1 \left(\gamma_1 (\bar{C}_{26} - \bar{C}_{16}) - 2 \right) r^{-2}, \qquad (3\text{-}10)$$

where the constants F_k ($k = 1, 2, \ldots, 6$) are given explicitly in the Appendix.

The form of the solution (3-8) makes immediately understood that change of both the inner and outer radii of the tube is generally always possible in this case, in which the material of the tube is completely unconstrained. For, if the particular set of boundary conditions (3-4) and (3-5) is considered, all four of

the arbitrary constants appearing in (3-8) and (3-10) acquire the unique nonzero values given explicitly in equations (A-3); hence, $u(r) \neq 0$ for all $1 \leq r \leq \beta$. In this case, the radial displacement of the inner and outer boundary of the tube is determined by setting $r = 1$ and $r = \beta$ in (3-8)$_1$, respectively.

If, on the other hand, the set of boundary conditions (3-4) and (3-6) is taken instead into consideration, the alternative set of unique nonzero values associated with the aforementioned arbitrary constants — see equations (A-5) — suggests that, although $u(r) \neq 0$ for $1 < r < \beta$ and, therefore, there is coupling between azimuthal shear and radial stretching in the interior of the tube, the boundary radii of the tube can be kept unchanged during deformation. Conditions of pure azimuthal shear can therefore also be observed in this case, although these require simultaneous action of appropriate nonzero normal tractions on the tube inner and outer boundaries; those tractions are determined by setting $r = 1$ and $r = \beta$ in (3-10)$_1$.

It should be finally noted that in the particular case that $\alpha = 0$ (radial fibres), equations (3-7) become uncoupled and, as a result, azimuthal shear strain and radial stretching become completely uncoupled deformations. It will be seen in what follows that this result is valid regardless of whether the material is constrained or not and, hence, this particular case, in which the fibres are aligned along the radial direction of the tube cross-section, is discussed separately in Section 4.

3.2. Incompressible material.
Incompressibility is a kinematic constraint which requires an arbitrary hydrostatic pressure $p(r)$ to be superimposed on the stress field; $p(r)$ does no work in any deformation which is compatible with the incompressibility constraint (tr $e = 0$). In this case, Hooke's law (2-7) is modified and its in-plane part takes the nondimensional form (e.g., [Spencer 1972; 1984])

$$\begin{pmatrix} t_{rr} \\ t_{\theta\theta} \\ t_{r\theta} \end{pmatrix} = \begin{pmatrix} \bar{C}_{11} & \bar{C}_{12} & \bar{C}_{16} \\ \bar{C}_{12} & \bar{C}_{22} & \bar{C}_{26} \\ \bar{C}_{16} & \bar{C}_{26} & 1 \end{pmatrix} \begin{pmatrix} e_{rr} \\ e_{\theta\theta} \\ 2e_{r\theta} \end{pmatrix} - p \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \tag{3-11}$$

where, as already mentioned, the appearing quantities are all nondimensionalized according to (3-2) and (3-3) before asterisks are dropped. The nondimensional form of the corresponding Navier-type governing equations then becomes

$$\bar{C}_{11}r(ru')' - \bar{C}_{22}u + \bar{C}_{16}r^2v'' - \bar{C}_{26}(rv' - v) - r^2p' = 0,$$
$$\bar{C}_{16}(r^2u')' + \bar{C}_{26}(ru)' + r^2v'' + rv' - v = 0, \tag{3-12}$$

and are accompanied by the incompressibility constraint $e_{rr} + e_{\theta\theta} = 0$, which yields the additional equation

$$u' + \frac{u}{r} = 0. \tag{3-13}$$

Solution of (3-13) yields $u(r)$ which is then inserted into (3-12)$_2$. The latter yields thus an inhomogeneous Euler ODE which is solved in the standard manner for the determination of $v(r)$. Solution of (3-12)$_1$ becomes next possible for $p(r)$ and, hence, the general solution of the system of simultaneous ODEs (3-12) and (3-13) is found to be

$$u(r) = \tilde{A}_1 r^{-1}, \quad v(r) = \tilde{A}_2 r + \tilde{A}_3 r^{-1}, \quad p(r) = \tilde{A}_4 - \tfrac{1}{2}\big(\tilde{A}_1(\bar{C}_{11} - \bar{C}_{22}) + 2\tilde{A}_3(\bar{C}_{16} + \bar{C}_{26})\big)r^{-2}, \tag{3-14}$$

where $\tilde{A}_1$, $\tilde{A}_2$, $\tilde{A}_3$ and $\tilde{A}_4$ are arbitrary constants of integration. Use of the constitutive equation (3-11) yields further the associated in-plane nondimensional stress components:

$$t_{rr} = \left(\tfrac{1}{2}\tilde{A}_1(2\bar{C}_{12} - \bar{C}_{11} - \bar{C}_{22}) + \tilde{A}_3(\bar{C}_{26} - \bar{C}_{16})\right)r^{-2} - \tilde{A}_4,$$
$$t_{\theta\theta} = \left(-\tfrac{1}{2}\tilde{A}_1(2\bar{C}_{12} - \bar{C}_{11} - \bar{C}_{22}) - \tilde{A}_3(\bar{C}_{26} - \bar{C}_{16})\right)r^{-2} + \tilde{A}_4,$$
$$t_{r\theta} = \left(\tilde{A}_1(\bar{C}_{26} - \bar{C}_{16}) - 2\tilde{A}_3\right)r^{-2}. \tag{3-15}$$

By considering the particular set of boundary conditions (3-4) and (3-5), $\tilde{A}_1$, $\tilde{A}_2$ and $\tilde{A}_3$ take unique nonzero values — see (A-7) — while $\tilde{A}_4 = 0$. Hence, $u(r) \neq 0$ for all $1 \leq r \leq \beta$ and a change of both the inner and outer radii of the tube is generally again possible in this case. It is also noted that, despite the nonzero values of both $\tilde{A}_1$ and $\tilde{A}_3$, $t_{rr} = t_{\theta\theta} = 0$ throughout the tube cross-section. Finally, if $\alpha = \pi/4$, then $\bar{C}_{16} = \bar{C}_{26}$ and, therefore, $u(r) = 0$ throughout the tube cross-section, thus causing conditions of pure azimuthal shear strain in this particular case.

Conditions of pure azimuthal shear strain are also possible for $\alpha \neq 0$, if only one of the two boundary conditions (3-6) is satisfied along with (3-4). Due to the form of (3-14)$_1$, the unused of the geometrical boundary conditions (3-6) is satisfied automatically and should therefore be replaced by its natural boundary condition counterpart detailed in (3-5). If, for instance, the set of mixed boundary conditions

$$u(1) = t_{rr}(\beta) = 0, \tag{3-16}$$

is chosen to replace (3-6), while (3-4) still hold, then $\tilde{A}_1 = 0$ and, therefore, $u(r) = 0$ throughout the tube cross-section ($\tilde{A}_2$, $\tilde{A}_3$ and $\tilde{A}_4$ are given by (A-8)). Nevertheless, a nonzero normal traction should act in this case in the radial direction of the tube inner boundary; this is determined by setting $r = 1$ in (3-15)$_1$. Similar arguments hold true if (3-16) are replaced by $t_{rr}(1) = u(\beta) = 0$. Pure azimuthal shear strain is observed again in this case, though a normal traction should be applied radially on the outer tube boundary.

In the particular case that $\alpha = 0$, equation (3-12)$_2$ becomes uncoupled from the set of equations (3-12)$_1$ and (3-13). Hence, azimuthal shear strain and radial stretching become again completely uncoupled deformations. Moreover, since the incompressibility constraint is associated with radial stretching only, the azimuthal strain problem becomes identical with its unconstrained material counterpart; this is the case discussed separately in Section 4.

3.3. Inextensible fibres.

The constraint of fibre inextensibility ($a^T e a = 0$) requires an arbitrary tension $T(r)$ to be superimposed on the stress field; this acts along the fibre direction and does no work in any deformation which conforms with this constraint. In this case Hooke's law takes the nondimensional form (e.g., [Spencer 1972; 1984])

$$\begin{pmatrix} t_{rr} \\ t_{\theta\theta} \\ t_{r\theta} \end{pmatrix} = \begin{pmatrix} \bar{C}_{11} & \bar{C}_{12} & \bar{C}_{16} \\ \bar{C}_{12} & \bar{C}_{22} & \bar{C}_{26} \\ \bar{C}_{16} & \bar{C}_{26} & 1 \end{pmatrix} \begin{pmatrix} e_{rr} \\ e_{\theta\theta} \\ 2e_{r\theta} \end{pmatrix} + T \begin{pmatrix} \cos^2\alpha \\ \sin^2\alpha \\ \sin\alpha\cos\alpha \end{pmatrix}, \tag{3-17}$$

where the appearing quantities are again nondimensionalized according to (3-2) and (3-3) before asterisks are dropped. The nondimensional form of the corresponding Navier-type governing equations then

becomes

$$\bar{C}_{11}r(ru')' - \bar{C}_{22}u + \bar{C}_{16}r^2v'' - \bar{C}_{26}(rv' - v) + r(rT)'\cos^2\alpha - rT\sin^2\alpha = 0,$$

$$\bar{C}_{16}(r^2u')' + \bar{C}_{26}(ru)' + r^2v'' + rv' - v + \tfrac{1}{2}\sin 2\alpha(r^2T)' = 0. \tag{3-18}$$

These are accompanied by the inextensibility constraint $e_{rr}\cot\alpha + e_{\theta\theta}\tan\alpha + 2e_{r\theta} = 0$ which, with use of (2-10), yields the additional equation

$$u'\cot\alpha + \frac{u}{r}\tan\alpha + v' - \frac{v}{r} = 0. \tag{3-19}$$

Equation (3-18)$_2$ can immediately be integrated once. A subsequent elimination of $v(r)$ and $T(r)$ from the resulting equations yields an inhomogeneous second-order Euler-type ODE for $u(r)$ which can be solved in the standard manner. The general solution of the system of simultaneous ODEs (3-18) and (3-19) can then be obtained with relative ease, to yield

$$u(r) = \frac{\hat{A}_1}{\hat{\gamma}_1(1 - m^2)}r^{-1} + \hat{A}_2r^m + \hat{A}_3r^{-m},$$

$$v(r) = \hat{A}_4r - \frac{\hat{A}_1\cot 2\alpha}{\hat{\gamma}_1(1 - m^2)}r^{-1} - \hat{A}_2\beta_1r^m - \hat{A}_3\beta_2r^{-m},$$

$$T(r) = \frac{2\hat{A}_1}{\sin 2\alpha}\left(1 + \frac{\bar{C}_{16} - \bar{C}_{26} - 2\cot 2\alpha}{\hat{\gamma}_1(1 - m^2)}\right)r^{-2}$$

$$+ \frac{2\hat{A}_2}{\sin 2\alpha}\left(\beta_1(m - 1) - m\bar{C}_{16} - \bar{C}_{26}\right)r^{m-1} + \frac{2\hat{A}_3}{\sin 2\alpha}\left(m\bar{C}_{16} - \beta_2(m + 1) - \bar{C}_{26}\right)r^{-m-1}, \tag{3-20}$$

where $\hat{A}_1$, $\hat{A}_2$, $\hat{A}_3$, $\hat{A}_4$ are arbitrary constants of integration, the constants β_1, β_2, $\hat{\gamma}_1$ and $\hat{\gamma}_2$ are given explicitly in the Appendix and

$$m^2 = \frac{\hat{\gamma}_2}{\hat{\gamma}_1}. \tag{3-21}$$

It can be shown that, due to the positive definiteness of the matrix $[C]$ and the fact that $0 \le \alpha < \pi/2$, both the numerator and the denominator in the right-hand side of (3-21) are positive and, therefore, the constant m is always real. Use of (3-17) yields next the associated in-plane nondimensional stress components as follows:

$$t_{rr} = \frac{\hat{A}_1}{\hat{\gamma}_1(1 - m^2)}H_1r^{-2} + \hat{A}_2H_2r^{m-1} + \hat{A}_3H_3r^{-m-1},$$

$$t_{\theta\theta} = \frac{\hat{A}_1}{\hat{\gamma}_1(1 - m^2)}H_4r^{-2} + \hat{A}_2H_5r^{m-1} + \hat{A}_3H_6r^{-m-1},$$

$$t_{r\theta} = \hat{A}_1r^{-2}, \tag{3-22}$$

where the constants H_k ($k = 1, 2, \ldots, 6$) are given explicitly in (A-10).

If the set of boundary conditions (3-4) and (3-5) is associated with this solution, all four arbitrary constants appearing in (3-20) and (3-22) take unique nonzero values (see (A-11)); hence, $u(r) \ne 0$ for all $1 \le r \le \beta$. It follows that a solution anticipating that both the inner and the outer tube radii change is again possible in this case. If, on the other hand, the alternative set of boundary conditions

(3-4) and (3-6) is instead considered, the corresponding set of unique nonzero values associated to those constants — see (A-13) — suggests that conditions of pure azimuthal shear strain can instead also be observed. Nevertheless, although the boundary radii of the tube do not change during the observed pure azimuthal shear deformation, it is again $u(r) \neq 0$ for $1 < r < \beta$ and, therefore, there is again coupling between azimuthal shear strain and radial stretching in the interior of the tube. It is again anticipated that appropriate nonzero normal tractions should be applied on the tube inner and outer boundaries; these are determined by setting $r = 1$ and $r = \beta$ in (3-22)$_1$.

In the case $\alpha = 0$, radial stretching and azimuthal shear strain become again uncoupled deformations. The fibre inextensibility constraint becomes associated with radial stretching in this case, so the azimuthal shear strain problem becomes again identical with its unconstrained material counterpart; see Section 4.

3.4. *Ideal fibre-reinforced material.* In this case, the material is assumed to be incompressible and also reinforced by inextensible fibres. Hence, Hooke's law takes the nondimensional form (e.g., [Spencer 1972; 1984])

$$
\begin{pmatrix} t_{rr} \\ t_{\theta\theta} \\ t_{r\theta} \end{pmatrix} = \begin{pmatrix} \bar{C}_{11} & \bar{C}_{12} & \bar{C}_{16} \\ \bar{C}_{12} & \bar{C}_{22} & \bar{C}_{26} \\ \bar{C}_{16} & \bar{C}_{26} & 1 \end{pmatrix} \begin{pmatrix} e_{rr} \\ e_{\theta\theta} \\ 2e_{r\theta} \end{pmatrix} + \begin{pmatrix} T\cos^2\alpha - p \\ T\sin^2\alpha - p \\ T\sin\alpha\cos\alpha \end{pmatrix}, \tag{3-23}
$$

where the appearing quantities are all nondimensionalized according to (3-2) and (3-3), before asterisks are dropped. The nondimensional form of the corresponding Navier-type governing equations then becomes

$$
\bar{C}_{11} r (ru')' - \bar{C}_{22} u + \bar{C}_{16} r^2 v'' - \bar{C}_{26}(rv' - v) + r^2(T'\cos^2\alpha - p') + rT\cos 2\alpha = 0,
$$
$$
\bar{C}_{16}(r^2 u')' + \bar{C}_{26}(ru)' + r^2 v'' + rv' - v + \tfrac{1}{2}\sin 2\alpha (r^2 T)' = 0, \tag{3-24}
$$

and are accompanied by both constraint equations (3-13) and (3-19), thus forming a system of four simultaneous ODEs for a total of four unknown functions, namely u, v, p and T.

Solution of Equation (3-13) yields $u(r)$, which is then inserted into (3-19) for the determination of $v(r)$. With the form of $u(r)$ and $v(r)$ becoming thus known, $T(r)$ and $p(r)$ are next obtained by consecutively solving (3-24)$_2$ and (3-24)$_1$, respectively. Hence, the general solution of the system of simultaneous ODEs (3-24), (3-13) and (3-19) is found to be

$$
u(r) = A_1 r^{-1},
$$
$$
v(r) = A_2 r - A_1 r^{-1} \cot 2\alpha,
$$
$$
p(r) = A_3 + \tfrac{1}{2}\left(A_4 - A_1(\bar{C}_{11} - \bar{C}_{22} - 2(\bar{C}_{16} + \bar{C}_{26})\cot 2\alpha) \right) r^{-2},
$$
$$
T(r) = A_4 r^{-2}, \tag{3-25}
$$

where A_1, A_2, A_3, A_4 are arbitrary constants. Use of (3-23) yields next the associated in-plane nondimensional stress components as follows:

$$
t_{rr} = \left(\tfrac{1}{2} A_1(2\bar{C}_{12} - \bar{C}_{11} - \bar{C}_{22} + 2(\bar{C}_{16} - \bar{C}_{26})\cot 2\alpha) + \tfrac{1}{2} A_4 \cos 2\alpha \right) r^{-2} - A_3,
$$
$$
t_{\theta\theta} = \left(-\tfrac{1}{2} A_1(2\bar{C}_{12} - \bar{C}_{11} - \bar{C}_{22} + 2(\bar{C}_{16} - \bar{C}_{26})\cot 2\alpha) - \tfrac{1}{2} A_4 \cos 2\alpha \right) r^{-2} + A_3,
$$
$$
t_{r\theta} = \left(A_1(2\bar{C}_{66} + \bar{C}_{26} - \bar{C}_{16}) + \tfrac{1}{2} A_4 \sin 2\alpha \right) r^{-2}. \tag{3-26}
$$

If the solution (3-25) is associated with the set of boundary conditions (3-4) and (3-5), A_1, A_2 and A_4 take unique nonzero values — (A-15) — while $A_3 = 0$. Hence, $u(r) \neq 0$ for all $1 \leq r \leq \beta$ and a change of both the inner and outer radii of the tube is again possible; it is also noted that, despite of the nonzero values of A_1 and A_4, $t_{rr} = t_{\theta\theta} = 0$ throughout the tube cross-section. On the other hand, pure azimuthal shear strain deformation is not possible in this case. For, simultaneous satisfaction of the boundary conditions (3-4)$_1$ and (3-6)$_1$ yields $u(r) = v(r) = 0$ throughout the tube cross-section.

It is again seen that in the particular case that $\alpha = 0$, radial stretching and azimuthal shear strain become again uncoupled problems, while both constraints involved associate themselves with radial stretching. Hence, the azimuthal shear strain problem becomes again identical with its unconstrained material counterpart which is discussed in the next section.

4. Perfectly flexible radial fibres

Interest is now focused in the particular case in which the fibres are straight and aligned along the radial direction of the tube cross-section. This corresponds to the choice $G(r) = \alpha = 0$ and, since the local and the polar coordinate systems coincide, the elastic behaviour of the material is described by (2-6), provided that the appearing local in-plane stress and strain components are replaced by their polar counterparts appearing in (2-7). It is already seen that, when $\alpha = 0$, radial stretching and azimuthal shear strain become completely uncoupled deformations regardless of whether the material is constrained or not. Restricting, for instance, attention to the unconstrained material case discussed in Section 3.1, one finds that equations (3-7) become uncoupled when $\alpha = 0$ and, hence, the resulting azimuthal shear strain problem is completely described by the Euler differential equation

$$r^2 v'' + r v' - v = 0. \tag{4-1}$$

This second-order ODE can describe pure azimuthal shear deformation only and, hence, it is associated with the pair of boundary conditions (3-4) only. Solution of this boundary value problem yields

$$v(r) = \frac{\beta}{\beta^2 - 1} \left(r - r^{-1} \right), \tag{4-2}$$

and Hooke's law (2-6)$_1$ reveals further that the nondimensional azimuthal shear stress

$$t_{r\theta} = \frac{2\beta}{\beta^2 - 1} r^{-2}, \tag{4-3}$$

is the only nonzero stress component associated with this deformation.

Interestingly enough, when the tube material is constrained in the manner suggested in Section 3, every single one of the problems discussed in sections 3.2, 3.3 and 3.4 provides precisely the same description for the pure azimuthal shear strain problem, namely (4-1) and (3-4); hence, in each case, it yields precisely the same solution with the outlined above on the basis of the unconstrained material version of the problem. It also becomes evident that, since material anisotropy does not enter the outlined problem description, the obtained solution (4-2) and (4-3) is identical to its isotropic material counterpart. However, it is already known [Soldatos 2009a; Soldatos 2010] that pure azimuthal shear deformation of a radially reinforced tube made of ideal fibre reinforced material (see Section 3.4) is not possible in the finite deformations regime. In contrast, finite pure azimuthal shear strain is indeed possible if the tube

material is incompressible but the radial fibres involved are free to extend or contract [Kassianidis et al. 2008]. Moreover, the material incompressibility constraint affects in a clear and obvious manner the relevant finite elasticity solution obtained in this latter reference, whereas this fact is not observed in the present case of small deformations (Section 3.2). It is instead observed that conventional linear elasticity theory cannot adequately account for the effects that material anisotropy and/or either of the constraints of material incompressibility and fibre inextensibility have on the azimuthal shear deformation problem considered when the tube is reinforced along the radial direction of its cross-section. In what follows, the infinitesimal strain problem described by (4-1) and (3-4), as well as its relatively simple solution given by (4-2) and (4-3), will therefore be mainly associated with the case discussed in Section 3.1, where the material is assumed completely unconstrained.

5. Radial fibres with bending stiffness

When the perfectly flexible fibres considered in the previous sections are replaced with fibres possessing bending stiffness, the theory is required to account further for possible action of couple-stress and therefore asymmetric stress [Spencer and Soldatos 2007]. The linearized version of the relevant hyper elasticity theory presented in that paper is considered for a study of the effects that fibre bending stiffness has on the azimuthal shear strain problem discussed in the preceding section. For convenience, the restricted part of that linearized theory, which requires use of only one additional elastic modulus, is employed here.

5.1. *Problem formulation.* The form of Hooke's law presented in Section 2 for $\alpha = 0$, before the nondimensional quantities (3-2) were introduced, is now suitable only for description of the symmetric part of the stress tensor. Hence,

$$\begin{pmatrix} t_{rr} \\ t_{\theta\theta} \\ t_{(r\theta)} \end{pmatrix} = \begin{pmatrix} C_{11} & C_{12} & 0 \\ C_{12} & C_{22} & 0 \\ 0 & 0 & C_{66} \end{pmatrix} \begin{pmatrix} e_{rr} \\ e_{\theta\theta} \\ 2e_{r\theta} \end{pmatrix}, \tag{5-1}$$

where $t_{(r\theta)}$ denotes the symmetric part of the shear stress component $t_{r\theta}$, the matrix $[C]$ is identical to its counterpart involved in (2-6) and the appearing strain components are given according to (2-10).

The antisymmetric part of $t_{r\theta}$ is caused by the action of a relevant couple-stress component, m_{rz}, arising when fibres resist bending. In terms of a notation similar to that adopted in [Soldatos 2009b], this couple-stress component is given according to

$$m_{rz} = d^f k_\theta^f, \qquad k_\theta^f = v''(r), \tag{5-2}$$

where d^f is the aforementioned additional elastic modulus (fibre bending stiffness) and k_θ^f represents the in-plane curvature component of the fibre in the linear elasticity regime. The antisymmetric part of the shear stress component is then expressed as

$$t_{[\theta r]} = -t_{[r\theta]} = \tfrac{1}{2}m'_{rz} = \tfrac{1}{2}d^f v'''(r). \tag{5-3}$$

It thus becomes immediately understood that radial stretching and azimuthal shear strain remain completely uncoupled deformations regardless of whether the material is constrained or not. Hence, the azimuthal equation of equilibrium (2-11)$_2$, which is again the only governing equation to be considered,

yields

$$\frac{d^f}{2C_{66}}(r^2 v'''' + 2r v''') - (r^2 v'' + r v' - v) = 0, \tag{5-4}$$

and makes obvious that the existing material anisotropy is indeed accounted for in this case.

Equation (5-4) is a fourth-order ODE for the azimuthal displacement component $v(r)$ and a unique determination of its solution requires specification of four boundary conditions. Two of them are evidently (2-12) while two more boundary conditions can be deduced from the relevant discussion detailed in [Soldatos 2009b]. Accordingly, the outer boundary ($r = B_1$) is assumed free of couple-stress, thus leading to the additional boundary condition

$$m_{rz}(B_1) = 0, \tag{5-5}$$

while the inner boundary ($r = B_0$) may be assumed either restrained against rotation or free of couple stress. It follows that the last of the four boundary conditions sought is one of the following alternatives:

$$\text{either} \quad v'(B_0) = 0 \quad \text{or} \quad m_{rz}(B_0) = 0. \tag{5-6}$$

It is observed that unlike its perfectly flexible fibres counterpart in (4-1), Equation (5-4) depends on the tube material properties. In this regard, the notation

$$\frac{d^f}{C_{66}} = 2l(B_1 - B_0), \tag{5-7}$$

introduces an intrinsic material length parameter l, which may be considered relevant to the fibre thickness. It is evident that when $l = 0$, (5-4) reduces to its perfectly flexible fibres counterpart. The role of nonzero values of l will become clearer in what follows.

5.2. *Nondimensional form of governing equations.* The Navier-type governing differential equation (5-4) and the boundary conditions associated to it are next nondimensionalized with use of the nondimensional quantities introduced in (3-2) and the additional nondimensional parameter

$$\lambda = \frac{l}{B_0} = \frac{d^f}{2C_{66}B_0(B_1 - B_0)}. \tag{5-8}$$

Nevertheless, asterisks are again dropped for convenience and, hence, all relevant quantities appearing without an asterisk in the remaining of this section, as well as Section 6, are those defined in (3-2). Accordingly, with use of (5-2), the nondimensional version of the present boundary value problem is found to be

$$\lambda(\beta - 1)(r^2 v'''' + 2r v''') - (r^2 v'' + r v' - v) = 0,$$
$$v(1) = 0, \quad v(\beta) = 1, \quad v''(\beta) = 0, \tag{5-9}$$

and

$$\text{either} \quad v'(1) = 0 \quad \text{or} \quad v''(1) = 0. \tag{5-10}$$

It is worth noting that, with simultaneous consideration of (5-9)$_2$, either condition in (5-10) resembles the boundary condition imposed at the end of an elastic slender beam which is clamped or simply supported, respectively, at $\beta = 1$.

With $l = 0$ in the case of perfectly flexible fibres, the conventional (symmetric) theory of linear elasticity (Sections 2–4) implied that the tube cross-section contains an infinite number of fibres. By associating the intrinsic length parameter l with the fibre thickness, the nonsymmetric stress theory employed in this section implies that the number of fibres as well as their density within the cross-section may be accounted for. Accordingly, if a fibre in the present plane strain case is thought of as a slender rectangle of length $B_1 - B_0$, then its area may be represented by the product $l(B_1 - B_0)$ and, hence, the finite number of fibres, N^f, that can be fitted into the tube cross-section is estimated to be

$$N^f = \frac{2\pi B_0}{l} = \frac{2\pi}{\lambda}. \tag{5-11}$$

It follows that the value of total fibre area is $2\pi B_0(B_1 - B_0)$ and this is independent of the fibre thickness. It is also of interest to note that, by dividing the total fibre area by the tube cross-sectional area, the fibre area fraction is estimated to be

$$S^f = \frac{2B_0}{B_1 + B_0} = \frac{2}{\beta + 1}, \tag{5-12}$$

and depends solely on the value of β. It is, therefore, seen that the value of β is indicative of the density (sparsity) of the fibre distribution within the tube cross-section. This result is illustrated in Table 1, where the total fibre area is calculated as a percentage of the area of the tube cross-section. It is finally noted that (5-11) implies $\lambda \leq 2\pi$. With $l \ll B_0$ and, therefore, $\lambda \ll 1$ in many practical applications, this inequality may be perceived as a natural consequence of the fact that thickness of common structural fibres is much smaller than the inner tube radius. However, since B_0 may in principle be smaller than l even if the fibre thickness is of the order of 10μ, values of $\lambda > 1$ are also anticipated by the present theory. In this context, the right-hand-side of (5-8) suggests that different physical interpretations of λ and/or l might also be possible, particularly when $O(\lambda) = 1$.

β	1	1.5	2	2.5	3	4	5	7.5	10	20	50	100
S^f %	100	80	66.7	57.1	50	40	33.3	23.5	18.2	9.5	3.9	2

Table 1. Estimated total area of radial fibres with bending stiffness as a percentage of the area of the tube cross-section.

Solution of the boundary value problem (5-9) and (5-10) is next achieved analytically, via the power series method, and computationally, with use of the successive approximation method (SAM) introduced in [Soldatos and Hadjigeorgiou 1990] (see also [Shuvalov and Soldatos 2003; Ye 2003; Soldatos 2003]).

5.3. _Power series solution._ Application of the power series method is based on the following Taylor-type series expansion of the solution sought around $r = 1$:

$$v(r) = \sum_{n=0}^{\infty} a_n (r - 1)^n, \tag{5-13}$$

where the constant coefficients a_n ($n = 0, 1, 2, \ldots$) are to be determined. Introduction of (5-13) into (5-9)$_1$, followed by nullification of the coefficients of like powers of $r - 1$, leads to the recurrence

relations

$$a_4 = \frac{a_1 + 2a_2 - 2\lambda(\beta - 1)3!a_3}{\lambda(\beta - 1)4!},$$

$$a_5 = -\frac{a_0 + 4a_1 + 6a_2 - (6\lambda(\beta - 1) + 1)3!a_3}{\lambda(\beta - 1)5!},$$

$$a_{n+4} = \frac{(n-1)(n+1)!}{\lambda(\beta-1)(n+4)!}a_n + \frac{(2n+1)(n+1)!}{\lambda(\beta-1)(n+4)!}a_{n+1}$$
$$- \left(\frac{(n+1)!}{(n-1)!} - \frac{1}{\lambda(\beta-1)}\right)\frac{(n+2)!}{(n+4)!}a_{n+2} - \frac{2(n+1)}{(n+4)}a_{n+3}, \quad n \geq 2, \quad (5\text{-}14)$$

where a_0, a_1, a_2 and a_3 are arbitrary constants. Use of the boundary conditions $(5\text{-}9)_2$ and $(5\text{-}10)_1$ yields

$$a_0 = 0, \qquad a_1 = 0, \tag{5-15}$$

while use of $(5\text{-}9)_2$ and $(5\text{-}10)_2$ yields

$$a_0 = 0, \qquad a_2 = 0. \tag{5-16}$$

In both cases, the values of the remaining constants are then determined numerically with use of the boundary conditions $(5\text{-}9)_{3,4}$, after the series expansion (5-13) is truncated to an appropriate number of terms that guarantee convergence of the obtained numerical results to a desired accuracy.

5.4. *Successive approximation solution.* Application of the well established successive approximation method introduced in [Soldatos and Hadjigeorgiou 1990] requires initially the conversion of $(5\text{-}9)_1$ into a system of four simultaneous first-order linear ODEs. In matrix form, these may be arranged as

$$\{X(r)\}' = [D(r)]\{X(r)\}, \qquad \{X(r)\}^T = \{v, v', v'', v'''\}, \tag{5-17}$$

where the nonzero elements of the matrix $[D(r)]$ are

$$D_{12} = D_{23} = D_{34} = 1, \qquad D_{43} = rD_{42} = -r^2D_{41} = \frac{1}{\lambda(\beta-1)}, \qquad D_{44} = -\frac{2}{r}. \tag{5-18}$$

For sufficiently thin tubes, an approximate solution is obtained by replacing the variable r appearing in (5-18) with the nondimensional cross-sectional middle-radius parameter $R = (\beta + 1)/2$. The resulting approximate system of four simultaneous linear ODEs with constant coefficients may then be written in the form

$$\{X(r)\}' = [D(R)]\{X(r)\}, \tag{5-19}$$

and its general solution can be expressed as follows:

$$\{X(r)\} = [K(r)]\{X(1)\}, \qquad 1 \leq r \leq \beta. \tag{5-20}$$

Here $\{X(1)\}$ denotes the value that the vector $\{X\}$ takes at the inner boundary of the tube, while the elements of the exponential matrix $[K(r)] = \exp[(r - 1)D(R)]$ can be evaluated analytically in the manner detailed in [Ye 2003].

If the tube is thick, it is divided in N fictitious, successive and coaxial layers having the same constant thickness, represented by the nondimensional thickness parameter $h = (\beta - 1)/N$, and the same material

properties. Upon choosing a suitably large value of N, each individual layer becomes itself a sufficiently thin elastic tube and, as a result, an approximate solution of the form (5-20) is considered satisfactory for the study of its behaviour. The approximate solutions thus obtained for all fictitious layers are then suitably connected by means of appropriate continuity conditions imposed on their fictitious interfaces, thus providing an arbitrarily close solution to that of the exact system (5-17) — see [Shuvalov and Soldatos 2003]. For an illustration of the relevant algorithm, consider the j-th such individual fictitious layer ($j = 1, 2, \ldots, N$), the nondimensional middle-radius parameter of which is given by

$$R^{(j)} = 1 + (h/2)(2j - 1), \tag{5-21}$$

giving thus rise to the exponential matrix

$$[K^{(j)}(r)] = \exp[(r - R^{(j)} + h/2)D(R^{(j)})], \qquad R^{(j)} - h/2 \leq r \leq R^{(j)} + h/2. \tag{5-22}$$

By requiring continuity of the azimuthal displacement component, the in-plane rotation component and the nonzero components of the stress and couple-stress tensors, one obtains the following continuity conditions on the $N - 1$ fictitious interfaces:

$$\{X^{(j)}(R^{(j)} + h/2)\} = \{X^{(j+1)}(R^{(j+1)} - h/2)\}, \qquad j = 1, 2, \ldots, N - 1. \tag{5-23}$$

Hence, with recursive use of (5-20), (5-22) and (5-23), the solution sought is constructed as follows:

$$\begin{aligned}
\{X^{(N)}(R^{(N)} + h/2)\} &= [K^{(N)}(R^{(N)} + h/2)]\{X^{(N-1)}(R^{(N-1)} + h/2)\} \\
&= [K^{(N)}(R^{(N)} + h/2)][K^{(N-1)}(R^{(N-1)} + h/2)]\{X^{(N-2)}(R^{(N-2)} + h/2)\} \\
&= \cdots = [\bar{K}]\{X^{(1)}(R^{(1)} - h/2)\},
\end{aligned} \tag{5-24}$$

where

$$[\bar{K}] = \left[\prod_{j=N}^{1} [K^{(j)}(R^{(j)} + h/2)] \right]. \tag{5-25}$$

With further use of the boundary conditions (5-9)$_{2,3,4}$ and (5-10), Equation (5-24) leads to a linear algebraic system (see [Soldatos and Ye 1995], for example), whose solution yields the distribution of the azimuthal displacement component throughout the tube cross-section. Note that the solution of the problem has been obtained by making use of algebraic manipulations involving 4×4 matrices only.

6. Numerical results and discussion

Equations (4-2) and (4-3) make clear that a convenient way for presentation of numerical results when fibres are perfectly flexible ($\lambda = 0$) is associated with the use of the nondimensional quantities

$$\bar{v}(r) = \frac{\beta^2 - 1}{\beta} v(r) = r - r^{-1}, \qquad \bar{t}_{r\theta}(r) = \frac{\beta^2 - 1}{\beta} t_{r\theta}(r) = 2r^{-2}. \tag{6-1}$$

It is observed that neither of these nondimensional quantities depends on the nondimensional radius parameter β of the tube outer boundary. This observation suggests that both $\bar{v}$ and $\bar{t}_{r\theta}$ maintain the same distribution profile regardless of the tube thickness. This is of course not the case when fibres resist bending but, for convenience in the presentation of numerical results, the azimuthal displacement and

the shear stress distributions predicted in Section 5 are also nondimensionalized further in accordance with (6-1). In this context the additional nondimensional stress parameter

$$\bar{t}_{\theta r} = \frac{\beta^2 - 1}{\beta}\, t_{\theta r}$$

is also employed where necessary. It should be emphasized that the profiles of the nondimensional displacement and stress distributions employed do depend on β when fibres resist bending. For a straightforward interpretation of the presented numerical results, it is convenient to assume that $\psi > 0$ and, hence, that the cause of the deformation is applied anticlockwise on the outer tube boundary.

Most of the numerical results shown next are related with relatively thick tubes and they are mainly produced by solving equation (5-9)$_1$ on the basis of the SAM outlined in Section 5.4. It is worth noting that corresponding numerical results obtained on the basis of the power series method (Section 5.3) are practically identical to those based on SAM and, hence, in line with the conclusions made in [Shuvalov and Soldatos 2003], the two methods are found to be computationally equivalent. However, due mainly to its slow convergence, the power series method seems to be computationally reliable for relatively thin tubes only. In contrast, SAM converges faster and is computationally reliable for a much wider range of the tube thicknesses. Very satisfactory convergence of SAM and accuracy of the obtained results was achieved by choosing $h/R^{(j)} < 0.01$ where h and $R^{(j)}$ are defined in Section 5.4; this fact is also in agreement with similar observations made in previous studies that were based on SAM (e.g., [Soldatos and Ye 1995]). It is convenient at this point to also note that numerical results shown in Figures 2–5 are obtained under the assumption that the geometric boundary condition (5-10)$_1$ is applied at the fibre root, $r = 1$, while corresponding results plotted in Figures 6-8 are obtained by assuming that the natural boundary condition (5-10)$_2$ is applied there.

Figure 2 depicts the first quadrant of the tube cross-section and, for different values of λ, shows the deformation pattern of a fibre initially aligned along the horizontal radius of the tube cross-section having its outer boundary at $\beta = 2.5$. It is recalled that $\lambda = 0$ represents the deformation pattern of a perfectly
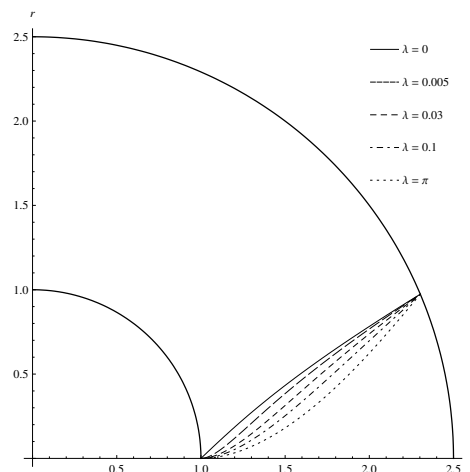

Figure 2. Deformation pattern of a fibre initially aligned along the horizontal radius of the tube cross-section. The fibre is assumed clamped in the inner tube boundary.

flexible fibre described by (4-2). As already mentioned, this pattern is identical to that attained by the horizontal radius of a corresponding isotropic material; having no bending stiffness, a perfectly flexible radial fibre just seems to follow passively the deformation pattern of an isotropic material. Corresponding deformation patterns for a fibre that resists bending ($\lambda \neq 0$) are plotted in Figure 2 under the assumption that the additional boundary condition $(5\text{-}10)_1$ is applied on the tube inner boundary; hence, the fibre is essentially assumed clamped there. It is observed that the slope of the perfectly flexible fibre is nonzero at the inner boundary of the tube cross-section and is monotonically decreasing with increasing β. Hence, a deformed perfectly flexible fibre is concave downwards. In contrast, in their deformed configuration, fibres with bending stiffness are initially concave upwards and, hence, they resist the applied azimuthal shear deformation. This resistance increases with λ and reflects on the slope of the deformed fibre which is initially increasing from its zero value imposed on the inner tube boundary. Although the slope of the fibre begins afterwards to decrease again when λ is small, the region of monotonically increasing slope values becomes larger with increasing λ and, therefore, with increasing fibre bending resistance, in line with physical expectation. For sufficiently large values of λ, the slope of the fibre deformation pattern seems to become monotonically increasing throughout the tube cross-section ($1 \leq r \leq 2.5$) and, hence, the whole fibre is concave upwards. It is finally noted that the deformed fibre pattern shown in Figure 2 for $\lambda = \pi$ remains practically unchanged if λ is increased further.

The dimensionless azimuthal displacement $\bar{v}$ is plotted in Figure 3 against r, for different values of β and for $\lambda = 0.1$. As already mentioned, in the perfectly flexible fibres case ($\lambda = 0$), $\bar{v}$ maintains the same distribution profile regardless of the value of β. However, each one of the dashed lines ($\lambda \neq 0$) begins at the tube inner boundary, as required by the boundary condition $(5\text{-}9)_2$, and ends at some different point of the solid line; namely, at the point β where the external azimuthal displacement ψ is applied on. It is seen that, for $\lambda \neq 0$, $\bar{v}$ is decreasing at the vicinity of the tube inner boundary with increasing β. This is in line with the expectation that, upon increasing the tube outer radius, the effect of the external cause of the deformation is decreasing at the root of the clamped fibre where the highest bending resistance is observed.
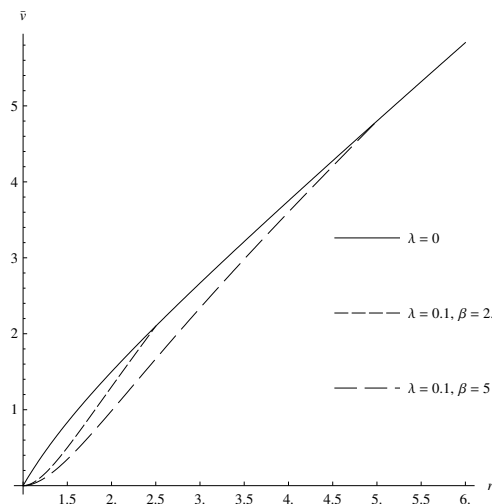

Figure 3. Nondimensional displacement $\bar{v}$ as a function of r for different values of β ($\lambda = 0.1$).
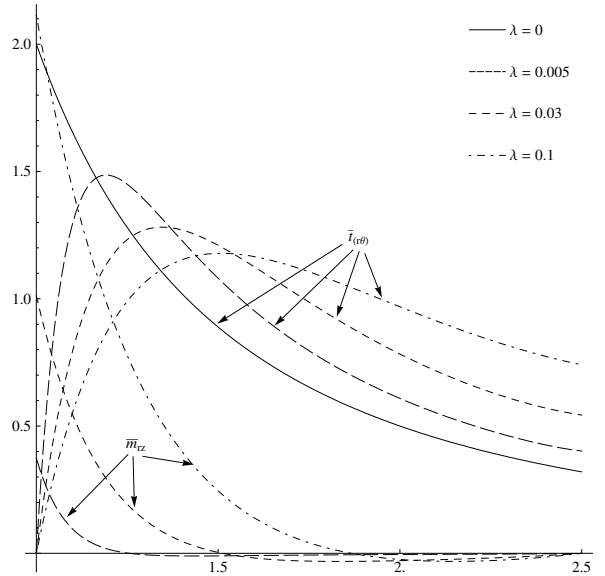
Figure 4. Symmetric part of nondimensional shear stress $\bar{t}_{(r\theta)}$ and nondimensional couple stress $\bar{m}_{rz} = \frac{\beta^2-1}{\beta\psi C_{66}} m_{rz}$ as a function of r for different values of λ ($\beta = 2.5$).

For different values of λ and for $\beta = 2.5$, Figure 4 shows the variation of the symmetric part of the nondimensional shear stress $\bar{t}_{r\theta}$ and the nondimensional couple-stress $\bar{m}_{rz}$ against r. It is recalled that $\bar{t}_{r\theta} = \bar{t}_{\theta r}$ for the perfectly flexible fibre case ($\lambda = 0$); this fact is represented by the solid line in the figure. As a result of the displacement boundary conditions (5-9)$_2$ and (5-10)$_1$, $\bar{t}_{(r\theta)}$ is zero at the inner tube boundary; see also (5-1). It follows that the couple-stress $\bar{m}_{rz}$ and, therefore, the antisymmetric part of the shear stress, $\bar{t}_{[r\theta]}$, are dominant at the vicinity of the inner tube boundary; in fact the couple-stress takes naturally its maximum value at the inner tube boundary where highest fibre bending resistance is observed. However, as distance from the inner tube boundary is increasing, the contribution of $\bar{t}_{(r\theta)}$ is increasing while that of $\bar{m}_{rz}$ and $\bar{t}_{[r\theta]}$ is decreasing fast and becomes gradually negligible. As λ increases, $\bar{t}_{(r\theta)}$ decreases within the inner part of the tube cross-section. This decrease of $\bar{t}_{(r\theta)}$ is compensated by the increasing contribution of $\bar{m}_{rz}$ while the outlined trend is reversed within the outer part of the tube cross-section.

For different values of λ and for $\beta = 2.5$, Figure 5, top, shows the distribution of the shear stresses $\bar{t}_{r\theta}$ and $\bar{t}_{\theta r}$ within the tube cross-section. It is seen that, maximum absolute shear stress occurs always at the inner tube boundary, though $\bar{t}_{r\theta}$ and $\bar{t}_{\theta r}$ take opposite values there for $\lambda \neq 0$; this is due to the fact that $\bar{t}_{(r\theta)} = 0$ at $r = 1$, as observed in Figure 4. The absolute value of maximum shear stress increases with increasing the value of λ but, away from the fibre root, $\bar{t}_{r\theta}$ decreases gradually from its maximum positive value while $\bar{t}_{\theta r}$ increases sharply from its corresponding negative minimum value. This is due to the fact that fibre bending resistance has not a dominant effect away from the inner tube boundary. Hence, the stress tensor becomes nearly symmetric outside a certain layer in the vicinity of the inner tube boundary, where negative $\bar{t}_{\theta r}$ values of large magnitude are observed; the width of that layer is naturally increasing with increasing the fibre bending stiffness. For different values of β and for $\lambda = 0.1$, Figure 5, bottom, shows the distribution of $\bar{t}_{r\theta}$ and $\bar{t}_{\theta r}$ within the tube cross-section. As already mentioned, $\bar{t}_{r\theta}$ maintains
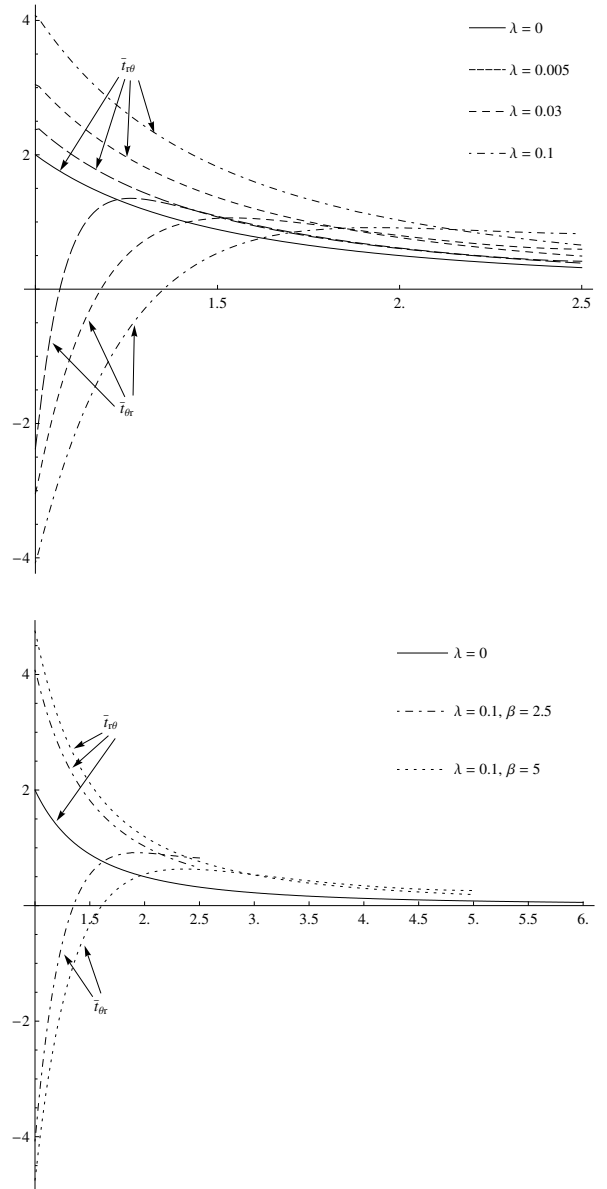
Figure 5. Distribution of shear stresses $\bar{t}_{r\theta}$ and $\bar{t}_{\theta r}$ for $\beta = 2.5$ and different values of λ (top), and for $\lambda = 0.1$ and different values of β (bottom).

again the same distribution profile in the perfectly flexible fibres case (solid line) regardless of the value of β. The figure shows that, when fibres resist bending ($\lambda \neq 0$), $\bar{t}_{r\theta}$ is increasing while $\bar{t}_{\theta r}$ is decreasing near the tube inner boundary with increasing β. It is observed that, the thicker is the tube the nearer the dashed lines approach the solid line at the outer tube boundary. Hence, for sufficiently thick tubes, the effects of fibre bending resistance are essentially confined within the aforementioned layer formed in the neighbourhood of the inner tube boundary; they are not felt in the vicinity of the outer tube boundary.
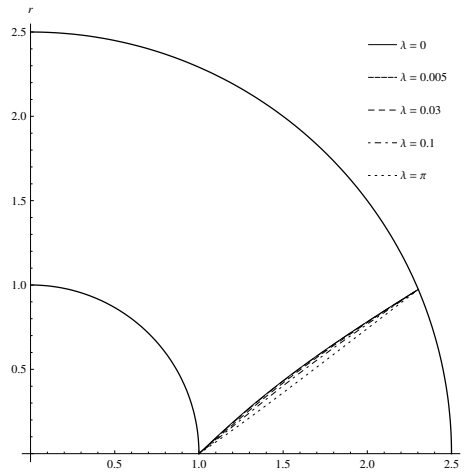
Figure 6. Deformation pattern of a fibre initially aligned along the horizontal radius of the tube cross-section; the fibre is assumed simply supported in the inner tube boundary.

When the geometrical boundary condition $(5\text{-}10)_1$ is replaced by the natural boundary condition $(5\text{-}10)_2$, the numerical values and the physical trends of the results plotted in Figures 2, 4 and 5, top, change dramatically and transform into those depicted in Figures 6, 7 and 8, respectively. Since the root of the fibre is essentially subjected to a simply supported type of boundary condition, the slope of the fibre is nonzero at $r = 1$; see Figure 6. However, the fibre still exhibits signs of bending resistance which, according to Figure 7, seem to emerge slightly further away from the fibre root, as soon as nonzero values of $\bar{m}_{rz}$ become influential. Nevertheless, the total bending resistance of the fibre influences considerably
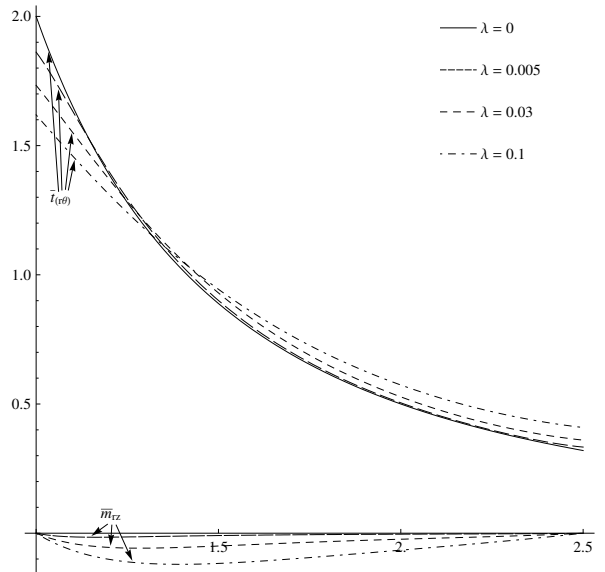

Figure 7. Symmetric part of nondimensional shear stress $\bar{t}_{(r\theta)}$ and nondimensional couple stress $\bar{m}_{rz} = \frac{\beta^2 - 1}{\beta \psi C_{66}} m_{rz}$ as a function of r for different values of λ ($\beta = 2.5$).
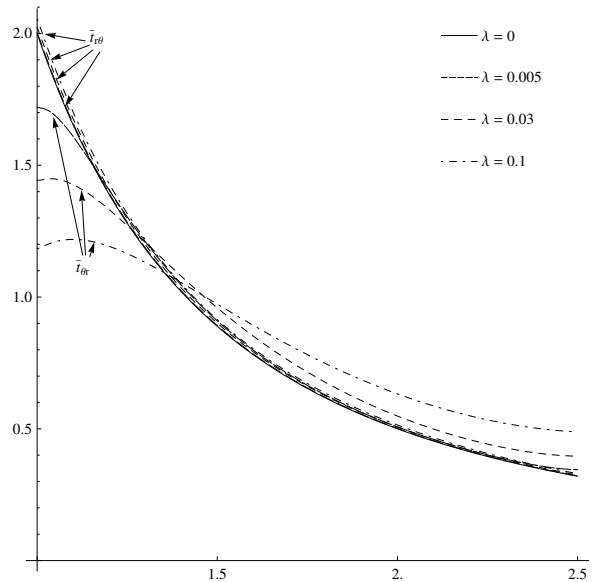
Figure 8. Distribution of shear stresses $\bar{t}_{r\theta}$ and $\bar{t}_{\theta r}$ for different values of λ ($\beta = 2.5$).

the value of the fibre slope at $r = 1$; see Table 2. As a result, the main difference between the values of $\bar{t}_{r\theta}$ and $\bar{t}_{\theta r}$ is again confined within the aforementioned layer near the inner tube boundary (Figure 8). It is also seen (Figure 6) that the maximum possible deviation of the deformed fibre from the deformed shape of its perfectly flexible counterpart is now naturally smaller than that observed in Figure 2.

λ	0	0.005	0.03	0.1	π
$v'(1)$	0.952	0.887	0.825	0.771	0.674

Table 2. The value of the boundary slope of a fibre supported according to (5-10)$_2$.

It is worth observing in this regard that for large values of λ (e.g., $\lambda = \pi$ in Figure 6) the shape of the deformed fibre is approximately still that of a straight line (that shape remains practically unchanged if the value of λ is increased further). A connection is therefore made between this observation and a conclusion drawn in [Soldatos 2009a; 2010], according to which, in the case of an ideal fibre-reinforced material, the inextensible radial fibres involved do not bend during azimuthal shear deformation; instead they remain straight during deformation and they force the tube cross-section to undergo area-preserving azimuthal shear strain, by changing their slope only. Nevertheless, small area-preserving and small pure azimuthal shear strain are essentially identical deformations in the present case of interest. Hence, radial fibres possessing high bending resistance ($\lambda \geq \pi$) appear to remain practically straight during deformation but they also extend in a manner that satisfies the conditions of pure azimuthal shear strain. If the fibres were inextensible, they would necessarily force the inner and outer tube boundaries to move, as observed in those same references, where the tube material was also assumed to be incompressible. However, as already mentioned, linear elasticity cannot adequately account for the effects that the constraints, of material incompressibility and/or fibre inextensibility have on the azimuthal shear deformation when the

tube is reinforced along the radial direction of its cross-section. In this regard, recall that the relevant area preserving azimuthal shear deformation observed in [Soldatos 2009a; 2010] for ideal fibre-reinforced materials takes place only in the finite elastic deformation regime.

7. Closure

In dealing with the principal questions addressed in the Introduction, it has been found that, when spiral fibres are perfectly flexible, both the inner and the outer radii of the tube exhibit a change for all four versions of the azimuthal shear strain problem considered in Section 3; this change is caused by the existence of coupling between azimuthal shear strain and radial stretching. Nevertheless, it has been also observed that conditions of pure azimuthal shear are possible in all but one of the cases considered and studied in Section 3, though different relevant conditions and/or requirements may apply to different versions of the problem. The only version of the problem for which small pure azimuthal shear strain is not possible is that of the ideal fibre-reinforced material discussed in Section 3.4. This result is in complete agreement with the relevant conclusion made in [Soldatos 2009a; 2010] according which, a tube made from an ideal fibre-reinforced material should instead be expected to undergo area preserving azimuthal shear strain. It is however also noted that area preserving azimuthal shear strain for ideal fibre-reinforced materials [Soldatos 2009a; 2010] is possible only within the finite elastic deformation regime.

When the fibres are straight and aligned along the radial direction of the tube cross-section, radial stretching and azimuthal shear strain become completely uncoupled deformations regardless of whether the tube material is constrained or not, and regardless of whether fibres are perfectly flexible or resist bending. In the perfectly flexible fibres case, the description and, hence, the solution of the problem becomes identical to that met in isotropic elasticity. It is therefore observed that conventional linear elasticity theory cannot adequately account for the effects that the material anisotropy and/or either of the constraints of material incompressibility and fibre inextensibility have on the azimuthal shear strain problem considered when the tube is reinforced along the radial direction of its cross-section.

On the other hand, effects of material anisotropy can be accounted for when fibres posses bending stiffness, by taking into consideration the action of couple-stress and therefore asymmetric stress. It is also seen that the natural appearance of an intrinsic material length parameter, which is representative of the fibre thickness, provides ability for consideration of the manner in which the fibres are supported on the tube boundaries. Hence, when the fibres are assumed clamped in the tube inner boundary, considerable fibre bending resistance is observed within a certain layer in the neighbourhood of the inner tube boundary; there, the fibres appear concave against the imposed deformation, which is in line with physical expectations. It is also observed that the absolute maximum values of the couple-stress and the shear stresses occur at the inner tube boundary where highest bending resistance is anticipated; these values stay influential within the aforementioned layer near the inner tube boundary. When the fibres are assumed simply supported at the inner tube boundary, their slope is naturally nonzero there. However, the fibres still exhibit bending resistance which emerges slightly further away from the fibre root, as soon as nonzero values of the couple-stress become influential. It is also seen that the maximum possible deviation of the deformed simply supported fibre from the deformed shape of its perfectly flexible counterpart is naturally smaller than that observed in the clamped fibre case.

It is also observed that, for large values of fibre bending stiffness, the deformed simply supported radial fibre remains approximately straight. This observation is in line with a conclusion drawn in [Soldatos 2009a; 2010] according to which, in the case of an ideal fibre-reinforced material, the inextensible radial fibres involved do not bend during azimuthal shear deformation; instead they remain straight during deformation and they force the tube cross-section to undergo area-preserving azimuthal shear strain by changing their slope only. Nevertheless, small area-preserving and small pure azimuthal shear strain are essentially identical deformations. Hence, radial fibres possessing high bending resistance appear to remain practically straight during deformation but they also extend in a manner that satisfies the conditions of pure azimuthal shear strain. If the fibres were inextensible, they would necessarily force the inner and outer tube boundaries to move, as observed in [Soldatos 2009a; 2010], where the tube material was also assumed to be incompressible.

Appendix: Explicit form of auxiliary parameters and formulas

The constants γ_1, γ_2 and γ_3 appearing in (3-8) are as follows:

$$\gamma_1 = \frac{2(\bar{C}_{16} + \bar{C}_{26})}{\bar{C}_{22} - \bar{C}_{11}}, \qquad \gamma_2 = -\frac{(\eta\bar{C}_{16} + \bar{C}_{26})}{\eta - 1}, \qquad \gamma_3 = -\frac{(\eta\bar{C}_{16} - \bar{C}_{26})}{\eta + 1}. \qquad (\text{A-1})$$

The constants F_k $(k = 1, 2, \ldots, 6)$ appearing in (3-10) are as follows:

$$\begin{aligned}
&F_1 = \gamma_1(\bar{C}_{12} - \bar{C}_{11}) - 2\bar{C}_{16}, &&F_4 = \gamma_1(\bar{C}_{22} - \bar{C}_{12}) - 2\bar{C}_{26}, \\
&F_2 = \bar{C}_{12} + \eta\bar{C}_{11} + \gamma_2\bar{C}_{16}(\eta - 1), &&F_5 = \bar{C}_{22} + \eta\bar{C}_{12} + \gamma_2\bar{C}_{26}(\eta - 1), \\
&F_3 = \bar{C}_{12} - \eta\bar{C}_{11} - \gamma_3\bar{C}_{16}(\eta + 1), &&F_6 = \bar{C}_{22} - \eta\bar{C}_{12} - \gamma_3\bar{C}_{26}(\eta + 1).
\end{aligned} \qquad (\text{A-2})$$

For the set of boundary conditions (3-4) and (3-5), the arbitrary nonzero constants appearing in (3-8) are found to be

$$\bar{A}_1 = \frac{\beta F_2 F_3(\beta^{2\eta} - 1)}{\bar{A}}, \qquad \bar{A}_2 = \frac{\beta F_1 F_3(\beta^{\eta-1} - 1)}{\bar{A}}, \qquad \bar{A}_3 = \frac{-\beta^\eta F_1 F_2(\beta^{\eta+1} - 1)}{\bar{A}},$$

$$\bar{A}_4 = \frac{\beta F_3(F_2 - \gamma_2 F_1) - \beta^{2\eta+1} F_2(F_3 - \gamma_3 F_1) + \beta^\eta F_1(\gamma_2 F_3 - \gamma_3 F_2)}{\bar{A}}, \qquad (\text{A-3})$$

where

$$\bar{A} = \beta^2 F_3(F_2 - \gamma_2 F_1)(\beta^{2\eta-2} + 1) - F_2(F_3 - \gamma_3 F_1)(\beta^{2\eta+2} + 1) + 2\beta^{\eta+1} F_1(\gamma_2 F_3 - \gamma_3 F_2). \qquad (\text{A-4})$$

For the set of boundary conditions (3-4) and (3-6), the corresponding nonzero constants are found to be

$$\bar{A}_1 = \frac{-\beta(\beta^{2\eta} - 1)}{\bar{A}}, \qquad \bar{A}_2 = \frac{\gamma_1\beta(\beta^{\eta-1} - 1)}{\bar{A}}, \qquad \bar{A}_3 = \frac{\beta^\eta(\beta^{\eta+1} - 1)}{\bar{A}},$$

$$\bar{A}_4 = \frac{\beta(\gamma_1\gamma_2 - 1) - \beta^\eta\gamma_1(\gamma_2 - \gamma_3) + \beta^{2\eta+1}(1 - \gamma_1\gamma_3)}{\bar{A}}, \qquad (\text{A-5})$$

where

$$\bar{A} = \beta^2\gamma_1\gamma_2(\beta^{\eta-1} - 1)^2 - \gamma_1\gamma_3(\beta^{\eta+1} - 1)^2 + (\beta^2 - 1)(\beta^{2\eta} - 1). \qquad (\text{A-6})$$

For the set of boundary conditions (3-4) and (3-5), the arbitrary constants appearing in (3-14) are found to be

$$\tilde{A}_1 = \frac{2\beta(\bar{C}_{16} - \bar{C}_{26})}{(\beta^2 - 1)(\bar{C}_{11} + \bar{C}_{22} - 2\bar{C}_{12})}, \qquad \tilde{A}_2 = \frac{\beta}{\beta^2 - 1}, \qquad \tilde{A}_3 = \frac{-\beta}{\beta^2 - 1}. \qquad \text{(A-7)}$$

For the set of the boundary conditions (3-4) and (3-16), these constants are found to be

$$\tilde{A}_2 = \frac{\beta}{\beta^2 - 1}, \qquad \tilde{A}_3 = \frac{-\beta}{\beta^2 - 1}, \qquad \tilde{A}_4 = \frac{\bar{C}_{16} - \bar{C}_{26}}{\beta(\beta^2 - 1)}. \qquad \text{(A-8)}$$

The constants β_1, β_2, $\hat{\gamma}_1$ and $\hat{\gamma}_2$ appearing in (3-20) are as follows:

$$\beta_1 = \frac{m \cot \alpha + \tan \alpha}{m - 1}, \qquad \beta_2 = \frac{m \cot \alpha - \tan \alpha}{m + 1},$$

$$\hat{\gamma}_1 = \frac{C_{22}}{2} \sin 2\alpha + \frac{\cos^3 \alpha}{\sin \alpha}(\cos^4 \alpha + \sin^4 \alpha),$$

$$\hat{\gamma}_2 = \frac{C_{22}}{2} \sin 2\alpha + \frac{\sin^3 \alpha}{\cos \alpha}(\cos^4 \alpha + \sin^4 \alpha). \qquad \text{(A-9)}$$

The constants H_k $(k = 1, 2, \ldots, 6)$ appearing in (3-22) are as follows:

$$H_1 = \bar{C}_{12} - \bar{C}_{11} + 2\bar{C}_{16} \cot 2\alpha + \cot \alpha \left(\bar{C}_{16} - \bar{C}_{26} - 2 \cot 2\alpha + \hat{\gamma}_1(1 - m^2)\right),$$

$$H_2 = m\bar{C}_{11} + \bar{C}_{12} - \bar{C}_{16}\beta_1(m - 1) + m \cot^2 \alpha + 1 - \cot \alpha(m\bar{C}_{16} + \bar{C}_{26}),$$

$$H_3 = \bar{C}_{12} - m\bar{C}_{11} + \bar{C}_{16}\beta_2(m + 1) + 1 - m \cot^2 \alpha + \cot \alpha(m\bar{C}_{16} - \bar{C}_{26}),$$

$$H_4 = \bar{C}_{22} - \bar{C}_{12} + 2\bar{C}_{26} \cot 2\alpha + \tan \alpha \left(\bar{C}_{16} - \bar{C}_{26} - 2 \cot 2\alpha + \hat{\gamma}_1(1 - m^2)\right),$$

$$H_5 = m\bar{C}_{12} + \bar{C}_{22} - \bar{C}_{26}\beta_1(m - 1) + m + \tan^2 \alpha - \tan \alpha(m\bar{C}_{16} + \bar{C}_{26}),$$

$$H_6 = \bar{C}_{22} - m\bar{C}_{12} + \bar{C}_{26}\beta_2(m + 1) + \tan^2 \alpha - m + \tan \alpha(m\bar{C}_{16} - \bar{C}_{26}). \qquad \text{(A-10)}$$

For the set of boundary conditions (3-4) and (3-5), the arbitrary constants appearing in (3-20) are found to be

$$\hat{A}_1 = \frac{\beta\hat{\gamma}_1 H_2 H_3 (1 - m^2)(\beta^{2m} - 1)}{\hat{A}},$$

$$\hat{A}_2 = \frac{-\beta H_1 H_3 (\beta^{m-1} - 1)}{\hat{A}}, \qquad \hat{A}_3 = \frac{-\beta^m H_1 H_2 (\beta^{m+1} - 1)}{\hat{A}},$$

$$\hat{A}_4 = \frac{\beta \left(H_2 H_3 (\beta^{2m} - 1) \cot 2\alpha - \beta^{m-1} H_1 (H_3\beta_1 - H_2\beta_2) - H_1(\beta^{2m} H_2\beta_2 - H_3\beta_1)\right)}{\hat{A}}, \qquad \text{(A-11)}$$

where

$$\hat{A} = H_1(\beta^{2m} H_3\beta_1 - H_2\beta_2) + H_2 H_3(\beta^2 - 1)(\beta^{2m} - 1) \cot 2\alpha$$
$$- \beta^2\left(2\beta^{m-1} H_1(H_3\beta_1 - H_2\beta_2) + H_1(\beta^{2m} H_2\beta_2 - H_3\beta_1)\right). \qquad \text{(A-12)}$$

For the set of boundary conditions (3-4) and (3-6), these constants are found to be

$$\hat{A}_1 = \frac{\beta\hat{\gamma}_1(1-m^2)(\beta^{2m}-1)}{\hat{A}}, \qquad \hat{A}_2 = \frac{-\beta(\beta^{m-1}-1)}{\hat{A}}, \qquad \hat{A}_3 = \frac{-\beta^m(\beta^{m+1}-1)}{\hat{A}},$$

$$\hat{A}_4 = \frac{\beta^m(\beta_1-\beta_2) - \beta(\beta_1 - \cot 2\alpha) + \beta^{2m+1}(\beta_2 - \cot 2\alpha)}{\hat{A}}, \tag{A-13}$$

where

$$\hat{A} = \beta_1\beta^2(\beta^{m-1}-1)^2 - \beta_2(\beta^{m+1}-1)^2 + (\beta^2-1)(\beta^{2m}-1)\cot 2\alpha. \tag{A-14}$$

For the set of boundary conditions (3-4) and (3-5), the arbitrary nonzero constants appearing in (3-25) are found to be

$$A_1 = \frac{\beta\tan 2\alpha}{\beta^2-1}, \qquad A_2 = \frac{\beta}{\beta^2-1},$$

$$A_4 = \frac{\beta\sec 2\alpha\tan 2\alpha}{\beta^2-1}\left(\bar{C}_{11} + \bar{C}_{22} - 2\bar{C}_{12} - 2(\bar{C}_{16} - \bar{C}_{26})\cot 2\alpha\right). \tag{A-15}$$

References

[Dorfmann et al. 2010] A. Dorfmann, J. Merodio, and R. W. Ogden, "Non-smooth solutions in the azimuthal shear of an anisotropic nonlinearly elastic material", *J. Eng. Math.* **68**:1 (2010), 27–36.

[Jones 1998] R. M. Jones, *Mechanics of composite materials*, Taylor & Francis, 1998.

[Kassianidis et al. 2008] F. Kassianidis, R. W. Ogden, J. Merodio, and T. J. Pence, "Azimuthal shear of a transversely isotropic elastic solid", *Math. Mech. Solids* **13**:8 (2008), 690–724.

[Rivlin 1949] R. S. Rivlin, "Large elastic deformations of isotropic materials, VI: further results in the theory of torsion, shear and flexure", *Philos. Tr. R. Soc. S. A* **242**:845 (1949), 173–195.

[Shuvalov and Soldatos 2003] A. L. Shuvalov and K. P. Soldatos, "On the successive approximation method for three-dimensional analysis of radially inhomogeneous tubes with an arbitrary cylindrical anisotropy", *J. Sound Vib.* **259**:1 (2003), 233–239.

[Soldatos 2003] K. P. Soldatos, "Accurate stress analysis of laminated composite structures", pp. 69–132 in *Modern trends in composite laminates mechanics*, edited by H. Altenbach and W. Becker, Courses and lectures / International Centre for Mechanical Sciences **448**, Springer, 2003.

[Soldatos 2009a] K. Soldatos, "Azimuthal shear deformation of an ideal fibre-reinforced tube according to a second gradient hyper-elasticity theory", session GS–CM – Continuum Mechanics, ID 0017 in *Proc. 7th EUROMECH Solid Mech. Conf.* (Lisbon, 2009), edited by J. Ambrósio et al., 2009.

[Soldatos 2009b] K. P. Soldatos, "Towards a new generation of 2D mathematical models in the mechanics of thin-walled fibre-reinforced structural components", *Int. J. Eng. Sci.* **47**:11-12 (2009), 1346–1356.

[Soldatos 2010] K. Soldatos, "Second-gradient plane deformations of ideal fibre-reinforced materials: implications of hyper-elasticity theory", *J. Eng. Math.* **68**:1 (2010), 99–127.

[Soldatos and Hadjigeorgiou 1990] K. P. Soldatos and V. P. Hadjigeorgiou, "Three-dimensional solution of the free vibration problem of homogeneous isotropic cylindrical shells and panels", *J. Sound Vib.* **137**:3 (1990), 369–384.

[Soldatos and Ye 1995] K. P. Soldatos and J. Q. Ye, "Axisymmetric static and dynamic analysis of laminated hollow cylinders composed of monoclinic elastic layers", *J. Sound Vib.* **184**:2 (1995), 245–259.

[Spencer 1972] A. J. M. Spencer, *Deformations of fibre-reinforced materials*, Clarendon Press, 1972.

[Spencer 1984] A. J. M. Spencer (editor), *Continuum theory of the mechanics of fibre-reinforced composites*, Courses and lectures / International Centre for Mechanical Sciences **282**, Springer, 1984.

[Spencer and Soldatos 2007] A. J. M. Spencer and K. P. Soldatos, "Finite deformations of fibre-reinforced elastic solids with fibre bending stiffness", *Int. J. Non-Linear Mech.* **42**:2 (2007), 355–368.

[Ting 1996] T. C. T. Ting, *Anisotropic elasticity: theory and applications*, Oxford Engineering Science Series **45**, Oxford University Press, New York, 1996.

[Ye 2003] J. Q. Ye, *Laminated composite plates and shells: 3D modelling*, Springer, 2003.

MOHAMED A. DAGHER: pmxmd2@nottingham.ac.uk
Theoretical Mechanics, School of Mathematical Sciences, University of Nottingham, University Park, Nottingham, MG7 2RD, United Kingdom

and

Department of Science and Engineering Mathematics, Faculty of Petroleum and Mining Engineering, Suez Canal University, Suez, Egypt

KOSTAS P. SOLDATOS: kostas.soldatos@nottingham.ac.uk
Theoretical Mechanics, School of Mathematical Sciences, University of Nottingham, University Park, Nottingham, MG7 2RD, United Kingdom

■msp

MEDIA WITH SEMIHOLONOMIC INTERNAL STRUCTURE

MARCELO EPSTEIN

This article is dedicated to the memory of Marie-Louise Steele.

The notions of semiholonomic and quasiholonomic Cosserat media are introduced and their differences outlined. Contrary to the classical holonomic and nonholonomic counterparts, the definition of semi and quasiholonomic media is not kinematic but constitutive. Possible applications include granular media embedded in a rigid matrix and colloidal suspensions in an ideal incompressible fluid.

1. Introduction

The publication in 1909 of the French version of the second volume of Chwolson's *Traité de physique* was generally well received by the scientific establishment of the time. Nevertheless, we find in the issue of *Nature* of July 21, 1910, a remark to the effect that a "note" by MM. E. and F. Cosserat at the end of this volume is 220 pages long and "does not in any sense harmonise" with Chwolson's work, which is "emphatically experimental in character", while the note is "strikingly mathematical". The remark concludes with the suggestion that "MM. Cosserat's note is a distinct and useful treatise, and should be able to stand on its own feet." And so it was, although the author of the review didn't seem to have taken notice. The celebrated book by the Cosserat brothers [1909], was, in fact, identical to that long note plus an additional 67 pages, which had already appeared in the previous volume of Chwolson's treatise. A similar critical remark can be found two years later in a review by Edwin B. Wilson of the Massachusetts Institute of Technology published in the *Bulletin of the American Mathematical Society* (July 1912, pp. 497–508), although six months later (February 1913, pp. 242–246) the same reviewer found it necessary to publish a separate review of the stand-alone book. Both the detailed content and the title of this review ("An advance in theoretical mechanics") demonstrate clearly that the quality of the enterprise of the brothers Cosserat did not go unnoticed.

Today we mainly recognise the contribution of the Cosserats to the modelling of media with internal structure. But this does not seem to have been the main intention of the authors, who had in mind a much wider scope. Their fundamental idea was to try to encompass all physical theories (including perhaps relativity) under the umbrella of a single principle of Euclidean action, a quantity that is invariant under Euclidean transformations. They also advanced the notion of the formal equivalence between a static theory of a deformable n-dimensional manifold and a dynamic theory of a deformable manifold of dimension $n + 1$, whereby the action is interpreted as a space-time entity. Another striking feature of the work is its error-free geometrically nonlinear formulation, particularly in view of the many mistakes made by succeeding generations, for example, in the realm of shell theory. To recover the conventional theory

of elasticity, the Cosserats propose the idea of the hidden triad ("trièdre caché"). Similarly, theories with particular constraints can be recovered by means of the concept of hidden action. Thus, rigid-body mechanics arises as a particular case of their theory. Had they not insisted on the rigidity of the triad, they could have also recovered the theory of second-grade materials by considering what we would call today second-order holonomic frames. Between these frames and the completely nonholonomic frames (corresponding to a micromorphic, rather than micropolar, continuum), there exists an intermediate type consisting of the semiholonomic frames. An intriguing question is whether the "hidden action" corresponding to this type of frames might lead to a physically meaningful interpretation. In this short article it is shown that this is indeed the case if one is willing to consider a material consisting entirely of a micromedium supported by an incoherent matrix. A distinction is drawn between strictly semiholonomic Cosserat media and a related category referred to as quasiholonomic. The presentation is intended to be as self-contained as possible.

2. A one-dimensional picture

The variety of possible Cosserat media[1] resists any attempt at a simple pictorial representation, and this deficiency is exacerbated when the picture is limited to the one-dimensional realm. Thus, for example, rigid-body motions in one dimension are unable to convey the possibility of rotations. Nevertheless, a picture may be of help in providing some insight and motivating further rigorous investigations:
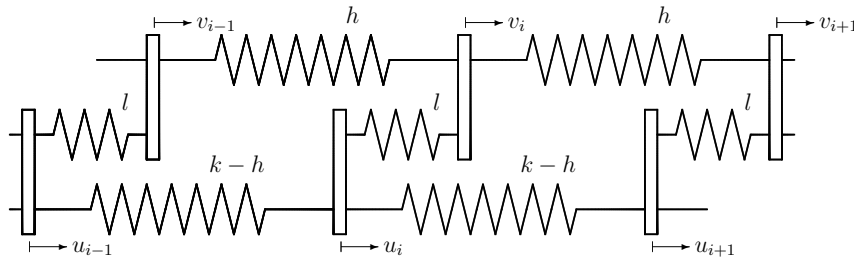
The lower row of material points (represented by the lower three elongated rectangles) symbolizes the macromedium, while the upper counterpart is the micromedium. The corresponding degrees of freedom are indicated, respectively, with the letters u_i and v_i, where the subscript runs over the number of particles. An eventual passage to the continuous limit is suggested, but not directly described. For simplicity, we represent the constitutive equations by means of linear elastic springs, so that, measured from an assumed stress-free reference configuration (which may or may not exist), the elastic energy is given, up to an irrelevant additive constant, by the expression

$$W = \tfrac{1}{2} \sum_i (k-h)(u_{i+1}-u_i)^2 + l(v_i-u_i)^2 + h(v_{i+1}-v_i)^2. \tag{2-1}$$

The following particular values of the stiffness constants k, h, and l are of interest:

• *Arbitrary k, l, and h*: This is the case of the *micromorphic medium*, also called the *nonholonomic* Cosserat medium.

---

[1]Although the terminology "Cosserat medium" is usually reserved for the particular case of a rigid triad (or micropolar continuum), in this paper we use it to denote the general case.

• $h = 0$: This case can be designated as a first-grade macromedium carrying a *zero-grade* micromedium. The grains, so to speak, do not interact elastically with each other. Note that the same physical result is obtained by specifying $h = k$ instead of $h = 0$.

• $l \to \infty$: In this case we must necessarily have $u_i = v_i$, for all i. This condition spells the disappearance of the micromedium. The energy expression reduces to the standard form,

$$W = \tfrac{1}{2} \sum_i k(u_{i+1} - u_i)^2, \tag{2-2}$$

of an elastic *first-grade* material. In the one-dimensional case, the specification $l \to \infty$, by making it rigid, renders the micromedium superfluous. Clearly, in a two- or three-dimensional situation, to achieve the same effect one would need to also specify $h = 0$. Otherwise, the mere rigidity of the micromedium would still allow for an interaction between rigidly rotating microparticles (or *grains*). This is, in fact, the definition of a *micropolar medium*, which is the material originally conceived by the Cosserat brothers in their magnum opus.

• $l = 0$: For the system to remain connected, we attach (in the unstressed state) the (upper) grain to the midpoint of the corresponding lower spring. The energy expression is given by

$$W = \tfrac{1}{2} \sum_i 2(k - h)\big((v_{i+1} - u_{i+1})^2 + (u_{i+1} - v_i)^2\big) + h(v_{i+1} - v_i)^2. \tag{2-3}$$

The lower springs (now double in number) connect between contiguous particles, while the upper springs connect between every second particle. This is the standard representation of a *second-grade* material. Notice that in this case the elastic energy (2-3) can also be written more suggestively as

$$W = \tfrac{1}{2} \sum_i 2k\big((v_{i+1} - u_{i+1})^2 + (u_{i+1} - v_i)^2\big) - h(v_{i+1} - 2u_{+1} + v_i)^2. \tag{2-4}$$

• $k \to \infty$: This is the case of a *rigid macromedium*. We will soon demonstrate that this situation may correspond mathematically to a genuinely *semiholonomic* Cosserat medium.

• $k = 0$: Physically, this case corresponds to an *incoherent matrix* within which the micromedium provides the only degree of elastic coherence. We will use also the terminology *quasiholonomic* medium to refer to this type of material.

3. Cosserat bodies

In continuum mechanics, a *material body* $\mathscr{B}$ is defined as a three-dimensional differentiable manifold that can be covered with a single coordinate chart. A *configuration* κ is defined as an embedding of $\mathscr{B}$ into the three-dimensional Euclidean space $\mathbb{E}^3$:

$$\kappa : \mathscr{B} \longrightarrow \mathbb{E}^3. \tag{3-1}$$

In terms of coordinate charts X^I ($I = 1, 2, 3$) and x^i ($i = 1, 2, 3$) in the body and in space, respectively, the configuration κ is given by three smooth functions:

$$X^I \mapsto x^i = \kappa^i(X^I). \tag{3-2}$$

To convey the presence of extra kinematic degrees of freedom, however, these definitions need to be expanded so that the differential geometry can properly reflect the existence of the microstructure and its

possible deformability. We recall that, given an m-dimensional differentiable manifold $\mathcal{M}$, its *principal frame bundle* $F\mathcal{M}$ is obtained by adjoining at each point $x \in \mathcal{M}$ the collection $F_x\mathcal{M}$ of all the possible bases of its tangent space $T_x\mathcal{M}$. The set thus obtained has a canonical structure of a differentiable manifold of dimension $m + m^2$. It is endowed with the *natural projection* map

$$\pi_M : F\mathcal{M} \longrightarrow \mathcal{M}, \tag{3-3}$$

which assigns to each point $p \in F\mathcal{M}$ the point $\pi(p) \in \mathcal{M}$ to which it is attached. If x^i ($i = 1, \ldots, m$) is a coordinate chart on $U \subset \mathcal{M}$ with natural basis $\boldsymbol{e}_i = \partial/\partial x^i$, we can construct an associated chart in $F\mathcal{M}$ by assigning to each point $p \in \pi_M^{-1}(U)$ the numbers $\{x^i, p_j^i\}$, where p_j^i is the i-th component of the j-th vector of the frame p in the natural basis $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m\}$. Expressed in terms of coordinates, the natural projection is given by

$$x^i, p_j^i \mapsto x^i. \tag{3-4}$$

We define a *Cosserat body* as the principal frame bundle $F\mathcal{B}$ of an ordinary material body $\mathcal{B}$. The physical intent is that, while the underlying body $\mathcal{B}$ represents the *macromedium*, each *fibre* $F_x\mathcal{B}$ represents the *microparticle* or *grain* at $x \in \mathcal{B}$.

Concomitantly with the enlargement of the scope of material bodies, we need to introduce a more general definition of the notion of configuration. To this end, we consider *fibre-preserving* maps

$$K : F\mathcal{B} \longrightarrow F\mathbb{E}^3 \tag{3-5}$$

such that K is a *fibre-bundle morphism* between $F\mathcal{B}$ and its image. By fibre preservation, we mean the commutativity of the diagram:

$$\begin{array}{ccc} F\mathcal{B} & \xrightarrow{\ K\ } & F\mathbb{E}^3 \\ \downarrow{\scriptstyle \pi_B} & & \downarrow{\scriptstyle \pi_E} \\ \mathcal{B} & \xrightarrow{\ \kappa\ } & \mathbb{R}^3, \end{array} \tag{3-6}$$

where κ is a well-defined map between the base manifolds. Thus, a *Cosserat deformation K* automatically implies the existence of an ordinary deformation κ, representing the deformation of the macrostructure. By fibre-bundle morphism we imply that, fibre by fibre, each of the restrictions $K|_X$ ($X \in \mathcal{B}$) commutes with the multiplicative right action of the general linear group $\mathrm{GL}(3; \mathbb{R})$. In terms of coordinates, this means that there exists an X-dependent matrix K_I^i such that any Cosserat configuration is completely defined by twelve smooth functions

$$x^i = \kappa^i(X^J) \tag{3-7}$$

and

$$K_I^i = K_I^i(X^J). \tag{3-8}$$

The physical meaning of these assumptions is that each grain can undergo only homogeneous deformations, as represented by the local matrix K_I^i. In other words, each grain behaves as a pseudorigid body. A more detailed treatment can be found in [Epstein and de León 1996; 1998; Epstein and Elżanowski 2007].

Remark 3.1. As already pointed out, the original formulation by the Cosserat brothers considered the case in which K_I^i is orthogonal. In the terminology of [Eringen 1999], this case corresponds to the

micropolar continuum. The more general case in which K_I^i is an arbitrary nonsingular matrix corresponds to the micromorphic continuum of Eringen. We use the term "Cosserat body" in this more general sense.

We can see that in a Cosserat body there exist two, in principle independent, mechanisms for dragging vectors by means of a deformation: The first mechanism is the ordinary dragging of vectors by means of the deformation gradient of the macromedium, represented by the matrix with entries $F_I^i = x_{,I}^i$. The second mechanism is the one associated with the deformation of the microparticle or grain, and is represented by the matrix with entries K_I^i. Note that in a second-grade body these two mechanisms are identified with each other, thus suggesting that different kinds of Cosserat media may be obtained by either kinematic restrictions of this kind or by constitutive restrictions. In fact, the Cosserat brothers themselves already advanced these possibilities and introduced the outmoded terminology of "trièdre caché" (hidden triad) and "W caché" (hidden strain-energy function) to refer, respectively, to these kinematic or constitutive restrictions. We will follow in their steps.

4. Nonholonomic, semiholonomic, and holonomic jets

Given two smooth manifolds, $\mathcal{M}$ and $\mathcal{N}$, of dimensions m and n, respectively, we say that two maps $f, g : \mathcal{M} \longrightarrow \mathcal{N}$ have the same k-jet at a point $X \in \mathcal{M}$ if: (i) $f(X) = g(X)$, and (ii) in a coordinate chart in $\mathcal{M}$ containing X and a coordinate chart in $\mathcal{N}$ containing the image $f(X)$, all the partial derivatives of f and g up to and including the order k are respectively equal.

Although the above definition is formulated in terms of charts, it is not difficult to show by direct computation that the property of having the same derivatives up to and including order k is in fact independent of the coordinate systems used in either manifold. Notice that, in order for this to work, it is imperative to equate *all* the lower-order derivatives. If, for example, we were to equate just the second derivatives, without regard to the first, the equality of the second derivatives would not be preserved under arbitrary coordinate transformations.

The property of having the same k-jet at a point is, clearly, an equivalence relation. The corresponding equivalence classes are called *k-jets at X*. Any function in a given k-jet is then called a *representative* of the k-jet. The k-jet at X of which a given function $f : \mathcal{M} \longrightarrow \mathcal{N}$ is a representative is denoted by $j_X^k f$. The collection of all k-jets at $X \in \mathcal{M}$ is denoted by $J_X^k(\mathcal{M}, \mathcal{N})$. The point X is called the *source* of $j_X^k f$ and the image point $f(X)$ is called its *target*.

Let a smooth map $f : \mathcal{M} \longrightarrow \mathcal{N}$ be given in terms of coordinates X^I ($I = 1, \ldots, m$) and x^i ($i = 1, \ldots, n$) in $\mathcal{M}$ and $\mathcal{N}$, respectively, by the functions

$$x^i = x^i(X^1, \ldots, X^m), \quad i = 1, \ldots, n. \tag{4-1}$$

The jet $j_X^2 f$, for example, is then given by the coordinate expressions

$$x^i(X^1, \ldots, X^m), \quad \left[\frac{\partial x^i}{\partial X^I}\right]_X, \quad \left[\frac{\partial^2 x^i}{\partial X^J \partial X^I}\right]_X, \tag{4-2}$$

a total of $n + mn + m^2 n$ numbers.

We are particularly interested in the case of 1-jets. Let us evaluate, accordingly, the coordinate expression of $j_X^1 K$, where K is a Cosserat configuration, as defined in coordinates by (3-7) and (3-8). Notice that the dimension of both the source and the target manifolds in this case is 12. Following the definition,

we conclude that $j_X^1 K$ consists of the elements

$$x^i, \quad K_I^i, \quad \left[\frac{\partial x^i}{\partial X^I}\right]_X, \quad \left[\frac{\partial K_I^i}{\partial X^J}\right]_X, \tag{4-3}$$

which we can abbreviate as

$$x^i, \quad K_I^i, \quad F_I^i = x_{,I}^i, \quad K_{I,J}^i. \tag{4-4}$$

If no further restrictions are imposed on K, we speak of the components (4-4) as the representatives of a *nonholonomic* 1-jet at $X \in \mathcal{B}$. It is possible, however, to demand in an intrinsic manner, independent of the coordinates, that the functions K under consideration satisfy the following compatibility requirement in a neighbourhood of X:

$$K_I^i \equiv x_{,I}^i. \tag{4-5}$$

In this case, the collection of 1-jets obtained is smaller. Not only are the second and third entries in (4-4) the same, but also, by virtue of the identical satisfaction of (4-5) in a neighbourhood of X, we must have:

$$K_{I,J}^i = x_{,IJ}^i = K_{J,I}^i. \tag{4-6}$$

In other words, the last element of the jet is symmetric with respect to its lower indices. We will indicate the coordinate expression of these *holonomic jets* as follows:

$$x^i, \quad F_I^i, \quad K_{I,J}^i = K_{J,I}^i. \tag{4-7}$$

Finally, there exists a third type of jet, somewhat intermediate between the two extremes just presented. It is obtained when the potential representatives K are restricted to satisfy the condition

$$K_I^i(X) = x_{,I}^i(X). \tag{4-8}$$

Thus, we demand the satisfaction of (4-5) not identically in a neighbourhood of X, but just at the point X itself. The 1-jets thus obtained are known as *semiholonomic jets*. The coordinate expression of a semiholonomic jet is

$$x^i, \quad K_I^i, \quad K_{I,J}^i. \tag{4-9}$$

Notice that the last entry is no longer necessarily symmetric.

Remark 4.1. Given an actual arbitrary configuration K, it will give rise automatically to point-wise nonholonomic jets. If the configuration is restricted so that condition (4-5) is satisfied over the whole base manifold $\mathcal{B}$, it will give rise to everywhere holonomic jets. In this sense, it is possible to speak of nonholonomic or holonomic configurations, respectively. On the other hand, it is not possible to define semiholonomic configurations. Indeed, if condition (4-8) were to be imposed at each point, we would immediately revert to condition (4-5), thus obtaining a holonomic configuration.

5. Semiholonomic Cosserat media

The last section ended in a definitely pessimistic note. Indeed, if semiholonomic configurations cannot be properly defined, there seems to be no point in attempting a definition of semiholonomic media. This kinematic impasse, however, can perhaps be resolved by means of a constitutive statement. We could say, for example, that a nonholonomic Cosserat medium is semiholonomic if its constitutive equation

involves only the semiholonomic part of the 1-jet of the configuration. Physically, this would correspond to a response that is in some sense oblivious of the presence of the macromedium. In this section, we look into this and other possibilities with some care.

Since we are contemplating a particular case of nonholonomic Cosserat media, it will be useful to record the law governing the change of constitutive law of such a medium under a change of reference configuration. For specificity, we will limit ourselves to a single scalar constitutive law, such as the free-energy density per unit mass ψ. Let the constitutive law with respect to a reference configuration K_0 be given in a coordinate system X^I by the expression

$$\psi = \psi_0(K_I^i, F_I^i, K_{I,J}^i; X^I), \tag{5-1}$$

and let the counterpart for a reference configuration K_1 with coordinates Y^A be given by

$$\psi = \psi_1(K_A^i, F_A^i, K_{A,B}^i; Y^A), \tag{5-2}$$

with an obvious notational scheme. The deformation from K_0 to K_1 is given by twelve quantities, written

$$Y^A = Y^A(X^I), \quad K_I^A(X^J). \tag{5-3}$$

By the law of composition of jets (or derivatives), we obtain between the constitutive expressions the relation

$$\psi_1(K_A^i, F_A^i, K_{A,B}^i; Y^A(X^J)) = \psi_0(K_A^i K_I^A, F_A^i F_I^A, K_{A,B}^i K_I^A F_J^B + K_A^i K_{I,J}^A; X^J), \tag{5-4}$$

where $F_I^A = Y_{,I}^A$.

The point of bringing this transformation equation to bear is the following result, whose proof is an immediate consequence of the transformation law (5-4).

Proposition 5.1. *If the constitutive law* (5-1), *in the reference configuration K_0, is independent of the second argument* (F_I^i), *so is the expression of the same constitutive law in any other reference configuration K_1 independent of the second argument* (F_A^i).

As a direct corollary of this proposition, we can propose the following definition.

Definition 5.2. A nonholonomic Cosserat medium is said to be *semiholonomic* at X if its constitutive law at X is independent of the deformation gradient of the macromedium.

From the mathematical standpoint, it is necessary to note that this definition does not imply the existence of a *canonical* projection of a nonholonomic jet onto a semiholonomic part. In fact, such a canonical projection does not exist. What the definition implies is that once a noncanonical choice is effected in one particular reference configuration, this choice can be convected to all other configurations by means of the correct application of the transformation (5-4). In particular, this convection involves the gradient of the change of reference configuration (F_I^A). Another way to state the choice of a particular "projection" is to say that a particular parallelism (whose physical meaning may, for example, be related to the existence of some particular stress-free configuration) must be chosen as part and parcel of the constitutive law of a semiholonomic Cosserat medium.

From the physical point of view, a semiholonomic Cosserat medium may be said to consist of an incoherent matrix upon which a coherent micromedium has been installed. The interaction between the

grains, however, "remembers" the existence of a particular configuration of the macromedium as the only remaining physical trace of its existence. It is interesting to remark that, since the macromedium plays no other role, the configuration space of a semiholonomic medium may be, in a possible application, assumed to be rigid.

The converse of the above statement is not true: a nonholonomic Cosserat medium with a rigid matrix is not automatically semiholonomic. Indeed, by a direct application of the principle of frame indifference, the constitutive law (5-1) can be reduced to the form

$$\psi = \psi(\boldsymbol{R}^T \boldsymbol{K}, \boldsymbol{U}, \boldsymbol{R}^T \nabla \boldsymbol{K}; X), \tag{5-5}$$

where the polar decomposition $\boldsymbol{F} = \boldsymbol{R}\boldsymbol{U}$ has been exploited and where block letters stand for the collections of homonymous indexed quantities used in previous formulas. Using now the polar decomposition

$$\boldsymbol{K} = \boldsymbol{R}'\boldsymbol{U}', \tag{5-6}$$

we may write (5-5) as

$$\psi = \psi(\boldsymbol{r}\boldsymbol{U}', \boldsymbol{U}, \boldsymbol{r}\boldsymbol{R}'^T \nabla \boldsymbol{K}; X), \tag{5-7}$$

where

$$\boldsymbol{r} = \boldsymbol{R}^T \boldsymbol{R}' \tag{5-8}$$

is the (referential) *relative rotation* of the grain with respect to the macromedium. If the macromedium is rigid, we must have necessarily $\boldsymbol{U} = \boldsymbol{I}$. But for a semiholonomic body the constitutive law must be independent of *both* components $\boldsymbol{U}$ and $\boldsymbol{R}$ of the polar decomposition of $\boldsymbol{F}$. It follows, therefore, that rigidity alone does not imply semiholonomy. If, on the other hand, the constitutive law of a rigid-matrix Cosserat medium is independent of the rotation $\boldsymbol{R}$, we may choose $\boldsymbol{R} = \boldsymbol{R}'$ (or, equivalently, $\boldsymbol{r} = \boldsymbol{I}$), thereby leading to the following reduced equation of a semiholonomic Cosserat body:

$$\psi = \psi(\boldsymbol{U}', \boldsymbol{R}'^T \nabla \boldsymbol{K}; X). \tag{5-9}$$

In the physical interpretation, we may say that the grains are attached to the rigid macromedium by means of ideal frictionless pins, so that there is no energetic cost to produce a relative rotation between them. In the admittedly imperfect pictorial representation of the figure below, the grains in the reference configuration are depicted as squares pin-jointed at their centres to the rigid matrix and connected to their neighbours by means of springs (represented by broken lines) designed to detect differential stretches
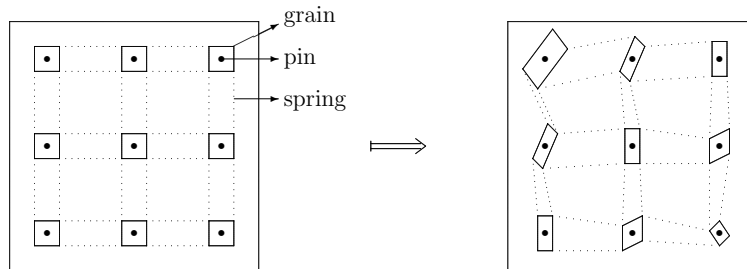
A rigid-matrix semiholonomic Cosserat medium.

and rotations between contiguous grains. The grains themselves behave as *pseudorigid bodies*, so that their deformed versions are represented by parallelograms.

The reduced form (5-9) of the constitutive law of a semiholonomic Cosserat material applies whether or not the matrix is rigid, since in either case the response is independent of both U and R.

6. Quasiholonomic Cosserat media

As defined, a semiholonomic Cosserat medium may not necessarily have any material symmetries. We want to contrast the above definition with the following one that, by demanding the maximum possible symmetry of the macromedium, appears to carry the same physical meaning.

Definition 6.1. A nonholonomic Cosserat medium is said to be *quasiholonomic* at X if, for some (local) reference configuration, its symmetry group $\mathcal{H}$ at X contains the subgroup given by

$$\mathcal{G} = \left\{ \{I, G, 0\} \mid G \in \mathrm{GL}(3; \mathbb{R}) \right\}, \tag{6-1}$$

where I is the unit of $\mathrm{GL}(3; \mathbb{R})$.

The reason to suspect that this definition might be equivalent to the previous one is that, due to the assumed arbitrariness of G, it seems to imply that the deformation of the macromedium plays no role in the constitutive response. A direct application of the definition of a nonholonomic symmetry, however, leads to the conclusion that a quasiholonomic medium must have a constitutive law of the form

$$\psi = \psi(K_I^i, K_{I,J}^i F_j^{-J}; X^I) \tag{6-2}$$

in the special reference configuration used in the definition[2].

Physically, this means that the price to pay for this large symmetry group is, surprisingly, the reappearance of the deformation gradient of the macromedium in the last argument of the constitutive law so as to permit the interaction between the grains to take into account their relative spatial locations (rather than those pulled back to some putative, perhaps unstressed, reference configuration).

The purpose of the following simple example is to shed light on the subtle difference between semiholonomic and quasiholonomic media, as conceived in Definitions 5.2 and 6.1, respectively. To this end, we consider the successive application of two deformations, the first of which can be regarded as a change of reference configuration so as to bring the notation in line with that of the previous section. The (Cartesian) coordinate systems X^I, Y^A, and x^i are assumed to coincide with each other. The first deformation is a uniaxial contraction along the X^1-axis, namely

$$Y^1 = 0.8X^1, \quad Y^2 = X^2, \quad Y^3 = X^3, \quad K_I^A = \delta_I^A. \tag{6-3}$$

The second deformation is a microrotation about the Y^3 axis that increases linearly with Y^1. Specifically:

$$x^1 = Y^1, \quad x^2 = Y^2, \quad x^3 = Y^3, \quad \{K_A^i\} = \begin{bmatrix} \cos\left(\frac{\pi}{3}Y^1\right) & -\sin\left(\frac{\pi}{3}Y^1\right) & 0 \\ \sin\left(\frac{\pi}{3}Y^1\right) & \cos\left(\frac{\pi}{3}Y^1\right) & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{6-4}$$

---

[2]In any other reference configuration, the symmetry group will contain a conjugate of the group $\mathcal{G}$ and the form of the constitutive law will be, accordingly, somewhat more involved.
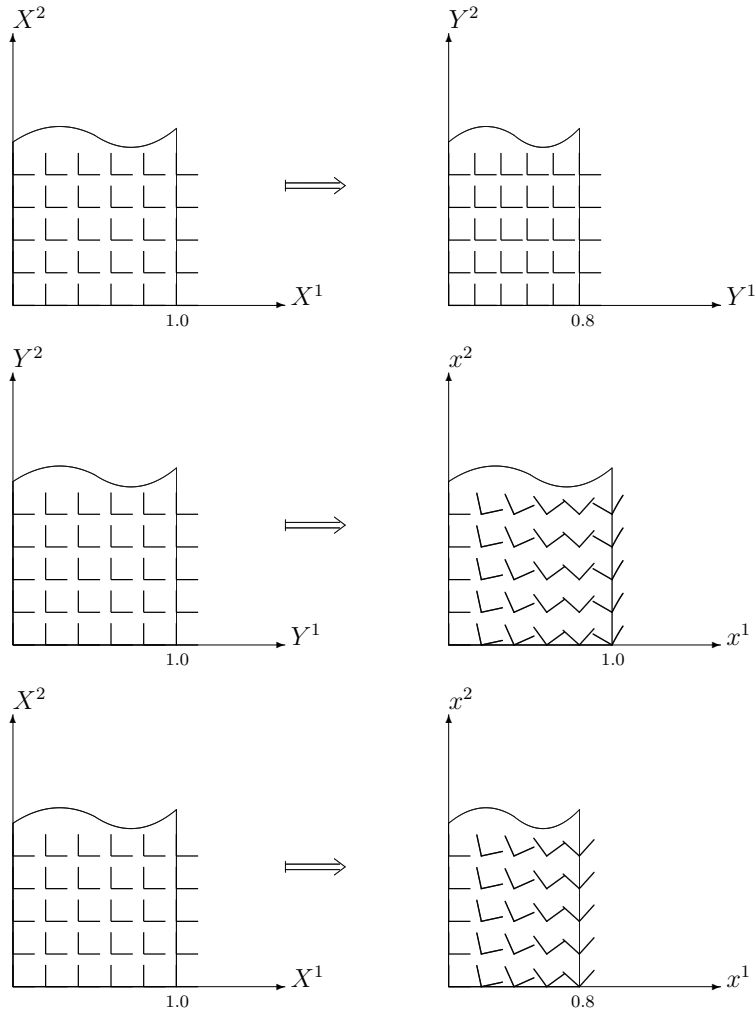
First deformation (top), second deformation (middle), and their composition (bottom).

The effect of each of the two deformations on a unit-width strip in the X^1, X^2 and Y^1, Y^2 planes is shown in the figure above (top and middle parts), together with their composition (bottom part). Notice that, at the moment of composition, it is the *already contracted* strip that encounters the values of the rotation field already in place (as dictated by the second deformation), thus resulting in a maximum value for the rotation of the grain in the deformed strip of 48° rather than 60°, which was the value at the right-hand end of the strip as far as the second deformation alone was concerned. If the Cosserat body is semiholonomic, the gradient of the rotation would be obtained by dividing 48° by the original unit width. On the other hand, if the Cosserat body is quasiholonomic, it is the width measured in the final deformed configuration that matters in the calculation of the gradient. Since this width is of 0.8, we verify that the rotation gradient in the composite deformation turns out to be identical to the gradient in the second deformation. In other words, the preapplication of the first deformation (in this case a contraction of the macromedium) is irrelevant for a quasiholonomic medium.

Among various possible physical applications of both semiholonomic and quasiholonomic Cosserat media, beyond those with a rigid matrix, we mention the modelling of aggregates [Zhang et al. 2006], such as colloidal suspensions [Moosaiea and Atefia 2007], when the underlying continuum upon which the interacting particles dwell is, say, an ideal incompressible fluid. The choice of model depends on the physical nature of the interactions between the dispersed particles.

Acknowledgment

References

[Cosserat and Cosserat 1909] E. Cosserat and F. Cosserat, *Théorie des corps déformables*, Hermann, Paris, 1909.

[Epstein and de León 1996] M. Epstein and M. de León, "Homogeneity conditions for generalized Cosserat media", *J. Elasticity* **43**:3 (1996), 189–201.

[Epstein and de León 1998] M. Epstein and M. de León, "Geometrical theory of uniform Cosserat media", *J. Geom. Phys.* **26**:1-2 (1998), 127–170.

[Epstein and Elżanowski 2007] M. Epstein and M. Elżanowski, *Material inhomogeneities and their evolution: A geometric approach*, Interaction of Mechanics and Mathematics, Springer, Berlin, 2007.

[Eringen 1999] A. C. Eringen, *Microcontinuum field theories, I: Foundations and solids*, Springer, New York, 1999.

[Moosaiea and Atefia 2007] A. Moosaiea and G. Atefia, "Analysis of concentrated suspension flow by utilizing a Cosserat-type continuum theory", *J. Dispersion Sci. Technol.* **28**:6 (2007), 901–906.

[Zhang et al. 2006] X. Zhang, R. G. Jeffrey, and Y.-W. Mai, "A micromechanics-based Cosserat-type model for dense particulate solids", *Z. Angew. Math. Phys.* **57**:4 (2006), 682–707.

MARCELO EPSTEIN: mepstein@ucalgary.ca
Department of Mechanical and Manufacturing Engineering, University of Calgary, 2500 University Drive NW,
Calgary, AB T2N 1N4, Canada
http://www.ucalgary.ca/hpl/epstein

■msp

DIAGNOSIS OF CONCRETE DAMS BY FLAT-JACK TESTS AND INVERSE ANALYSES BASED ON PROPER ORTHOGONAL DECOMPOSITION

TOMASZ GARBOWSKI, GIULIO MAIER AND GIORGIO NOVATI

Dedicated to the memory of Marie-Louise Steele and to Charles Steele.

Flat-jack tests have been employed for decades for the assessment of stresses and Young moduli in possibly deteriorated concrete dams and masonry structures. We propose a procedure for such tests that includes several innovations: identification of Young moduli and shear modulus in the presence of orthotropy, of pre-existing normal and shear stresses, and of tensile and compressive strength and fracture energy; use of full-field displacement measurements by digital image correlation (instead of extensometers); computer simulations performed once-and-for-all and productive of results which are subsequently processed out by proper orthogonal decomposition and its truncation; and identification of parameters in situ, soon after the tests, by portable computer with software able to perform inverse analyses by mathematical tools newly introduced into this context. The proposed procedure is validated by means of pseudoexperimental numerical exercises, by employing comparatively, as central computational tools, artificial neural networks and a trust region algorithm implying only first-order derivatives (with respect to the sought parameters) of the discrepancy function to minimize.

1. Introduction

Many concrete dams built up several decades ago in developed countries are at present deteriorated, with possible consequent decrease of safety margins with respect to limit states or collapses. A relatively frequent cause of damage in dam concrete is the physicochemical process called alkali-silica reaction: after a dormant period of several years from casting, a substantial decay of mechanical properties and an expansion generating self-stresses may occur slowly along a period of, say, thirty or forty years, with nonuniform distribution over the dam volume and sometimes without external visible manifestations of aging; see, for example, [Swamy 1992; Ahmed et al. 2003; Comi et al. 2009]. Such possible occurrences clearly require a rather large number of diagnostic analyses on a possibly damaged dam.

Structural diagnostic procedures of a mechanical nature employed in state-of-the-art dam engineering can be classified as follows (see [Maier et al. 2004], for instance): (a) quasi-nondestructive experiments by flat jacks on the dam surface [Fedele and Maier 2007]; (b) in depth coring or overcoring tests (similar to those which are traditional in geomechanics), usually with extraction of specimens for the laboratory (and, hence, destructive) [Goodman 1989; Fedele et al. 2005; Leite and Corthesy 2001]; (c) overall dynamical inverse analyses based on excitations by vibrodynes or by ambient vibrations and measurements by

---

accelerometers [Loh and Wu 2000]; (d) statical overall inverse analyses under loading due to ad hoc changes of water level in the reservoir, with measurements of consequent displacements by means of pendulums, collimators and/or interferometric radar [Fedele et al. 2006]; (e) same as (d), but with loading provided by seasonal, in-service variation of reservoir level (such alternative is much more economical than (d), however it requires more complex inverse analyses taking into account thermal effects as well [Ardito et al. 2008]).

Diagnostic procedures (c), (d) and (e) lead to the assessment of Young modulus only. Method (b) is clearly destructive and expensive. The limitations of procedure (a), traditionally applicable only to the estimation of the Young modulus and normal stresses, have been recently mitigated, so far only theoretically, in [Fedele et al. 2005] by recourse to test simulations and inverse analysis.

A step ahead in the development of structural diagnosis based on flat-jack tests and inverse analysis is provided in this paper, where the following novelties are presented and proposed to the engineering practice: (i) less destructive geometric configuration of the slots; (ii) use of digital image correlation (DIC) as full-field displacement measurement technique which economically provides much more experimental data than extensometers; (iii) assessment of inelastic parameters additional to elasticity and of the existing stress state (assumed as uniform, locally); (iv) parameter identification performed in a fast and inexpensive fashion (possibly in situ, using a portable computer) by a numerical procedure resting on proper orthogonal decomposition (POD), radial basis functions (RBF) interpolations and a trust region algorithm (TRA) or by an alternative procedure based on an artificial neural network (ANN).

Section 2 is primarily devoted to the operative sequence of the procedure to be carried out in situ, namely in selected locations on the dam surface. The computational procedures proposed for the estimation of elastic moduli and stresses are described, with some details and numerical validations, in the successive two sections: the techniques centered on POD–RBF–TRA in Section 3, and those based on POD–ANN in Section 4. In Section 5 preliminary research results are presented on the identification of plastic and fracture parameters. Section 6 is primarily devoted to prospects of further research.

2. Experimental procedure and its modeling

2.1. *Constitutive models.* Simple traditional constitutive models, still popular nowadays in dam engineering practice, have been adopted for the present investigation and validation exercises on the proposed parameter identification methods. Let the x-axis be horizontal on the dam free surface, assumed vertical for simplicity, let the z-axis be orthogonal to that surface, and, hence, axis y vertical as well. Anisotropy, specifically orthotropy with transversal isotropy in the horizontal plane, may be generated in dam concrete by the casting process, especially in roller compacted concrete, more and more frequently adopted in the last decade. In this case linear elasticity can be described as follows:

$$\begin{pmatrix} \epsilon_x \\ \epsilon_y \\ \epsilon_z \\ \gamma_{xy} \\ \gamma_{yz} \\ \gamma_{xz} \end{pmatrix} = \begin{pmatrix} 1/E_H & -\nu_{VH}/E_V & -\nu_H/E_H & 0 & 0 & 0 \\ -\nu_{HV}/E_H & 1/E_V & -\nu_{HV}/E_H & 0 & 0 & 0 \\ -\nu_H/E_H & -\nu_{VH}/E_V & 1/E_H & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/G_V & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/G_V & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/G_H \end{pmatrix} \begin{pmatrix} \sigma_x \\ \sigma_y \\ \sigma_z \\ \tau_{xy} \\ \tau_{xy} \\ \tau_{xz} \end{pmatrix}, \qquad (1)$$

$$v_{HV} = v_{VH} \frac{E_H}{E_V}, \qquad G_H = \frac{E_H}{2(1 + v_H)}. \qquad (2)$$

The material parameters in (1) are the horizontal and vertical Young moduli, E_H and E_V; the in-plane (horizontal) and out-of-plane shear moduli, G_H and G_V; and the Poisson ratios v_H, v_{VH} and v_{HV}, constrained by the relationships (2) due to the symmetry of the elastic compliance matrix and to the in-plane elastic energy balance (same as in the isotropic case), respectively. Therefore only five of these parameters are independent. The identification procedures developed in Section 3 and 4 will concern the three parameters E_H, E_V and G_V, which are the ones reasonably expected to play the main role in the system response in the hypothesis of transversal isotropy (in the horizontal plane) for the dam concrete. We assume a priori that $v_H = v_{VH} = 0.2$; the other parameters in (1) depend on the previous ones through (2).

Isotropy reduces to two the independent elastic parameters, say E and G (since $E_V = E_H = E$, $G_V = G_H = G$, $v_H = v_{HV} = v_{VH} = E/(2G) - 1$). The isotropy hypothesis will be used in Section 5 when focus is set on the identification of inelastic parameters.

The stress field pre-existing to the test and sought for by the identification procedure can be reasonably assumed to be a plane stress state ($\sigma_z = \tau_{zx} = \tau_{zy} = 0$), assumed as uniform over the whole volume affected by the test, hence governed by the three components $\sigma_H (= \sigma_x)$, $\sigma_V (= \sigma_y)$ and $\tau_{HV} (= \tau_{xy})$.

For the plastic behavior of concrete the classical Drucker–Prager model is adopted here (perfect plasticity with nonassociated flow rule), depicted in Figure 1a. Fracture is supposed to be first-mode only and reducible to a cohesive crack model with linear softening, as shown in Figure 1b in terms of normal stress σ versus opening displacement v. The assumptions above on the inelastic behavior of concrete lead to the following mathematical model formulation

$$F = t - p \tan \beta - d \leq 0, \qquad d = \left(1 + \tfrac{1}{3} \tan \beta\right) f_t \qquad (3)$$

(see [Lubliner 1990; Jirasek and Bazant 2001], for instance). The meanings of the symbols in these formulas and in the figure are as follows: F is the yield function in the plane (Figure 1a) of the average normal stress p (compression positive) and of the equivalent shear t (I_1 being the first invariant of the stress tensor, J_2 the second invariant of the deviatoric stress tensor); the angles β and ψ represent internal friction and dilatancy of plastic flow, respectively; the cohesion d and the uniaxial strength f_t are related to each other by (3)$_2$, as a consequence of (3)$_1$. The dilatancy ψ is assumed to be $\psi = 40°$, a value frequently adopted in overall analyses of concrete dams in engineering practice. Thus the inelastic
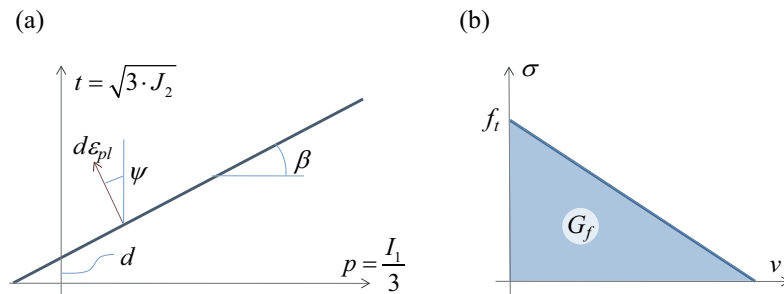

Figure 1. Material models and parameters relevant to diagnostic purposes: (a) Drucker–Prager no-hardening plasticity; (b) cohesive crack model with linear softening.
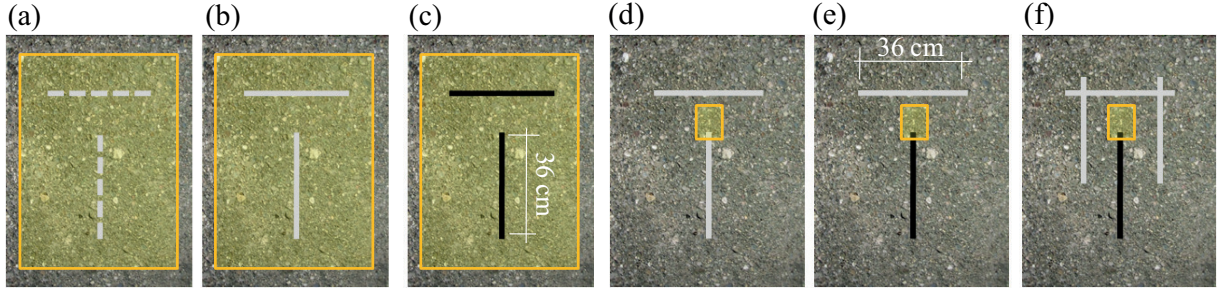
Figure 2. Sequence of steps in the parameter identification procedure: (a) positioning and (b) cutting of the slots (recording of displacements over the large ROI); (c) jacks inserted and pressurized (recording of displacements over the large ROI); (d) depressurization of jacks and change of ROI; (e) vertical jack pressurized to higher pressure level (displacements recorded over the smaller ROI).

parameters to estimate turn out to be three only: the internal friction β, the tensile strength f_t, and the fracture energy G_f. The uniaxial compressive strength f_c can be computed after model calibration through the following relationship, consequence of (3) for $F = 0$:

$$f_c = \frac{d}{1 - \frac{1}{3}\tan\beta} = \frac{\left(1 + \frac{1}{3}\tan\beta\right)f_t}{1 - \frac{1}{3}\tan\beta}. \tag{4}$$

2.2. *The test and its simulation.* The sequence of operative steps of the proposed diagnostic procedure based on flat-jack tests is outlined below with reference to Figure 2.

(a) In the selected place on the structure surface, the position of two future orthogonal slots (T-shaped geometry) is marked and a first photograph is taken by the DIC instrument over the region of interest (ROI) which is depicted in Figure 2a. In the DIC jargon (see [Hild and Roux 2006; Avril et al. 2008], for example), ROI is the area over which full-field measurements are performed; zone of interest (ZOI) is the image subset used by the image matching routine as a correlation window: comparing photos taken before and after the deformation, every ZOI of the initial image can be located in the deformed image through its "signature" in terms of gray-level values characterizing the pixel subset in point, as schematically illustrated in Figure 3. The displacement of the center point of a ZOI is computed as average of the displacements of the pixels inside the subset. A suitable grid is defined over the ROI, grid-nodes displacements being measurable by means of subsequent DIC photos.

(b) The two slots are cut (Figure 2b) and a second photo is shot by the DIC equipment so that, as mentioned, the displacements due to the release of the pre-existing stresses in the cut can be measured at all grid nodes and can be dealt with as accurately representative of the displacement full-field over the ROI.

(c) Two flat-jacks are inserted and pressurized by a piston-pumped liquid up to a pre-established pressure (Figure 2c). This pressure should correspond to a compromise between the needs to generate large enough, well measurable displacements in the ROI and to avoid nonnegligible inelastic strains. Again a DIC photo is taken over the ROI in order to capture the new displacement full-field.

(d) The horizontal jack is removed from its slot and the vertical one is depressurized. The original ROI employed in the preceding steps is replaced by a smaller ROI appropriate for focusing on one of the two
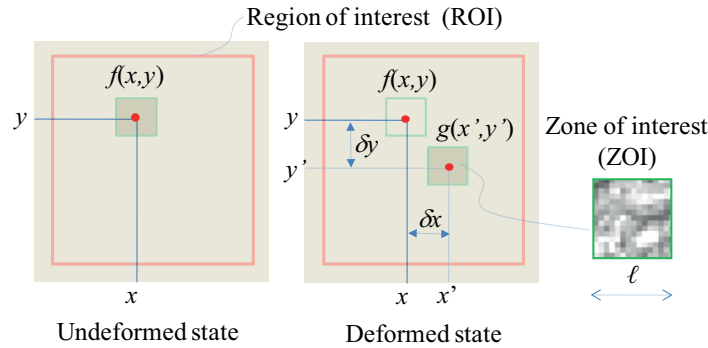
Figure 3. Schematic visualization of the DIC measurement procedure. The functions f and g are gray-level values; x, y are pixel coordinates in the ROI, and δx, δy are the displacement coordinates of the ZOI central point.

zones, located near the tips of the vertical slot, where inelastic deformations are expected (Figure 2d). A reference photo is taken of this ROI.

(e) The single flat-jack placed in the vertical slot is highly pressurized in order to generate plastic strains and a quasi-brittle fracturing process near the tips of the loaded slot where stress concentrations occur (Figure 2e). A sequence of DIC photos is taken to capture the nonlinear evolution of the displacements over the ROI, due to such pressure loading. The crack propagation is assumed as confined to the central vertical plane and to be stable, despite the softening in the cohesive crack behavior. Such localized fracture can be regarded as quasi-nondestructive in engineering terms. A possible provision to be adopted which can be beneficial in phase (e) for inelastic material characterization is the following one: two additional vertical slots (shown in Figure 2f) are cut in order to practically remove the influence on the fracture process of the horizontal stress; this, however, would make the test slightly more destructive. This variant is not investigated herein.

Based on these DIC measurements, the parameters identifications are performed according to the following sequence of phases (see caption of Figure 2): (I) elastic moduli, on the basis of experimental data concerning transition from stage (b) to (c); (II) stresses, on the basis of the elastic moduli estimates achieved in phase (I) and of data acquired at stages (a) and (b); (III) inelastic parameters, on the basis of data concerning the transition from stage (d) to the various deformation stages represented in the sequence of DIC photos taken in phase (e).

The finite element (FE) model here adopted for the subsequent computer simulations of the tests is depicted in Figure 4a as for the overall geometry and in Figure 4b as for the mesh. The FE discretization exhibits the following features: 97,600 tetrahedral elements with linear shape functions for displacements; 57,280 degrees of freedom; boundary conditions with vanishing displacements on the borders separating the domain from the surrounding volume supposed to be not perturbed by the test. It is worth noting that the ROI for DIC measurements is significantly reduced in moving from the identification of elastic moduli and stresses to the estimation of inelastic parameters: in fact for the latter inverse analyses concerning less uniform and more localized strains field, it is useful to exploit a higher density of the available DIC pixels.
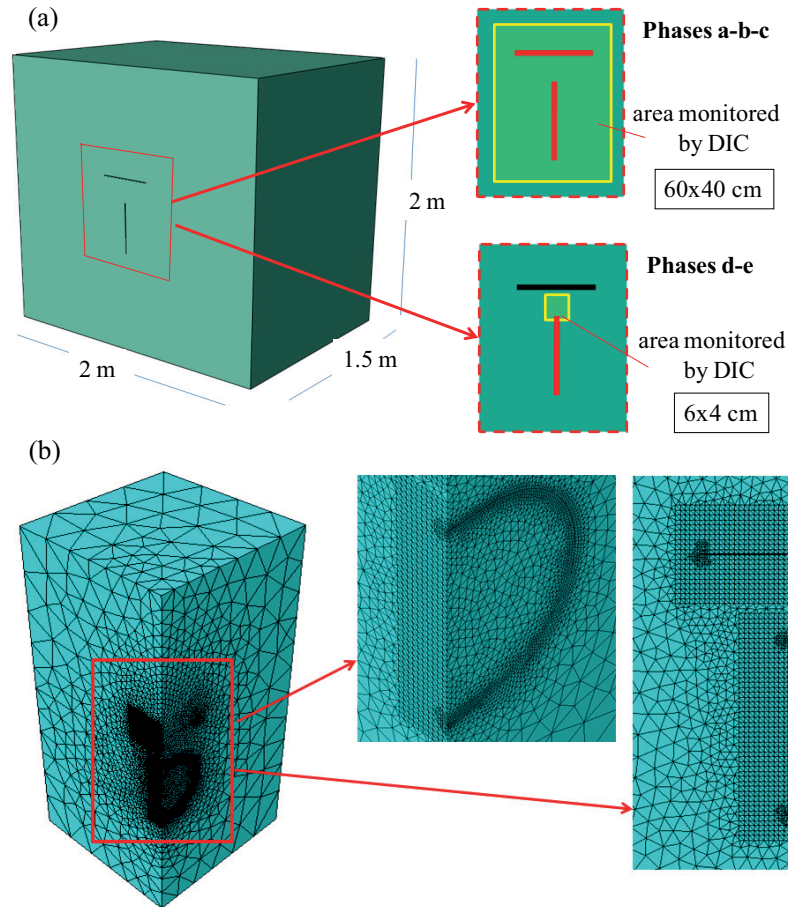
Figure 4. Finite element model for test simulations: (a) overall geometry and (b) mesh.

The simplicity of both the constitutive models and of the FE discretization for the numerical exercises presented in the two subsequent sections is motivated by the following circumstances: the present purpose is limited to a first numerical validation of novel procedures; consistency is pursued here with the assumptions and approximations at present adopted for real-life structural analyses of large concrete dams in view of inevitable uncertainties of several data.

A further simplification of test modeling concerns the flat-jacks: their configurations adopted at present would suggest to model the two steel laminae and the circumferential welds. Here uniform compressive traction is assumed throughout the concrete walls of the slot. Such assumption is reasonable: first, because it represents an acceptable approximation to the present purposes; second, because a new kind of flat jack based on textile laminae supported by a circumferential rod, now under investigation, is likely to be adopted in the future, thus making the assumption above more realistic.

For the numerical validation of the novel diagnostic procedures the adopted jack dimensions and test geometry are similar to those which are at present usual in traditional applications, consistent with usual sizes of aggregate in concrete. Clearly, larger dimensions with equal procedures can be adopted for dam concrete when the average aggregate size turns out to be larger.

3. Identification of elastic moduli and stresses by a POD–RBF–TRA procedure

3.1. *The inverse analysis procedure.* As usual in diagnostic inverse analyses, the domain in space over which the search for the material parameters can be confined should be preliminarily chosen according to expert judgment. For the present purposes, at first with reference to the estimation phase (I), the "feasible domain" is defined within the three-dimensional space of the elastic parameters in Table 1 by means of lower and upper bounds. Adopting the step values specified in this table, a grid with $M = 1089$ nodes is generated over the feasible domain.

For each of these sets of parameters the test simulation provides, by the FE model of Figure 4, the displacements measurable by DIC (over the ROI shown in Figure 2) relevant to the operative steps from (a) to (c), namely the displacements due to slot cutting and jack pressurization up to $p = 5$ MPa in the elastic range. The relative large size of the ROI makes immaterial, for identification by elastic modeling, the possible initial development of inelastic strains, localized near the slot tips, as the jack pressure approaches the maximum value ($p = 5$ MPa).

The DIC instruments here considered measure (and provide in digitalized form) $N = 2086$ displacement components (along horizontal and vertical axis on the dam surface, Figure 2a) at the nodes of the FE model (Figure 4) which coincide with the DIC grid nodes contained in the ROI (Figure 3).

The POD–RBF procedure. The clearly massive and heavy sequence of direct analyses based on the parameter vectors $\boldsymbol{p}_i$ ($i = 1, \ldots, M$) corresponding to grid nodes has been carried out by a commercial FE code (ABAQUS). Let the results be gathered in an $N \times M$ matrix $\boldsymbol{U}$: its i-th column $\boldsymbol{u}_i$ is a vector, or *snapshot*, containing the N pseudoexperimental data belonging to the (now elastic) response to the pressure loading of the system with elastic moduli gathered in vector $\boldsymbol{p}_i$.

Clearly, since the differences among the system responses arise only from variations of the sought parameters within the preselected domain in their space, the snapshots turn out to be correlated: they can be represented by "almost parallel" vectors in their space of $N = 2086$ dimensions. This correlation suggests the use of approximation, or *compression*, in the information contained in the snapshot matrix $\boldsymbol{U} = [\boldsymbol{u}_1 \cdots \boldsymbol{u}_M]$ relevant to the elastic behavior of the system, by recourse to a proper orthogonal decomposition (POD) and truncation based on it.

The mathematical theory and computational procedures related to POD have origins remote in time and in application fields, and are now presented in a vast and still growing literature, of which we mention [Chatterjee 2000; Wu et al. 2003; Ostrowski et al. 2008]. Here only the specific procedure selected and applied to the present purposes is outlined, without analytical details.

Starting from the matrix $\boldsymbol{U}$ defined above, preliminarily obtained from M test simulations by FEM, the symmetric, positive semidefinite or definite matrix $\boldsymbol{D} = \boldsymbol{U}^T \boldsymbol{U}$ is generated. Its (real, nonnegative)

	phase (I)			phase (II)					
	E_{H}/GPa	E_{V}/GPa	G_{V}/GPa	E_{H}/GPa	E_{V}/GPa	G_{V}/GPa	σ_{H}/MPa	σ_{V}/MPa	τ_{HV}/MPa
min	10	15	4	10	15	4	−6	−7	−2.5
step	1	1	1	5	5	4	1	1	0.5
max	20	25	12	20	25	12	−2	−1	−0.5

Table 1. Grid in parameter space for phases (I) and (II) of the identification procedure.

eigenvalues λ_i and the relevant normalized eigenvectors $\boldsymbol{v}_i$ are computed and employed to obtain the orthonormal basis

$$\boldsymbol{\Phi} = \begin{bmatrix} \boldsymbol{\Phi}_1 & \cdots & \boldsymbol{\Phi}_M \end{bmatrix}, \qquad \text{with} \qquad \boldsymbol{\Phi}_i = \boldsymbol{U}\boldsymbol{v}_i\lambda_i^{-1/2}. \tag{5}$$

In view of orthonormality, the $M \times M$ matrix $\boldsymbol{A}$ encoding the amplitudes $\boldsymbol{a}$ of the snapshots $\boldsymbol{u}$ is characterized by the relations

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_M \end{bmatrix} = \boldsymbol{\Phi}^T\boldsymbol{U}, \qquad \boldsymbol{U} = \boldsymbol{\Phi}\boldsymbol{A}. \tag{6}$$

The correlation among the snapshots (earlier noticed and reasonably expected for measurable responses to tests at varying parameters over the domain) make negligible many amplitudes $\boldsymbol{a}_i$ in the new basis $\boldsymbol{\Phi}$. It was mathematically proved (see e.g., [Nabney 2002; Liang et al. 2002]) that the eigenvalues λ_i quantify the negligibility of such amplitudes. By preserving only the $\overline{M}$ largest eigenvalues, with $\overline{M} \ll M$, an approximation $\overline{\boldsymbol{U}}$ is achieved of the $N \times M$ snapshot matrix $\boldsymbol{U}$ by means of a truncated basis $\overline{\boldsymbol{\Phi}}$ ($N \times \overline{M}$) and the relevant truncated $\overline{M} \times M$ amplitudes $\overline{\boldsymbol{A}}$, namely

$$\overline{\boldsymbol{U}} = \overline{\boldsymbol{\Phi}}\,\overline{\boldsymbol{A}} \cong \boldsymbol{U}, \qquad \text{whence} \qquad \overline{\boldsymbol{A}} = \overline{\boldsymbol{\Phi}}^T\boldsymbol{U}. \tag{7}$$

The approximation in $(7)_1$ is regarded as acceptable and negligible in subsequent developments, so that the amplitudes in $\overline{\boldsymbol{A}}$ are henceforth computed by $(7)_2$.

These developments often turn out to be computationally heavy but need to be done once-and-for-all as preparatory work generating matrices $\overline{\boldsymbol{\Phi}}$ and $\overline{\boldsymbol{A}}$: these condense ("compress") the information on the system behavior contained in the much larger matrix $\boldsymbol{U}$ and can be accommodated in a small computer to be employed "in situ".

The snapshot $\boldsymbol{u}$ corresponding to any "new" parameter vector $\boldsymbol{p}$ (new because not included as node in the preselected grid over the feasible domain in the parameter space) now can be assessed through (7) by means of its "amplitude" $\overline{M}$-vector $\bar{\boldsymbol{a}}(\boldsymbol{p})$. This amplitude can be expressed by means of radial basis function (RBF) interpolation (see, e.g., [Wu 1995; Buhmann 2003; Hassing et al. 2010]) using the known amplitude vectors $\bar{\boldsymbol{a}}_i(\boldsymbol{p}_i)$ corresponding to grid-nodes vectors $\boldsymbol{p}_i$, $i = 1, \ldots, M$:

$$\bar{a}_k(\boldsymbol{p}) = \sum_{i=1}^{M} b_i^k \, g(\boldsymbol{p}, \boldsymbol{p}_i) = \boldsymbol{g}^T(\boldsymbol{p})\boldsymbol{b}^k \qquad k = 1, \ldots, \overline{M}, \tag{8}$$

where $\bar{a}_k(\boldsymbol{p})$ is the k-th component of $\bar{\boldsymbol{a}}(\boldsymbol{p})$, the b_i^k are the interpolation coefficients and $g(\boldsymbol{p}, \boldsymbol{p}_i)$ are the radial functions; the vectors $\boldsymbol{g}^T(\boldsymbol{p})$ and $\boldsymbol{b}^k$ gather functions and coefficients, respectively. The cubic RBFs

$$g(\boldsymbol{p}, \boldsymbol{p}^i) = \|\boldsymbol{p} - \boldsymbol{p}_i\|^3 \tag{9}$$

where $\| \cdot \|$ is the Euclidean norm, are adopted here for the interpolation above. For each component k, Equation (8), enforced at the j-th node ($j = 1, \ldots, M$) of the parameter grid, takes the form $[\bar{\boldsymbol{a}}_j(\boldsymbol{p}_j)]_k = \boldsymbol{g}^T(\boldsymbol{p}_j)\boldsymbol{b}^k$. Assembling all M equations of this type, the following linear system is obtained, with the coefficients b_i^k as unknowns:

$$\boldsymbol{G}\boldsymbol{b}^k = [\,k\text{-th row of matrix } \overline{\boldsymbol{A}}\,]^T, \qquad k = 1, \ldots, \overline{M}, \tag{10}$$

where $\boldsymbol{G} = (G_{ji})$ with $G_{ji} = g(\boldsymbol{p}_j, \boldsymbol{p}_i)$. All the $\overline{M}$ linear equation systems to be solved share the same coefficient matrix (which is independent of k).

When matrix $\boldsymbol{B} = [\boldsymbol{b}^1 \ldots \boldsymbol{b}^k \ldots \boldsymbol{b}^{\overline{M}}]$ is available, the interpolation, (8), provides the truncated amplitude vector $\bar{\boldsymbol{a}}$ corresponding to any new parameter vector $\boldsymbol{p}$: $\bar{a}_k(\boldsymbol{p}) = [\,\boldsymbol{g}(\boldsymbol{p})]^T \boldsymbol{b}^k$; and, hence, (7) leads to the snapshot $\boldsymbol{u}$ of the N measurable quantities, namely to an approximation consistent with the preceding POD process. Formally:

$$\boldsymbol{u}(\boldsymbol{p}) \cong \overline{\boldsymbol{\Phi}} \; \bar{\boldsymbol{a}}(\boldsymbol{p}) = \overline{\boldsymbol{\Phi}} \; \boldsymbol{B}^T \boldsymbol{g}(\boldsymbol{p}) \tag{11}$$

where the $M-$vector $\boldsymbol{g}$ contains the values of the considered RBFs at the parameter node $\boldsymbol{p}$.

Clearly, (11) may replace with substantial savings test simulations (direct analyses) by FE computer codes or by other computing tools. This circumstance becomes very meaningful and practically advantageous in inverse analysis of the present kind.

In the present context parameter identification means minimization of a traditional discrepancy function ω of the residuals $\boldsymbol{R}$ (differences between experimental data $\boldsymbol{u}_s$ and their computed counterparts $\boldsymbol{u}_c$). Traditionally, the function ω is formulated as quadratic form with the inverse of the covariance matrix $\boldsymbol{C}$ of the experimental measures:

$$\min_{\boldsymbol{p}} \; \omega\,(\boldsymbol{p}) = \min_{\boldsymbol{p}} \left\{ \boldsymbol{R}^T\,(\boldsymbol{p})\,\boldsymbol{C}^{-1}\boldsymbol{R}\,(\boldsymbol{p}) \right\} \tag{12}$$

where

$$\boldsymbol{R}\,(\boldsymbol{p}) = \boldsymbol{u}_s - \boldsymbol{u}_c\,(\boldsymbol{p})\,. \tag{13}$$

In view of the "pseudoexperimental" approach of the present investigation, it is assumed herein $\boldsymbol{C} = \boldsymbol{I}$ (identity matrix) without methodological limitations.

In order to numerically solve the problem formulated by (12), the popular trust region algorithm (TRA) is employed in this section and it is outlined below (for details see e.g., [Coleman and Li 1996; Conn et al. 2000]).

The Jacobian $\boldsymbol{J}$, the gradient $\partial\omega/\partial\boldsymbol{p}$ of the objective function ω and an approximation of the Hessian $\boldsymbol{H}$ read, respectively:

$$\boldsymbol{J}\,(\boldsymbol{p}) = \frac{\partial \boldsymbol{R}}{\partial \boldsymbol{p}}\,, \qquad \frac{\partial \omega}{\partial \boldsymbol{p}} = \boldsymbol{J}^T \boldsymbol{R}\,, \qquad \boldsymbol{H}\,(\boldsymbol{p}) = \frac{\partial^2 \omega}{\partial \boldsymbol{p}^T \partial \boldsymbol{p}} \simeq \boldsymbol{J}^T \boldsymbol{J}\,. \tag{14}$$

In each iteration, say the j-th one starting from the current point $\boldsymbol{p}_j$, the following two-variable mathematical programming problem has to be solved:

$$\min_{\alpha, \beta} \left\{ \tfrac{1}{2} \boldsymbol{s}_j^T \boldsymbol{J}^T \boldsymbol{J} \boldsymbol{s}_j + \left(\frac{\partial \omega}{\partial \boldsymbol{p}} \right)^T \boldsymbol{s}_j, \quad \|\boldsymbol{s}_j\| \leq \Delta_j \right\}, \tag{15}$$

where

$$\boldsymbol{s}_j = \alpha \boldsymbol{J}_j^T \boldsymbol{R}_j - \beta (\boldsymbol{J}_j^T \boldsymbol{J}_j)^{-1} \frac{\partial \omega}{\partial \boldsymbol{p}}. \tag{16}$$

In (15) Δ_j defines the "trust region" to be possibly adjusted after check according to criteria specified e.g., in [Nocedal and Wright 1999; Conn et al. 2000].

The iteration j of the first-order TRA described by (15) and (16) requires a multiplicity of first derivatives specified in (14) and numerically approximated here by forward finite differences. Therefore a high number of direct analyses are necessary in each application of TRA to the present inverse problem.

As for TRA and similar sequential mathematical programming technique, the computational burden is increased by possible nonconvexity (namely by possible local minima) of the objective function

$\omega(\boldsymbol{p})$ which may require different initializations (even if once-and-for-all for a specific kind of technical situations). In parameter identification problems where lack of discrepancy function convexity is expected, genetic algorithms may be an advantageous alternative (not examined herein) to mathematical programming, and the preliminary POD–RBF procedure above becomes very desirable for practical applications in view of the high number of required direct analyses. Computational savings may arise from POD–RBF also for sensitivity analysis of many measurable quantities (see [Kleiber et al. 1997], for example); however, we will not consider this topic further.

In the present diagnostic method the inevitably sequential estimation of parameters (elastic-stresses-inelastic, Section 2, Figure 2) implies an increase of the parameter space dimensionality and hence a growth (exponential at equal density) of the number M of grid nodes (and snapshots).

The advantages achievable by the procedure POD–RBF–TRA proposed herein in association with DIC and flat-jack tests for the identification, in sequence, of elastic moduli and stresses, are evidenced by numerical examples in Section 3.2.

In the POD–RBF procedure just outlined, truncation represents a clearly important step. The choice of number $\overline{M}$ of the preserved eigenvalues of matrix $\boldsymbol{D}$ (much below the number M of snapshots) can be decided by various criteria (besides the one based on eigenvalues λ_i, here without a direct physical meaning), two of which are particularly suitable to the present context: (i) meaningful difference between a snapshot resulting from a test simulation by FE analysis and the corresponding snapshot computed by POD–RBF with truncations; (ii) assessment of intrinsic accuracy of POD–RBF at varying the dimensionality $\overline{M}$ of the new basis $\overline{\boldsymbol{\Phi}}$. The latter has been adopted in the present study and is outlined in what follows.

Outside the nodes grid but still within the preselected domain, $\hat{M}$ parameter vectors $\hat{\boldsymbol{p}}_j$ are randomly generated (here $\hat{M} = 100$), and corresponding snapshots of pseudoexperimental data are computed through FE simulations of the same test. Random noise (here $\pm 5\,\mu$m) is added to each snapshot which thus becomes $\hat{\boldsymbol{u}}_j$ ($j = 1, \ldots, \hat{M}$). The amplitude vector corresponding to $\hat{\boldsymbol{u}}_j$ is computed using (7):

$$\bar{\boldsymbol{a}}_j(\hat{\boldsymbol{p}}_j) = \overline{\boldsymbol{\Phi}}^T \hat{\boldsymbol{u}}_j(\hat{\boldsymbol{p}}_j), \qquad j = 1, \ldots, \hat{M}. \tag{17}$$

Now an approximation to such amplitude vector has to be attainable by RBF interpolation through (8) (employing the matrix $\boldsymbol{B}$) in correspondence with some $\boldsymbol{p}$:

$$\bar{\boldsymbol{a}}^{\mathrm{RBF}}(\boldsymbol{p}) = \boldsymbol{B}\,\boldsymbol{g}(\boldsymbol{p}). \tag{18}$$

Such $\boldsymbol{p}$ can be singled out by defining (for each j-th off-grid node) a residual

$$\boldsymbol{R}_j(\boldsymbol{p}) = \bar{\boldsymbol{a}}_j(\hat{\boldsymbol{p}}_j) - \bar{\boldsymbol{a}}^{\mathrm{RBF}}(\boldsymbol{p}) \tag{19}$$

and then by solving the problem

$$\boldsymbol{p}_j^* = \text{value of } \boldsymbol{p} \text{ that minimizes } \boldsymbol{R}_j^T(\boldsymbol{p})\,\boldsymbol{R}_j(\boldsymbol{p}). \tag{20}$$

The j-th parameter vector $\hat{\boldsymbol{p}}_j$ originally considered is now compared to its counterpart $\boldsymbol{p}_j^*$ resulting from minimization (20). Their difference can be regarded as representative of the RBF approximation in this j-th numerical test. An overall quantification of such error for the whole set of $\hat{M}$ off-grid nodes, i.e.,
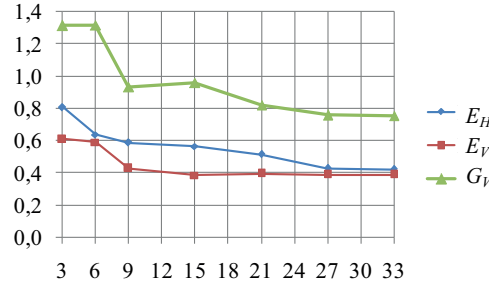
Figure 5. Errors, according to (21), in the elastic moduli due to the POD–RBF preliminaries as a function of the adopted truncation index $\bar{M}$.

the mean error according to (21) below, can be regarded as an assessment of the approximation induced by the POD–RBF scheme:

$$(\text{mean error})_l = \frac{1}{\hat{M}} \sum_{j=1}^{\hat{M}} \left| \frac{(\hat{\boldsymbol{p}}_j)_l - (\boldsymbol{p}_j^*)_l}{(\hat{\boldsymbol{p}}_j)_l} \right| \tag{21}$$

where subscript l denotes the l-th component of the parameters vectors. Figure 5 shows the mean error resulting for the elastic constants as a function of POD truncation index $\bar{M}$.

3.2. *Validation by computational examples.*
This subsection gathers a sequence of selected numerical checks of potentialities and limitations of the (deterministic, least squares) diagnostic method presented in what precedes.

(A) First, inverse analyses have been carried out comparatively by TRA only, with the following provisions: (i) various initializations in order to check the well-posedness of the problem and, particularly, absence of local minima; (ii) random perturbations of the (pseudo) experimental input data in order to check their consequences on the estimates; (iii) same estimation procedure applied comparatively on the basis of DIC data and on the basis of extensometric data which are traditionally acquired with standard flat-jack tests.

The reference values of the parameters used in identification examples, carried out by TRA only, concerning phase (I) and phase (II), are as follows:

elastic parameters: $E_\mathrm{H} = 15\,\mathrm{GPa}$ $E_\mathrm{V} = 20\,\mathrm{GPa}$ $G_\mathrm{V} = 8\,\mathrm{GPa}$ $\nu_\mathrm{H} = \nu_\mathrm{VH} = 0.2$

pre-existing stresses: $\sigma_\mathrm{H} = -3\,\mathrm{MPa}$ $\sigma_\mathrm{V} = -5\,\mathrm{MPa}$ $\tau_\mathrm{HV} = -1\,\mathrm{MPa}$

All three kinds of evaluations led to positive and encouraging results starting from these reference values, with pseudomeasurements by DIC and by the FE model (Figure 4) same as those specified in Section 2; pressure of 5 MPa in both jacks; random noises on input data over intervals of $\pm 2.5\,\mu\mathrm{m}$ and then of $\pm 5.0\,\mu\mathrm{m}$, with uniform probability distribution.

Figure 6 indicates satisfactory stability of the TRA identification with respect to realistic random errors of DIC pseudoexperimental data considered. For comparison purposes, Figure 7a shows possible measurements by extensometers (specifically, linear variable differential transducers) employed at present in dam engineering. The elongation of $8 + 8$ segments between marks are intended to be monitored
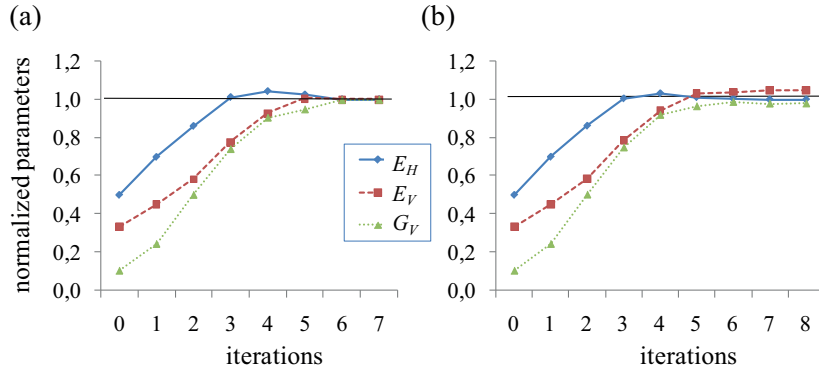
Figure 6. Identification of elastic parameters by TRA with different random perturbations of input data: (a) $\pm 2.5\,\mu$m; (b) $\pm 5.0\,\mu$m.
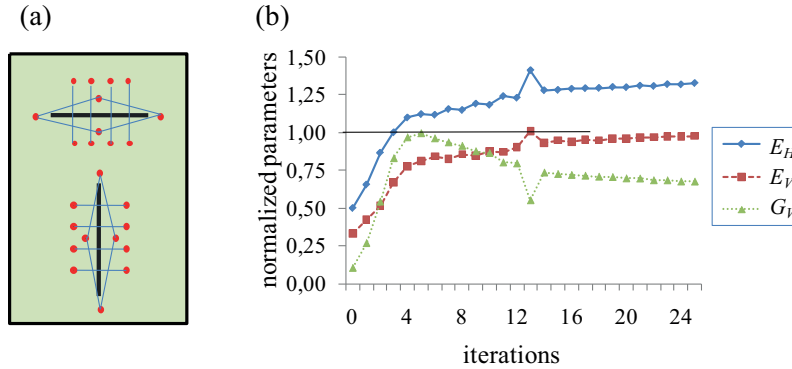

Figure 7. (a) Extensometric measurements and (b) consequent identification of elastic parameters (normalized by the reference values) by TRA inverse analysis.

and affected by $\pm 5.0\,\mu$m random errors. Actually the considered noise and measurement number are rather optimistic when compared to the current practice based on extensometers. A TRA inverse analysis similar to the one leading to Figure 6 has been performed, but now on the basis of snapshots consisting of the 16 extensometric measurements shown in Figure 7a. Lack of convergence turns out to occur with extensometers in the comparative example illustrated by Figures 6b and 7b.

(B) The potentiality and accuracy of the whole procedure POD–RBF–TRA have been tested by numerical exercises like those summarized here below.

With reference to phase (I) of the identification procedure, Table 1 specifies the uniform grid chosen over the three-dimensional domain of the elastic parameters: the number of nodes amounts to $M =$ 1089. Using this grid, the snapshots $\boldsymbol{u}$ (over the ROI) are computed by FE simulation as response to jack pressurization at 5 MPa of the two slots, and the corresponding matrices $\bar{\boldsymbol{\Phi}}$ and $\boldsymbol{B}$ are generated according to the POD–RBF procedure described in Section 3 (see page 187). Here the truncation adopted is quantified by $\bar{M} = 3$. On this basis, for any parameter vector $\boldsymbol{p}$ in the feasible domain, the corresponding snapshot $\boldsymbol{u}$ can be obtained in a fast way by (11) once the vector $\boldsymbol{g}(\boldsymbol{p})$ is computed.

Subsequently a set of $\hat{M} = 100$ off-grid parameter vectors $\hat{\boldsymbol{p}}_j$ is randomly generated and the corresponding snapshots $\hat{\boldsymbol{u}}_j$ (over the ROI) are computed through FE simulations analogous to the previous ones. Such snapshots, after perturbation by random noise addition, are used as pseudoexperimental data relevant to the transition from step (b) to step (c) of the test, Figure 2. For each $\hat{\boldsymbol{p}}_j$, the minimization (12) is carried out, with $\hat{\boldsymbol{u}}_j$ (now noised) playing the role of $\boldsymbol{u}_s$, vector $\boldsymbol{u}_c(\boldsymbol{p})$ being expressible via (11) on the basis of the chosen grid. Here the adopted noise value is $\pm 5\,\mu\mathrm{m}$.

The j-th parameter vector $\hat{\boldsymbol{p}}_j$ originally considered is now compared to its counterpart $\boldsymbol{p}_j^{**}$ identified through the minimization (12) and the accuracy of the estimation of each elastic parameter is assessed. For the j-th identification, the error of the computed estimate $(\boldsymbol{p}_j^{**})_l$ takes the form

$$(\text{error})_{jl} = \left| \frac{(\hat{\boldsymbol{p}}_j)_l - (\boldsymbol{p}_j^{**})_l}{(\hat{\boldsymbol{p}}_j)_l} \right|, \qquad j = 1, \ldots, 100, \qquad l = 1, \ldots, 3. \tag{22}$$

Figure 8a shows the results obtained for the three elastic moduli by means of the 100 inverse analyses discussed above. In each histogram, the abscissa gives a range for the error of the estimate (in percent), and the ordinate gives the percentage of cases in which the error lies in the corresponding interval. The mean error, for the set of $\hat{M}$ identifications, amounts to 1.02%, 0.99% and 1.87% for the moduli E_{H}, E_{V} and G_{V}, respectively.

We next discuss phase (II) of the diagnostic procedure, the estimation of the pre-existing stresses. The preliminary POD–RBF procedure must concern not only the three stress components, but also the three elastic moduli. The six-dimensional grid is specified in Table 1 and has $M = 3375$ nodes. To avoid too high a value of M and consequent high computational burden for the preliminary POD–RBF, the steps adopted for E_{H}, E_{V} and G_{V} in phase (II) are much larger than those in phase (I), as shown in the table. As for the truncation leading to the new basis $\overline{\boldsymbol{\Phi}}$, (7), the present choice is $\overline{M} = 4$.

Similarly to what was described above for phase (I), $\hat{M} = 100$ off-grid parameter vectors are considered and for each of them pseudoexperimental results are generated and utilized to perform an identification by minimizing function ω of (12). It is worth noting that such minimization is to be performed not in the six-dimensional parameter domain over which the grid is defined, but in its three-dimensional subdomain of the stress parameters.
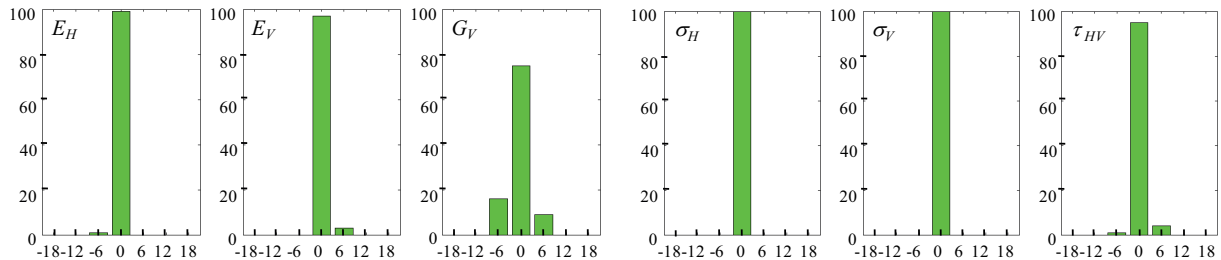

Figure 8. Performance of POD–RBF–TRA procedure for parameter identification based on pseudoexperimental DIC measurements affected by random noise of $\pm 5\,\mu\mathrm{m}$. In the abscissa, intervals of relative errors; in the ordinate, the percentage of errors found to fall within each such interval. Estimates of elastic moduli (first three diagrams) and stresses (remaining diagrams).

Figure 8b provides a visualization of the accuracy checks now concerning the POD–RBF–TRA estimation of the stresses σ_H σ_V τ_{HV} by means of $N = 2086$ measurements (over the ROI) of displacements, in transition from step (a) to step (b) (see Figure 2), after having introduced as preliminary data the estimates achieved for the elastic moduli by the preceding identification. The resulting mean error equals to: 0.65%, 0.96% and 1.51% for σ_H, σ_V and τ_{HV}, respectively.

The conclusion of the computations described (and other omitted for brevity) is rather encouraging: the devised method centered on POD, RBF and TRA, confined to elastic behavior, turns out to provide parameter estimations of both elastic moduli and pre-existing stresses which are expected to be fairly satisfactory in practical applications.

4. Identification of elastic moduli and stresses by a POD–ANN procedure

4.1. *The inverse analysis procedure.* Soft-computing methods, particularly artificial neural networks (ANN), are more and more frequently employed in engineering also to solve economically parameter identification problems (see e.g., [Waszczyszyn 1999; Fedele et al. 2005]). To the present purposes a conventional feed-forward neural network is adopted and optimized, in order to comparatively investigate an alternative to the POD–RBF–TRA procedure presented in what precedes. To achieve computational economies similar to those described in Section 3, POD is adopted again, and the compression of the information contained in the set of preliminary test simulations turns out to be useful for the reduction of the ANN input and for its useful "balance" with corresponding output.

The ANN is intended to provide the sought parameters $\boldsymbol{p}$ starting from the $\overline{M}$-dimensional amplitude vector $\bar{\boldsymbol{a}}$ which corresponds to the N-dimensional experimental data vector $\boldsymbol{u}$ (snapshot) in the truncated basis $\overline{\boldsymbol{\Phi}}$ with $\overline{M} \ll M$, if M is the number of snapshots preliminarily computed by a FE code, like in the POD–RBF–TRA of Section 3. Since orthonormality is preserved in $\overline{\boldsymbol{\Phi}}$ after truncation, (7) gives rise to the relationship

$$\overline{A} = \overline{\boldsymbol{\Phi}}^T U. \tag{23}$$

This equation can be specialized to the link between any snapshot $\boldsymbol{u}_h$ (corresponding, through experiment or its simulation, to a parameter vector $\boldsymbol{p}_h$) and its compacted amplitude $\bar{\boldsymbol{a}}_h$, namely

$$\bar{\boldsymbol{a}}_h\left(\boldsymbol{p}_h\right) = \overline{\boldsymbol{\Phi}}^T \boldsymbol{u}_h\left(\boldsymbol{p}_h\right) \qquad \text{where} \quad h = 1, \dots, M. \tag{24}$$

The $N \times \overline{M}$ matrix $\overline{\boldsymbol{\Phi}}$ (where N is the number of experimental or pseudoexperimental data and $\overline{M}$ the number of preserved eigenvalues of the matrix $\boldsymbol{D} = \boldsymbol{U}^T\boldsymbol{U}$) is computed once and for all by the same POD procedure presented in Section 3. Equation (24) is essential for the generation of an ANN able to routinely estimate, fast and possibly in situ, the parameters $\boldsymbol{p}$ on the basis of input $\bar{\boldsymbol{a}}$ representative of measurements $\boldsymbol{u}$ through (24) again.

The kind of ANN adopted here is a traditional one and can be outlined as follows (details in [Bishop 1995; Hagan et al. 1996; Bishop 2006])

The chosen ANN category is usually called *feed-forward multilayer perceptron.* Its architecture consists of a sequence of neurons layers (input, hidden, output). The transformation performed by each neuron in each active layers (not in the input layer) on received signals consists of a linear combination (by coefficients to be calibrated, called *weights* and *biases*, all gathered in a vector $\boldsymbol{w}$) and a sigmoid

transform. The vector w is computed by minimizing a function Ω that quantifies the discrepancy between expected parameters p_h (targets) related to the amplitude $\bar{a}_h$ through (24), and their counterparts generated by the ANN to be trained and, hence, related to the vector w of the minimization variables. Namely, the problem to solve for the ANN calibration is

$$\min_{w} \Omega(w) = \min_{w}\left\{ \sum_{h=1}^{M} \left\| p_h(\bar{a}_h) - p_h'(w, \bar{a}_h) \right\| \right\}. \tag{25}$$

The preselected parameter nodes p in the grid over the domain chosen in the p-space, and the corresponding, preliminarily computed, snapshots u together with the associated amplitudes $\bar{a}$ after POD, provide, without additional computations, a set of M *patterns* (pairs of $\bar{a}_h$ and corresponding targets p_h). These patterns are employed partly for ANN training as above (see (25)), and also partly for testing and validating the ANN. These computational processes are carried out once-and-for-all in order to optimize the architecture of the ANN for routine later use. ANN optimization is here intended, as usual, to avoid overfitting and roughness in the parameter estimations.

The solution w^* of problem (25) is achieved by the popular Levenberg–Marquardt algorithm, i.e., by first order mathematical programming in its back-propagation version, often employed for training feed-forward ANNs; see [Haykin 1998; Waszczyszyn 1999; Bishop 2006], for example.

4.2. *Validation by numerical examples.*
The chosen domain in the parameter space and grid are again those specified in Table 1. Therefore the $N \times M$ matrices U containing DIC pseudoexperimental data preliminarily computed once-and-for-all by FE simulations (Figure 4) have dimensions $N = 2086$, $M = 1089$ and $N = 2086$, $M = 3375$ for the identification of elastic moduli and of stresses, respectively.

Also the two POD procedures are the same as in Section 3 but now they are intended to compress the information contained in the snapshots in order to provide amplitudes in a new truncated basis $\overline{\Phi}$ as input for a suitably trained ANN. The set of snapshots is subdivided into three subsets for various roles in ANN calibration: 70% for training; 15% for testing; 15% for validation. All pseudoexperimental data in U are modified by random errors in the range $\pm 5\,\mu$m.

The number of amplitude components and the number of neurons in the ANN hidden layer (after a priori choice of an ANN with a single hidden layer architecture) are here established by the following optimization procedure: the mean error is assessed as a function of the neuron number in the hidden layer H and in the input layer I. The latter neuron number coincides with the dimensionality $\overline{M}$ of the amplitude vector generated by the truncated POD.

Figure 9 shows the influence of H with $I = 3$ and $I = 6$ on the ANN devoted to estimation of elastic parameters. The mean errors in Figure 9 are the averages in percentage of the differences between the expected exact parameters (called *targets*) and their values resulting from ANN training and testing. As a conclusion, the process of generating and checking the computational tool POD-ANN for later routine estimation of the three elastic parameters, leads to the choice of $I = 3$ neurons in input layer and $H = 6$ in hidden layer.

As for the identification of stresses σ_{H}, σ_{V} and τ_{HV} by means of the DIC measurements in phase (b) of the flat-jack test (Figure 2), the input layer of the second ANN must include three additional neurons to receive the estimates of elastic parameters based on DIC data from phase (c), besides the neurons that receive the amplitude $\overline{M}$-vector generated by the truncated POD, (23)+(24), of the snapshot u provided
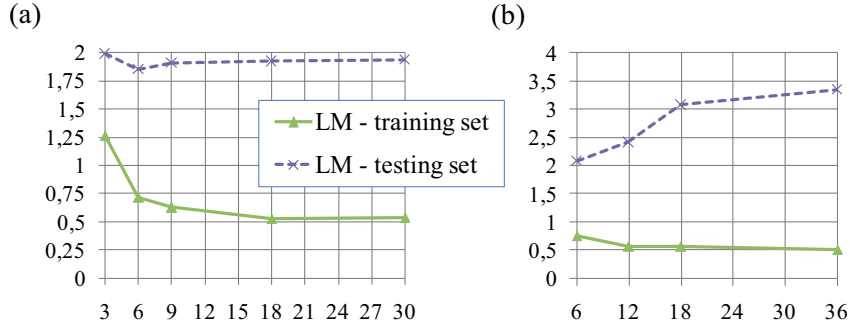
Figure 9. Influence of H, the number of neurons in the hidden layer, on the POD-ANN procedure for estimating the elastic parameters for (a) three and (b) six neurons in the input layer.
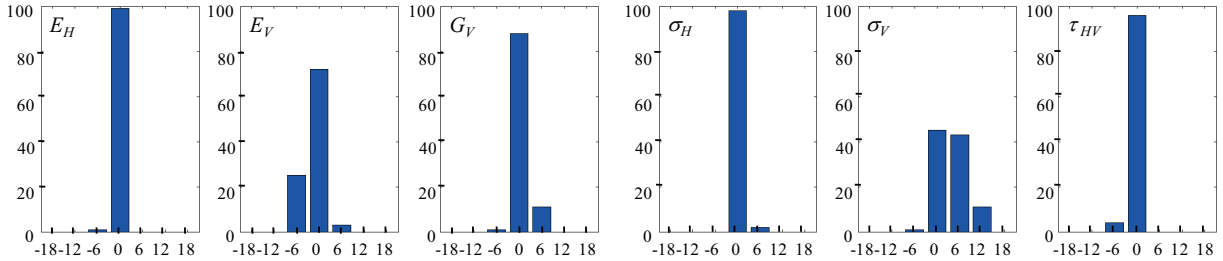

Figure 10. Performance of POD–ANN procedure for parameter identification based on pseudoexperimental DIC measurements affected by random noise of $\pm 5.0\,\mu$m. In the abscissa, intervals of relative errors; in the ordinate, the percentage of errors found to fall within each such interval. Estimates of elastic moduli (first three diagrams) and stresses (remaining diagrams).

by DIC. Now $M = 3375$, consistently with Table 1, and, hence, matrix U has dimensions $N \times M = 2086 \times 3375$. The design of the ANN for stresses, through considerations similar to those outlined above for elastic properties (again with random noise $\pm 5\,\mu$m) here omitted for brevity, leads to the following ANN architecture: 7–14–3. The set of 100 samples employed for the final accuracy checks of the POD–RBF–TRA inverse analysis procedure developed in Section 3, has been used for similar checks on the alternative procedure POD–ANN described in the present section. Results concerning elastic moduli and stresses are presented in Figure 10. A comparison of these results with those achieved in Section 3 and illustrated in Figure 8 leads to the conclusion that both procedures POD–RBF–TRA and POD–ANN associated with flat-jack test with T-geometry and to DIC full-field measurements are adequate to the identification of orthotropic elasticity parameters and pre-existing plane-stress state. Computer efforts and estimation accuracy are comparable.

5. Estimation of inelastic parameters

Inelastic parameters, not estimated by flat-jack tests in the state-of-the-art diagnostic methodology, are the objectives of the operative steps (d) and (e) and of the identification phase (III) in the present method

(Section 2.2 and Figure 2). A simplification which is adopted herein for inelastic parameter identification consists of the following two assumptions: (i) elastic isotropy, consistently with the isotropy attributed to the inelastic model; (ii) a pre-existing plane stress state in which σ_H is the only nonvanishing stress component. The elastic modulus E and the stress σ_H are supposed to have been identified in the previous phases of the procedure.

The concrete inelastic parameters to identify are internal friction angle β, tensile strength f_t and fracture energy G_f; the dilatancy $\psi = 40°$ is regarded as assumed a priori.

The experimental data supposed to be provided by a DIC photo, now over the smaller area (ROI) shown in Figure 2d-e, consist of $N = 210 \times 2 = 420$ displacement components on the dam surface. The reduction of the ROI implies a higher resolution in measuring the displacements due to inelastic strains and quasi-brittle fracture which occur when pressure is increased beyond the elastic threshold. Here it is assumed that the flat-jack pressure is increased up to the value of $10\,\text{MPa}$ and that a sequence of 6 photos is taken at pressure levels $p = 5, \dots, 10\,\text{MPa}$ (with increments of $1\,\text{MPa}$).

The inverse problem, (12), is now reformulated so that the displacements monitored at distinct load levels be all involved in the definition of the residual $\boldsymbol{R}$ and, hence, of the discrepancy function to minimize. To this purpose vector $\boldsymbol{u}$ becomes a vector consisting of 6 subvectors, each one representing the displacements of the ROI grid-nodes at one of the considered pressure levels. Therefore, henceforth in this section, the term snapshot will denote a vector $\boldsymbol{u}$ with $6N = 2520$ components. It is worth noting that the response of the system, now nonlinear under the applied pressure, is influenced by the effects (mainly the self-stresses) generated during the slot cutting, i.e during the operative step (b) of the experimental procedure (see Section 2.2). Such effects are accounted for in the FE simulations of the whole experimental procedure.

The FE model is substantially the same as the one employed for the preceding simulations, Figure 4a and Figure 4b; only a local mesh refinement is carried out over a volume in the vicinity of the free surface where the ROI is located, in order to accurately compute displacements at the ROI grid-nodes.

The expert-preselected domain for the three parameters sought is defined as follows:

$$1 \le f_t \le 6\,\text{MPa}, \qquad 50 \le G_f \le 300\,\text{Nm}^{-1}, \qquad 40° \le \beta \le 80°. \qquad (26)$$

In view of the novelty of flat-jack diagnostic tests and of inverse analyses based on them beyond the elastic range, the identifiability of the three parameters β, f_t and G_f has been investigated by various numerical exercises, some of which are outlined in what follows.

(A) Using the trust region algorithm (TRA) outlined in Section 3 (but here without POD preliminaries), parameter identifications are carried out using the following reference values of the constitutive parameters and the stress σ_H:

$$\text{elastic parameters and stress:} \quad E_H = 15\,\text{GPa} \quad \nu = 0.2 \quad \sigma_H = -3\,\text{MPa}$$
$$\text{inelastic parameters:} \quad f_t = 3\,\text{MPa} \quad \beta = 68° \quad G_f = 180\,\text{N/m}$$

The input consists of all 2520 pseudoexperimental DIC data generated by a simulation with the FE model of Figure 4, with the material model of Section 2; first inverse analysis is performed and illustrated in Figure 11a, starting from input data perturbed by a random noise governed by uniform probability density
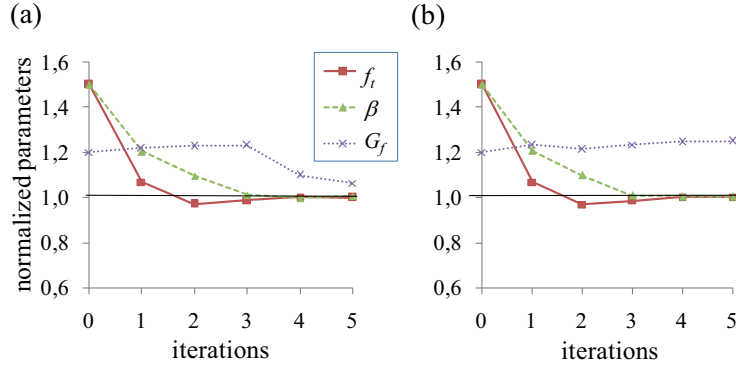
Figure 11. Convergence of TRA applied to inelastic parameter identification with input perturbed by random noise of (a) $\pm0.25\,\mu$m and (b) $\pm0.5\,\mu$m.

function over the interval $\pm0.25\,\mu$m. With the same initialization and similar input perturbation amplified to $\pm0.5\,\mu$m the TRA consequence is shown in Figure 11b.

This computational exercise (and others not considered here) shows stability of the estimation procedure with respect to (expected) experimental inaccuracies.

(B) In order to endorse the identification of inelastic parameters with the advantages of POD exhibited by the methodology presented in the preceding sections, the extended POD–ANN procedure (Section 4) is presented here below by means of a numerical exercise.

Clearly, in the present context, the ANN to design must be fed not only with the amplitude $\bar{a}$ (in the truncated basis $\bar{\Phi}$; see (6)) of the new experimental data vector u, but also with the results of the preceding identifications of the elastic modulus E (isotropic case) and of the pre-existing stress σ_{H}. Therefore a five-dimensional parameter space has to be considered; by choosing in such space the preselected domain and grid specified in Table 2, the number of grid nodes and the of corresponding snapshots (to be computed by FE test simulations once-and-for-all) reach now 5346.

Only $\bar{M} = 6$ eigenvalues of the matrix $D = U^T U$ of order $M = 5346$ are preserved in the POD compression (i.e., after POD truncation). Therefore, each snapshot vector u containing $N = 2520$ components is replaced by its amplitude $\bar{a}$ of 6 components only; the ANN input layer consists of $6 + 2 = 8$ neurons the additional two being grid values of E and σ_{H}.

The 5346 available patterns (node parameters p emerging from Table 2 and compressed amplitudes $\bar{a}$ of the corresponding snapshot u) have been employed to train (using 3742 randomly selected patterns), test (using 802 patterns) and validate (using 802 patterns) an ANN of the kind specified in Section 4,

	E/GPa	$\sigma_{\mathrm{H}}/\mathrm{MPa}$	f_t/MPa	$G_f/(\mathrm{N/m})$	$\beta/^\circ$
min	10	-6	1	50	40
step	5	2	1	25	5
max	20	-2	6	300	80

Table 2. Grid in parameter space for phase (III) of the identification procedure.
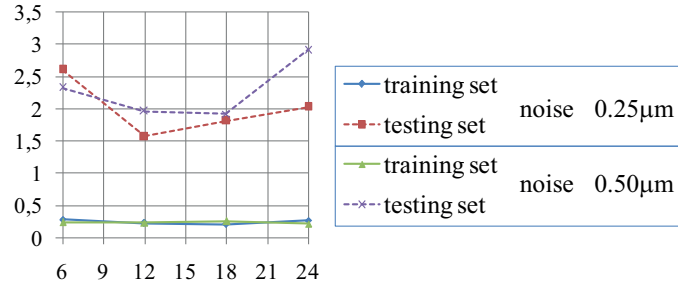
Figure 12. Influence of H, the number of neurons in the hidden layer, on the POD-ANN procedure for estimating the inelastic parameters for $I = 8$ input neurons.

by the same procedure outlined there. The optimal architecture turned out to consist of a single hidden layer with 12 neurons; see Figure 12.

The following numerical exercise (among others omitted for brevity) carried out by the POD–ANN procedure above, not only tests it but also provides motivation for further developments in terms of experimental techniques.

Over the five-dimensional domain specified by Table 2, a set of 100 points is randomly chosen. The corresponding snapshots are computed by the same FE and constitutive material models described in Section 2. Inverse analyses are performed, with perturbation noise of $\pm 0.25\,\mu$m and $\pm 0.50\,\mu$m, by POD–ANN calibrated in what precedes. The resulting estimates of inelastic parameters are compared to their "exact values" employed as input for the direct analyses which led to the snapshots. The differences are shown in Figure 13, in a fashion similar to that of Figures 8 and 10 for comparisons.

Clearly, the scatter of the estimation error for the fracture energy G_f is rather large. Such undesirable circumstance can be explained as follows: when preexisting horizontal compression σ_{H} or tensile strength f_t are rather large within the selected domain, with the assumed jack pressure the fracturing process is very limited and, therefore, fracture energy has small influence on the DIC measurable displacements; hence, larger becomes the sensitivity of its estimate to perturbations (including experimental errors and round-off errors in the numerical identification procedure).
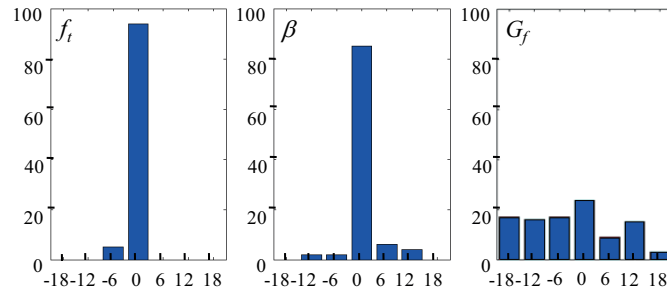

Figure 13. Performance of POD–ANN procedure for inelastic parameter identification based on pseudoexperimental DIC measurements affected by random noise of $\pm 0.5\,\mu$m. In the abscissa, intervals of relative errors; in the ordinate, the percentage of errors found to fall within each such interval.

Possible remedies, consistent with desirable progress in dam monitoring practice, are the following provisions: (i) an assessment of the crack opening displacement near the slot tip is extracted from the DIC measurements and employed to control the pressure growth in the flat-jack; the test is performed when the assessed displacement exceeds a pre-established threshold; (ii) two additional vertical slots are cut as shown in Figure 2f in order to practically eliminate the horizontal compressive stress.

Provision (i) would imply that the jack pressure is no longer selected a priori, once and for all, but becomes variable and represents an additional input into the ANN. The geometry change (ii) obviously makes the test somewhat more costly and more destructive. Both of these prospects could be worth of further investigations.

6. Closing remarks

The procedure resulting from this study on diagnostic analysis of concrete dams based on flat-jack tests, exhibits the following main advantages with respect to the present practice and also with respect to recent results achieved on the same subject (see e.g., [Fedele and Maier 2007]): (a) less damaging slots, due to the T-shaped geometry adopted here; (b) more experimental data which can be economically acquired by digital image correlation (DIC) and which regularize the inverse problem to solve; (c) assessment of inelastic properties of concrete, besides the elastic ones and the stress state considered so far; (d) computational efforts concentrated in a preparatory stage of multiple test simulations (say by a finite element commercial code) and much alleviated in applications, which can now be performed routinely in situ, either by soft computing, specifically by means of artificial neural networks (ANN), or by a popular mathematical programming procedure (TRA).

Besides obvious improvements of FE modeling and of constitutive models, meaningful further developments (which are partly in progress and partly desirable in the near future) can be listed as follows: (i) experimental validation in real-life engineering circumstances, particularly, as for the operative aspects of DIC measurements, on a concrete dam surface, after selection of suitable instruments and supports; (ii) improved jacks, with suitable textile instead of steel sheet in order to reduce friction and weld effects, and with larger radius complying with possible large size of aggregate in dam concrete; (iii) extension of the present approach and methodology to inhomogeneous periodic materials such as masonry, on which traditional flat-jack tests are frequently adopted to structural diagnosis purposes; (iv) transition from the present deterministic batch estimation methodology to the stochastic one (methodologically like, e.g., in [Fedele and Maier 2007] for a first step in this direction), with elaboration of both experimental data and their random errors in order to reach estimates accompanied by their covariance matrices.

An inverse analysis performable in situ with economy and brevity by exploiting preliminary computational results (and their POD approximation) achieved in a computing center, as proposed in this paper, turns out to be promising for structural diagnosis and for mechanical characterization, in dam engineering as well as in other technologies.

References

[Ahmed et al. 2003] T. Ahmed, E. Burley, S. Rigden, and A. Abu-Tair, "The effect of alkali reactivity on the mechanical properties of concrete", *Constr. Building Mater.* **17**:2 (2003), 123–144.

[Ardito et al. 2008] R. Ardito, G. Maier, and G. Massalongo, "Diagnostic analysis of concrete dams based on seasonal hydrostatic loading", *Eng. Struct.* **30**:11 (2008), 3176–3185.

[Avril et al. 2008] S. Avril, M. Bonnet, A. Bretelle, M. Grediac, F. Hild, P. Ienny, F. Latourte, D. Lemosse, S. Pagano, E. Pagnacco, et al., "Overview of identification methods of mechanical parameters based on full-field measurements", *Exp. Mech.* **48**:4 (2008), 381–402.

[Bishop 1995] C. M. Bishop, *Neural networks for pattern recognition*, The Clarendon Press Oxford University Press, New York, 1995. With a foreword by Geoffrey Hinton.

[Bishop 2006] C. M. Bishop, *Pattern recognition and machine learning*, Springer, New York, 2006.

[Buhmann 2003] M. D. Buhmann, *Radial basis functions: theory and implementations*, vol. 12, Cambridge Monographs on Applied and Computational Mathematics, Cambridge University Press, Cambridge, 2003.

[Chatterjee 2000] A. Chatterjee, "An introduction to the proper orthogonal decomposition", *Current Sci.* **78**:7 (2000), 808–817.

[Coleman and Li 1996] T. F. Coleman and Y. Li, "An interior trust region approach for nonlinear minimization subject to bounds", *SIAM J. Optim.* **6**:2 (1996), 418–445.

[Comi et al. 2009] C. Comi, R. Fedele, and U. Perego, "A chemo-thermo-damage model for the analysis of concrete dams affected by alkali-silica reaction", *Mech. Mater.* **41**:3 (2009), 210–230.

[Conn et al. 2000] A. R. Conn, N. I. M. Gould, and P. L. Toint, *Trust-region methods*, MPS/SIAM Series on Optimization, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2000.

[Fedele and Maier 2007] R. Fedele and G. Maier, "Flat-jack tests and inverse analysis for the identification of stress states and elastic properties in concrete dams", *Meccanica* **42**:4 (2007), 387–402.

[Fedele et al. 2005] R. Fedele, G. Maier, and B. Miller, "Identification of elastic stiffness and local stresses in concrete dams by in situ tests and neural networks", *Struct. Infrastruct. Eng.* **1**:3 (2005), 165–180.

[Fedele et al. 2006] R. Fedele, G. Maier, and B. Miller, "Health assessment of concrete dams by overall inverse analyses and neural networks", *Int. J. Fract.* **137**:1-4 (2006), 151–172.

[Goodman 1989] R. E. Goodman, *Introduction to rock mechanics*, 2nd ed., Wiley, New York, 1989.

[Hagan et al. 1996] M. T. Hagan, H. B. Demuth, and M. H. Beale, *Neural network design*, PWS Publishing, Boston, 1996.

[Hassing et al. 2010] P. M. Hassing, H. Fang, and Q. Wang, "Identification of material parameters for McGinty's model using adaptive RBFs and optimization", *Struct. Multidiscip. O.* **42**:2 (2010), 233–242.

[Haykin 1998] S. Haykin, *Neural networks: a comprehensive foundation*, 2nd ed., Prentice Hall, Upper Saddle River, NJ, 1998.

[Hild and Roux 2006] F. Hild and S. Roux, "Digital image correlation: from displacement measurement to identification of elastic properties: a review", *Strain* **42**:2 (2006), 69–80.

[Jirasek and Bazant 2001] M. Jirasek and Z. P. Bazant, *Inelastic analysis of structures*, Wiley, New York, 2001.

[Kleiber et al. 1997] M. Kleiber, H. Antunez, T. D. Hien, and P. Kowalczyk, *Parameter sensitivity in nonlinear mechanics: theory and finite element computations*, John Wiley, Chichester,NY, 1997.

[Leite and Corthesy 2001] M. H. Leite and R. Corthesy, "Stress measurements in concrete structures with modified doorstopper technique", *ACI Struct. J.* **98**:5 (2001), 619–628.

[Liang et al. 2002] Y. C. Liang, H. P. Lee, S. P. Lim, W. Z. Lin, K. H. Lee, and C. G. Wu, "Proper orthogonal decomposition and its applications. I. Theory", *J. Sound Vibration* **252**:3 (2002), 527–544.

[Loh and Wu 2000] C. H. Loh and T. C. Wu, "System identification of Fei-Tsui arch dam from forced vibration and seismic response data", *J. Earthquake Eng.* **4**:4 (2000), 511–537.

[Lubliner 1990] J. Lubliner, *Plasticity theory*, MacMillan, New York, 1990.

[Maier et al. 2004] G. Maier, R. Ardito, and R. Fedele, "Inverse analysis problems in structural engineering of concrete dams", pp. 97–107 in *Computational mechanics*, edited by Z. H. Yao et al., Tsinghua Univ. Press, Beijing, 2004.

[Nabney 2002] I. Nabney, *NETLAB: algorithms for pattern recognition*, Springer, London, 2002.

[Nocedal and Wright 1999] J. Nocedal and S. J. Wright, *Numerical optimization*, Springer Series in Operations Research, Springer-Verlag, New York, 1999.

[Ostrowski et al. 2008] Z. Ostrowski, R. Bialecki, and A. Kassab, "Solving inverse heat conduction problems using trained POD-RBF network inverse method", *Inverse Prob. Sci. Eng.* **16**:1 (2008), 39–54.

[Swamy 1992] R. N. Swamy, *The alkali-silica reaction in concrete*, Van Nostrand Reinhold, New York, 1992.

[Waszczyszyn 1999] Z. Waszczyszyn, *Neural networks in the analysis and design of structures*, Springer, New York, 1999.

[Wu 1995] Z. M. Wu, "Compactly supported positive definite radial functions", *Adv. Comput. Math.* **4**:3 (1995), 283–292.

[Wu et al. 2003] C. G. Wu, Y. C. Liang, W. Z. Lin, H. P. Lee, and S. P. Lim, "A note on equivalence of proper orthogonal decomposition methods", *J. Sound Vibration* **265**:5 (2003), 1103–1110.

TOMASZ GARBOWSKI: tomasz.garbowski@put.poznan.pl
Poznan University of Technology, Institute of Structural Engineering, ul. Piotrowo 5, 60-965 Poznan, Poland

GIULIO MAIER: giulio.maier@polimi.it
Dipartimento di Ingegneria Strutturale, Politecnico di Milano, Piazza Leonardo da Vinci 32, I-20133 Milan, Italy
giulio.maier@polimi.it

GIORGIO NOVATI: giorgio.novati@polimi.it
Dipartimento di Ingegneria Strutturale, Politecnico di Milano, Piazza Leonardo da Vinci 32, I-20133 Milan, Italy

■msp

A ZERO-STIFFNESS ELASTIC SHELL STRUCTURE

SIMON D. GUEST, ELIZBAR KEBADZE AND SERGIO PELLEGRINO

A remarkable shell structure is described that, due to a particular combination of geometry and initial stress, has zero stiffness for any finite deformation along a twisting path; the shell is in a neutrally stable state of equilibrium. Initially the shell is straight in a longitudinal direction, but has a constant, nonzero curvature in the transverse direction. If residual stresses are induced in the shell by, for example, plastic deformation, to leave a particular resultant bending moment, then an analytical inextensional model of the shell shows it to have no change in energy along a path of twisted configurations. Real shells become closer to the inextensional idealization as their thickness is decreased; experimental thin-shell models have confirmed the neutrally stable configurations predicted by the inextensional theory. A simple model is described that shows that the resultant bending moment that leads to zero stiffness gives the shell a hidden symmetry, which explains this remarkable property.

1. Introduction

A novel zero-stiffness structure is described. The structure is a thin shell that is initially straight in a longitudinal direction, but has a uniform, nonzero curvature in the transverse direction. The structure is prestressed, and the interaction of the elastic properties with the prestress is such that the structure can be deformed without any applied load; this is not a local phenomenon — the structure can continue to be deformed in a finite closed path. Equivalently, the structure is neutrally stable: there is no change in total internal strain energy as the structure is deformed, even though any particular component of strain energy will vary. An experimental model of this structure, made from a sheet of copper beryllium of thickness 0.1 mm and width 30 mm, is shown in Figure 1.

The ability to deform a structure without load is quite unexpected when initially observed, and is clearly a function of the prestress that the shell carries. Certainly it is well known that the stiffness of structures changes with applied load. Stable structures can become unstable when loaded: a simple example is the buckling of a strut through the application of axial load. At the cusp between stability and instability there may then be a point of neutral stability, where a structure has zero stiffness: to first order, there is no change in load with displacement. However, while typically for buckling phenomena this point of neutral stability/zero stiffness is isolated, it is also possible to engineer systems which, when they buckle, are neutrally stable for large deformations [Tarnai 2003].

The stiffness of structures also changes through *prestress*, where the structure is loaded against itself. A classical example of this behaviour is provided by tensegrity structures [Calladine 1978; Guest 2011], which typically rely on prestress in order to be able to act as structures at all. However, even for tensegrities with rigid compression members, increasing the relative level of prestress can reduce the stiffness
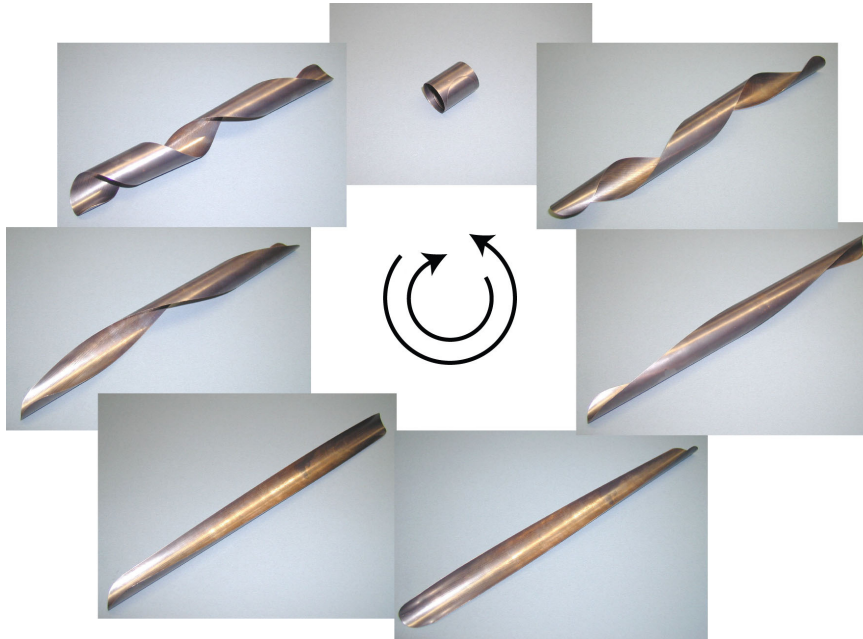
Figure 1. A series of different configurations of the *same* zero-stiffness shell structure. In each configuration, the shell is held in place by no more than the friction with the underlying surface. The shell can be transformed between configurations in both directions, clockwise and anticlockwise. The analytical model in Section 2 predicts that the shell will, in each case, be wrapped around an underlying cylinder (see Figure 3) of constant radius, and it can be seen that this is the case here.

of some modes of deformation, and for extreme levels of prestress, can also lead to structures with zero stiffness, even for large deformations [Schenk et al. 2007].

The present paper deals with thin shell structures that are straight longitudinally, but uniformly curved in the transverse direction, as shown in Figure 2a. Shell structures of this type are used in steel tape measures and also as lightweight deployable booms for spacecraft [Rimrott 1965]. The mechanics of
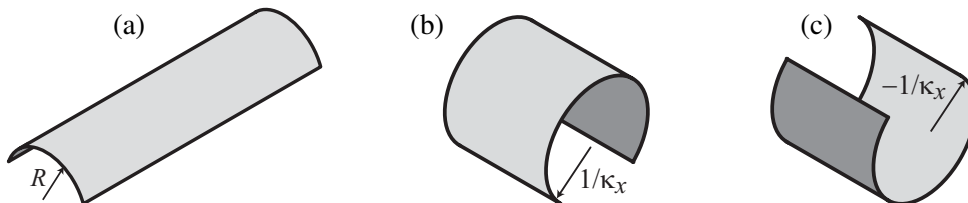

Figure 2. (a) A shell that is straight longitudinally, but curved transversely, with two coiling modes, (b) and (c). A configuration change from (a) to (b) involves *same-sense* bending: the centres of curvature are on the same side of the shell. A configuration change from (a) to (c) involves *opposite-sense bending*: the centres of curvature are on opposite sides of the shell.

such structures has been studied extensively [Mansfield 1973; Seffen and Pellegrino 1999], although until recently most studies have focussed on structures that are both isotropic and initially unstressed.

Recently, it has become clear that interesting properties, and in particular bistability, can be engendered if these curved thin shell structures are made to be anisotropic [Guest and Pellegrino 2006], or if the structures are prestressed [Kebadze et al. 2004]. If the shells are given the correct set of anisotropic bending properties, for example, through being manufactured in fibre-reinforced plastic, then the shell can be made bistable so that the second stable state has the *same* sense of bending, as shown in Figures 2a and 2b. Alternatively, if an isotropic shell is correctly prestressed, then the shell can be made bistable with the second stable state having the *opposite* sense of bending, as shown in Figures 2a and 2c. In this case, in the initial configuration, the shell is prestressed in bending, so that it wishes to coil up, but this is prevented by the structural depth of the curved shape. In [Pellegrino 2005] a shell is described that exploits this mode while ensuring that the two states have the same stored strain energy, so that a partially coiled shell can coil and uncoil without change of energy, and is neutrally stable. This is the only previous example of a zero-stiffness shell of which we are aware.

The present paper explores the case of an isotropic shell which is prestressed in the opposite sense to that studied in [Kebadze et al. 2004], so that the prestress favours same-sense bending. It will show that bistability cannot be engendered, but remarkably, for a particular value of prestress, the structure can be left without any torsional stiffness.

2. Analytical model

The basic analytical model that we use is essentially identical to that described in [Kebadze et al. 2004; Guest and Pellegrino 2006]. We make two geometric assumptions: that the shell is inextensional, and that the curvature of the shell is uniform across its midsurface. The inextensional assumption is valid for thin shells, where the energy required to stretch the shell dwarfs the energy required to bend the shell. A consequence of our assumptions is that we are neglecting boundary effects; for further discussion of this, see [Galletly and Guest 2004]. The two geometric assumptions together imply that we can consider the shell midsurface as lying on an underlying cylindrical surface, as shown in Figure 3. The (uniform) curvature of the surface can then be described in terms of two parameters, the nonzero principal curvature of the underlying cylinder, C, and the orientation of the local axes (x, y) with respect to the axis of the cylinder, defined by an angle θ.
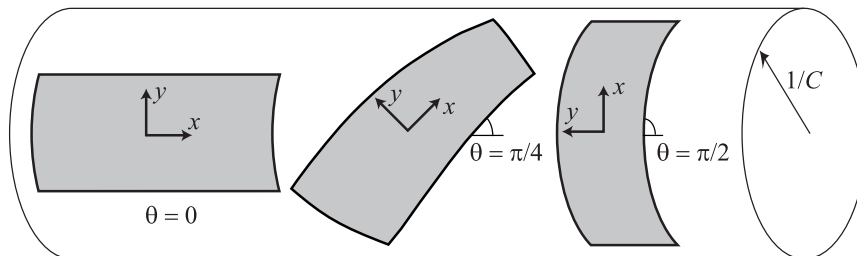

Figure 3. Definition of the geometry of the shell in terms of an underlying cylinder with curvature C. The angle θ specifies the orientation of the shell with respect to the cylinder.

The curvature of the shell can thus be described by the curvature vector $\boldsymbol{\kappa}$:

$$\boldsymbol{\kappa} = \begin{bmatrix} \kappa_x \\ \kappa_y \\ 2\kappa_{xy} \end{bmatrix} = \frac{C}{2} \begin{bmatrix} 1 - \cos 2\theta \\ 1 + \cos 2\theta \\ 2 \sin 2\theta \end{bmatrix}. \tag{1}$$

The curvature components are given by $\kappa_x = -\partial^2 w/dx^2$, $\kappa_y = -\partial^2 w/dy^2$, and $\kappa_{xy} = -\partial^2 w/dx\,dy$, where w is the relative displacement out of the plane defined by x and y. The transformation to the x, y curvilinear coordinates is obtained from, for example, a Mohr circle construction, as described in [Guest and Pellegrino 2006]. Note that the definition $\kappa_{xy} = -\partial^2 w/dx\,dy$ for the twisting curvature is standard in the plates and shells literature, including [Kebadze et al. 2004], but is only half the value commonly used in the composites literature, including [Guest and Pellegrino 2006].

We assume an initial configuration for the shell with $\theta_0 = 0$ and $C_0 = 1/R$, so the change in curvature to any other configuration is given by

$$\Delta\boldsymbol{\kappa} = \frac{C}{2} \begin{bmatrix} 1 - \cos 2\theta \\ 1 + \cos 2\theta - \dfrac{2}{CR} \\ 2 \sin 2\theta \end{bmatrix}. \tag{2}$$

We also assume that the shell is prestressed in the initial configuration. As the shell is straight in the x-direction, it cannot sustain any moment/unit length along the edge normal to the y-axis, so in the initial configuration $m_y = m_{xy} = 0$. However, because of the curvature in the y-direction, the depth of the cross-section allows a uniform initial moment $m_x = m$ to be equilibrated by midplane forces in the shell, as shown in Figure 4.

In a general configuration we define the moment/unit length carried by the shell as a vector $\boldsymbol{m}$,

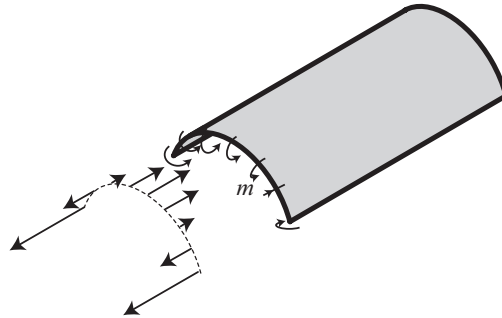$$\boldsymbol{m} = \begin{bmatrix} m_x \\ m_y \\ m_{xy} \end{bmatrix}, \tag{3}$$

Figure 4. Initial prestress in the shell. The uniform moment/unit length $m_x = m$ is equilibrated by a distribution of midplane forces in the shell. The moment m is uniform throughout the shell, except for a narrow boundary layer at the free ends.

with an initial value

$$\boldsymbol{m}_0 = \begin{bmatrix} m \\ 0 \\ 0 \end{bmatrix}. \tag{4}$$

We assume linear-elastic material behaviour, and therefore in a general configuration, the moment will be given by

$$\boldsymbol{m} = \boldsymbol{D}\Delta\kappa + \boldsymbol{m}_0. \tag{5}$$

The bending stiffness matrix $\boldsymbol{D}$ is given by

$$\boldsymbol{D} = D \begin{bmatrix} 1 & \nu & 0 \\ \nu & 1 & 0 \\ 0 & 0 & (1-\nu)/2 \end{bmatrix}, \tag{6}$$

where ν is the Poisson's ratio of the material, and D is the shell bending stiffness, defined in terms of the Young's modulus E, thickness t, and the Poisson's ratio as

$$D = \frac{Et^3}{12(1-\nu)^2}. \tag{7}$$

We define the strain energy U as the energy stored in the shell per unit area due to its deformation away from the initial configuration, and so

$$U = \tfrac{1}{2}\Delta\kappa^{\mathrm{T}}\boldsymbol{D}\Delta\kappa + \Delta\kappa^{\mathrm{T}}\boldsymbol{m}_0. \tag{8}$$

Finally, we write everything in a nondimensional form (with a hat, ˆ) in terms of the bending stiffness, D, and initial radius of curvature R:

$$\hat{U} = \frac{UR^2}{D}, \quad \hat{\boldsymbol{D}} = \frac{\boldsymbol{D}}{D}, \quad \hat{\kappa} = R\kappa, \quad \hat{m} = \frac{R}{D}m, \quad \hat{C} = CR.$$

In [Kebadze et al. 2004] the behaviour of this system was explored when the initial moment $\hat{m} = mR/D$ is positive, which leads to bistable behaviour. The present paper notes the remarkable behaviour associated with the value

$$\hat{m} = -(1-\nu). \tag{9}$$

Figure 5 shows the variation of $\hat{U}$ with $\hat{C}$ and θ for three values of $\hat{m}$: $\hat{m} = 0$, $\hat{m} = -(1-\nu)$, and $\hat{m} = -2(1-\nu)$.

The key plot is Figure 5b. For $\hat{m} = -(1-\nu)$, there is no change in stored energy $\hat{U}$ with θ. Thus, from the initial configuration, a series of new twisted configurations with the same underlying curvature $\hat{C} = 1$ are possible, and these are clearly shown in the experimental results shown in Figure 1. The structure is in a state of neutral equilibrium, and has zero stiffness, even for large excursions along this deformation path.

Figures 5a and 5c represent the behaviour of the shell for values of $\hat{m}$ respectively greater or smaller than the critical value of $-(1-\nu)$. Figure 5a shows the case where $\hat{m} = 0$, and is hence identical to the isotropic plot in [Guest and Pellegrino 2006]. There are two equilibrium configurations, marked M and N: M is the stable initial configuration, and N is an unstable coiled configuration, where $\theta = \pi/2$. In
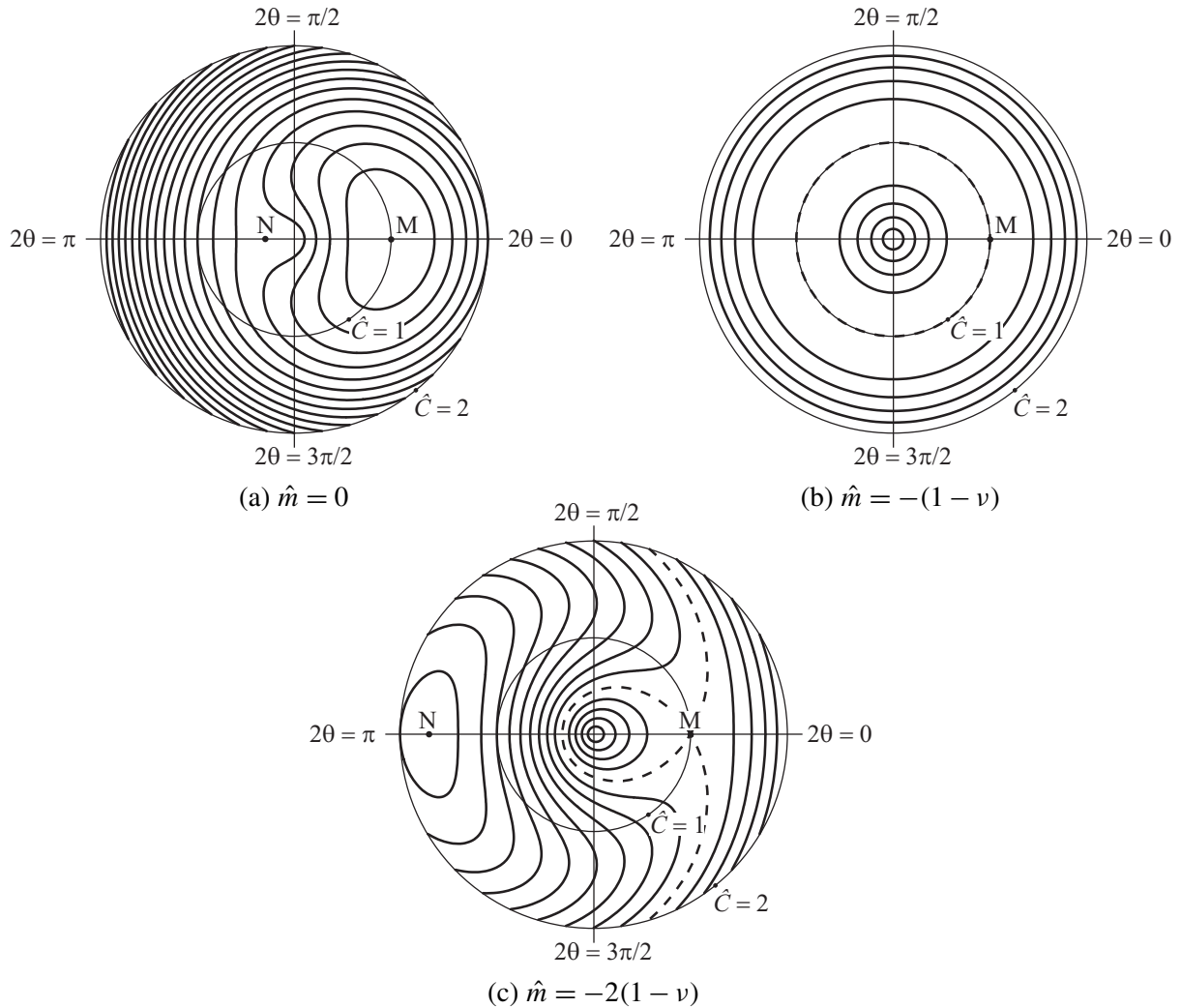
Figure 5. Polar plots of the nondimensional energy $\hat{U}$ plotted as a function of $\hat{C}$ and θ for three values of initial moment $\hat{m}$. The initial configuration is labelled as M, and $\hat{U} = 0$ at this point. Contours are plotted for $\hat{U} = \dots, -0.1, 0, 0.1, \dots$, with $\hat{U} = 0$ plotted as a dashed line.

fact, for any value of $\hat{m}$ in the range $0 \geq \hat{m} > -(1 - \nu)$ similar behaviour is observed, with a stable initial configuration, and an unstable coiled configuration.

Figure 5c shows the case where $\hat{m}$ is twice the critical value of $-(1 - \nu)$. Again there are two equilibrium configurations, marked M and N: M is the initial configuration, which is now unstable, and N is a coiled configuration with $\theta = \pi/2$, which is now stable. In fact, for any value of $\hat{m} < -(1 - \nu)$ similar behaviour is observed, with an unstable initial configuration, and a stable coiled configuration. The energy $\hat{U}$ is negative at N, but this is simply a consequence of arbitrarily setting $\hat{U} = 0$ at the original configuration M.

Note that in each of the plots, there appears to be a maximum in $\hat{U}$ at $\hat{C} = 0$, but this is simply an artefact of the way the data is plotted: $\hat{U}$ will continue to increase for $\hat{C} < 0$, but this portion of the data is not plotted in the present paper, as no interesting behaviour is observed in this regime (unlike in [Kebadze et al. 2004], where additional stable states are found with $\hat{C} < 0$ for $\hat{m}$ positive).

3. Experimental results

Experimental verification of the zero stiffness behaviour was obtained through models made from a thin sheet of copper beryllium (CuBe). The basic manufacturing protocol was as described in [Kebadze et al. 2004]. The shells were formed in a curved initial state from annealed CuBe with $t = 0.1$ mm and a width of 30 mm, and were then age-hardened to give a stress-free curved shell.

The prestress moment was imposed by passing the unstressed shell through a set of rollers, which leads to a residual moment m through the mechanism described in [Kebadze et al. 2004]. Although Section 2 gives a precise value of m for zero stiffness, it is actually difficult to predict the rolling parameters that will give this value of m, as this depends on the precise strain-hardening characteristics of the CuBe as it yields. Thus, in practice we proceeded by trial and error to fine-tune the rolling process to give shells that had no torsional stiffness. The final result of this process is shown in Figure 1, where a series of configurations of the same shell is shown, each with an underlying curvature $C \approx 1/12.5$ mm. In each configuration, the shell is only held in place by friction with the underlying surface. We have not attempted any detailed experimental measurements on these models.

4. Conceptual disk model

This section describes a simple conceptual model that has two aims: firstly it will reveal a "hidden" symmetry that provides an explanation for the particular value of prestress moment m that provides zero stiffness; and secondly, it describes a zero-stiffness shell structure that doesn't require an assumption that the shell is so thin that it can be considered to be inextensional.

Consider a thin circular bimetallic flat disk — for example, one made of brass and steel, as in Figure 6a. As the disk is heated from its initial flat stress-free state, the brass will want to expand more than the steel,
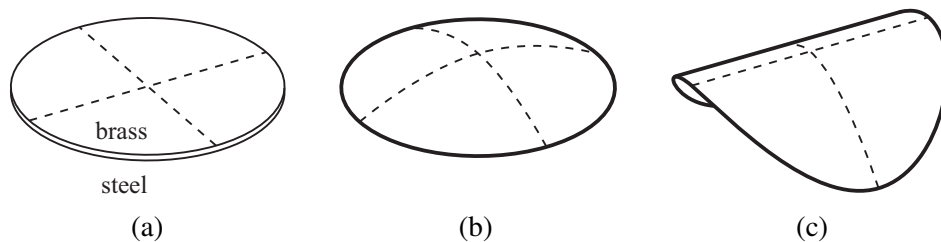

(a)   (b)   (c)

Figure 6. A conceptual model of the formation of a zero-stiffness shell from a bimetallic disk. As the disk is heated from (a), it will initially form a domed structure, as in (b). After further heating, the bending response will bifurcate and become nonisotropic; and in the limit the structure will become cylindrical; see (c). The principal directions of bending in (c) are arbitrary — any other choice of bending direction would lead to a twisted form of the structure having the same stored internal energy.

and if the disk is to remain flat, this will lead to a uniform residual moment in the shell. In practice, the disk will dome slightly, as shown in Figure 6b, but to do this the surface changes its Gaussian curvature, which requires in-plane stretching [Calladine 1988]. At some point (which depends on the thickness of the shell) a bifurcation will take place, following which the curvature will no longer be uniform in all directions: in some arbitrary principal direction the curvature will decrease, while the curvature will increase in the perpendicular direction. As heating is further increased, the disk will approach a cylindrical configuration, as shown in Figure 6c; this process is described in more detail in [Freund 2000; Seffen and McMahon 2007].

It is clear that the process of heating a bimetallic disk must lead to a zero-stiffness shell at any point after the disk has bifurcated. The bifurcation takes place about an arbitrary axis, and whichever axis is chosen the stored strain energy will be the same. Deforming the shell on the continuous path through states with varying axes of bifurcation will not change the stored energy, and hence it can be concluded that the path is neutrally stable. We shall see that in the extreme case of a thin shell with no in-plane deformation, this will reproduce the mode described in Section 2; but the bifurcated bimetallic disk does not require any assumption about being thin, or about boundary conditions, to have zero stiffness.

For the assumption of a thin inextensional shell, the bimetallic disk model can be used to calculate the critical value of prestress moment found in Section 2. Consider a preliminary state of the disk where the disk has been heated, but is held flat. In this state, there will be a uniform (negatively valued) moment due to the temperature change, m_t, but no curvature:

$$\boldsymbol{m}_i = \begin{bmatrix} m_t \\ m_t \\ 0 \end{bmatrix}, \quad \boldsymbol{\kappa}_i = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \tag{10}$$

Consider now that the disk is released, and is allowed to increase its curvature in the y-direction until the moment m_y becomes zero, at which point it has reached the initial state $(\boldsymbol{m}_0, \boldsymbol{\kappa}_0)$ considered in Section 2. As the change from the preliminary to the initial state is elastic, we can write

$$\boldsymbol{m}_0 = \boldsymbol{D}(\boldsymbol{\kappa}_0 - \boldsymbol{\kappa}_i) + \boldsymbol{m}_i, \tag{11}$$

and hence

$$\begin{bmatrix} m \\ 0 \\ 0 \end{bmatrix} = D \begin{bmatrix} 1 & \nu & 0 \\ \nu & 1 & 0 \\ 0 & 0 & (1-\nu)/2 \end{bmatrix} \begin{bmatrix} 0 \\ 1/R \\ 0 \end{bmatrix} + \begin{bmatrix} m_t \\ m_t \\ 0 \end{bmatrix}. \tag{12}$$

To satisfy this equation in the y-direction, we must have

$$m_t = -\frac{D}{R}, \tag{13}$$

and hence

$$m = \frac{D\nu}{R} - \frac{D}{R}. \tag{14}$$

Thus, the residual prestress moment, written in nondimensional form,

$$\hat{m} = \frac{mR}{D} = -(1-\nu), \tag{15}$$

is exactly the critical moment identified in Section 2.

5. Conclusion

The disk model presented in Section 4 has shown that the zero-stiffness mode identified in Section 2 can be explained simply by consideration of a hidden symmetry of the shell structure: if the structure is flattened, then the resultant moment in the shell doesn't vary with direction, and bending about any axis is equally preferable. There may be minor effects associated with boundary conditions, but these didn't have a noticeable effect on the experimental structures that we manufactured, and they will certainly not be present for a circular shell structure.

Acknowledgements

References

[Calladine 1978] C. R. Calladine, "Buckminster Fuller's "Tensegrity" structures and Clerk Maxwell's rules for the construction of stiff frames", *Int. J. Solids Struct.* **14**:2 (1978), 161–172.

[Calladine 1988] C. R. Calladine, *Theory of shell structures*, Cambridge University Press, 1988.

[Freund 2000] L. B. Freund, "Substrate curvature due to thin film mismatch strain in the nonlinear deformation range", *J. Mech. Phys. Solids* **48**:6-7 (2000), 1159–1174.

[Galletly and Guest 2004] D. A. Galletly and S. D. Guest, "Bistable composite slit tubes, II: a shell model", *Int. J. Solids Struct.* **41**:16–17 (2004), 4503–4516.

[Guest 2011] S. D. Guest, "The stiffness of tensegrity structures", *IMA J. Appl. Math.* **76**:1 (2011), 57–66.

[Guest and Pellegrino 2006] S. D. Guest and S. Pellegrino, "Analytical models for bistable cylindrical shells", *Proc. R. Soc. Lond. A* **462**:2067 (2006), 839–854.

[Kebadze et al. 2004] E. Kebadze, S. D. Guest, and S. Pellegrino, "Bistable prestressed shell structures", *Int. J. Solids Struct.* **41**:11-12 (2004), 2801–2820.

[Mansfield 1973] E. H. Mansfield, "Large-deflexion torsion and flexure of initially curved strips", *Proc. R. Soc. Lond. A* **334**:1598 (1973), 279–298.

[Pellegrino 2005] S. Pellegrino, "Bistable shell structures", in *46th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference* (Austin, TX, 2005), AIAA, Reston, VA, 2005.

[Rimrott 1965] F. P. J. Rimrott, "Storable tubular extendible member: a unique machine element", *Mach. Des.* **37** (1965), 156–163.

[Schenk et al. 2007] M. Schenk, S. D. Guest, and J. L. Herder, "Zero stiffness tensegrity structures", *Int. J. Solids Struct.* **44**:20 (2007), 6569–6583.

[Seffen and McMahon 2007] K. A. Seffen and R. A. McMahon, "Heating of a uniform wafer disk", *Int. J. Mech. Sci.* **49**:2 (2007), 230–238.

[Seffen and Pellegrino 1999] K. A. Seffen and S. Pellegrino, "Deployment dynamics of tape springs", *Proc. R. Soc. Lond. A* **455**:1983 (1999), 1003–1048.

[Tarnai 2003] T. Tarnai, "Zero stiffness elastic structures", *Int. J. Mech. Sci.* **45**:3 (2003), 425–431.

SIMON D. GUEST: sdg@eng.cam.ac.uk
Department of Engineering, University of Cambridge, Trumpington Street, Cambridge CB2 1PZ, United Kingdom
http://www.eng.cam.ac.uk/~sdg

ELIZBAR KEBADZE: buba.kebadze@uk.bp.com
Department of Engineering, University of Cambridge, Trumpington Street, Cambridge CB2 1PZ, United Kingdom
Current address: *BP Exploration, Chertsey Road, Sunbury-upon-Thames, TW16 7LN, Middlesex, United Kingdom*

SERGIO PELLEGRINO: sergiop@caltech.edu
Graduate Aerospace Laboratories, California Institute of Technology, Pasadena, CA 91125, United States
http://www.pellegrino.caltech.edu

∎msp

MODAL ANALYSIS OF LAMINATED BEAMS WITH FUZZY CORE STIFFNESS/FUZZY INTERLAYER SLIP

RUDOLF HEUER AND FRANZ ZIEGLER

Dedicated to the memory of the late Marie-Louise and in the honor of Professor Charles R. Steele

It is mainly the matrix in composite structures that exhibits fuzzy randomness of the material parameters. When extending the work on two-layer and symmetric, three-layer viscoelastic beam, plate, and shell structures based on the definition of an equivalent effective homogeneous model, to include either fuzzy pure elastic interface slip or fuzzy core stiffness, by means of modal analysis we succeed in working out the effects on the dynamic properties of these fuzzy structures. Modal coupling by the light damping forces is neglected. Fully analyzed within the scope of this paper is a simply supported sandwich beam with fuzzy elastic core material parameters. The analysis of this illustrative example is based on the interval representation (that is, on the set of α-cuts) with a triangular membership function of the core shear stiffness prescribed. Membership functions of the undamped natural frequencies are defined using fuzzy set theory, however, avoiding artificial uncertainties. Under time-harmonic excitation, the dynamic magnification factors and, with light and deterministic modal structural damping taken into account, the fuzzy phase angles of the steady modal response are evaluated. Where appropriate, envelope functions are defined.

1. Introduction

The material of the matrix in composite lightweight structures has a large volume fraction and thus should be of low density and low cost: as a consequence of the technological processes its material parameters are less standardized when compared to high-strength fibers or reinforcing sheets. To predict the safe limits of the structural response and to account for such a variety of the structural properties a fuzzy randomness in the material parameters is prescribed. Thus generally, analyses require application of the fuzzy finite element method; see, for example, [Hanss and Willner 2000; Möller and Beer 2004]. Alternatively, for the formulation using stochastic finite elements, see, for example, [Dasgupta 2008]. When extending the work on two-layer and symmetric viscoelastic three-layer beam, plate, and shell structures based on the essential definition of an effective equivalently homogenized model, see [Adam et al. 2000; Irschik et al. 2000; Hansen et al. 2005; Heuer 2007], to include either fuzzy pure elastic interface slip (the physical interface exhibits a large variability in the material parameters differing from those of the neighboring layers) or fuzzy core stiffness (the lightweight core material is of similar consistency to a matrix material), we can avoid numerical analysis schemes and analytically work out the effects on the dynamic properties of these fuzzy thin-walled structures. Within the scope of this paper, a simply supported sandwich beam with fuzzy elastic core material parameters is fully analyzed by way of example. It should be mentioned that the solution technique remains applicable to even polygonal

composite plates since a decomposition into two "membranes" has been explored [Heuer et al. 1992; Heuer 2007].

The analysis of the illustrative example, mentioned above, is based on the interval formulation by referring to the set of α-cuts with a (triangular) membership function of the core shear stiffness prescribed. Another interpretation of such an interval number in a closed set is a random variable whose probability density function is unknown but nonzero only in the range of the interval. A recent interval dynamic modal analysis of an uncertain cantilevered shear beam with prescribed lower and upper bounds of the closed set of the shear modulus, [Modares et al. 2010], transformed the deterministic natural frequencies to their inclusive set values. Similarly, membership functions of the more complex undamped natural frequencies of the fuzzy layered beam are defined using fuzzy set theory [Zadeh 1965; Dubois and Prade 1997; Viertl and Hareter 2006], however, avoiding artificial uncertainties. Under time-harmonic excitation the problem is solved in closed form: the dynamic magnification factors and, with light modal structural damping taken into account (thereby neglecting the effects of modal coupling by the light damping forces), the fuzzy phase angles of the modal response are evaluated. Where possible, envelope functions are defined. Such modal response studies of layered beams within the interval formulation allow us to consider the worst case, by either putting the lower limit of the shear modulus of the core material to zero or by taking into account fully delaminated layers. Thus, with the deterministic assigned stiffness known, the lower branch of the membership function can be defined without the requirement of additional data. However, precision engineering may require narrowing the limits of the set of α-cuts of selected natural frequencies: the consequence on the allowable material variability (lower and upper bounds) is addressed. Subsequently, some practically important effects caused by asymmetric uncertainty of the material parameters with respect to the upper branch of the membership function are discussed. Consequently, it can be concluded that the fuzzy set of the core material provides the most appropriate model: starting with the worst-case scenario, all more restrictive bounds can be considered in an inversely taken step of rather simple additional computation without prior knowledge of material data. Since the action of imposed eigenstrains (for example, in the case of nonstationary thermal loads or in the case of piezoelectric strains in smart layers) is considered in the homogenized equivalent fuzzy beam, a fuzzy controller can be designed to annihilate the forced vibrations even under the condition of no additional stress or, relaxed, of no additional stress resultants (see [Ziegler 2005] for "impotent eigenstrains"); the definition is given by [Mura 1991]. To explore the variability of the (light) damping effects related to a fuzzy retardation time of the linear viscous model is left for future investigations since modal decomposition is crucial for the analytical results presented in this paper.

2. Linear viscoelastic layered beams

2.1. *Three-layer beams.* Sandwich structures are commonly defined as three-layer type constructions consisting of two thin face layers of high-strength material attached to a moderately thick core layer of low strength and density [Plantema 1966; Stamm and Witte 1974; Altenbach et al. 2004]. Dynamic response analyses require higher-order theories. A review of the equivalent single layer and layerwise laminate theories is provided by [Reddy 1993]; see also [Irschik 1993; Backström and Nilsson 2005]. The effects of interlayer slip have been discussed for elastic bonding in [Hoischen 1954; Goodman and Popov 1968; Chonan 1982], and for more general interlayer slip laws in [Murakami 1984]. In [Heuer

2004], complete analogies are presented between various models of viscoelastic sandwich structures, with or without interlayer slip, with homogenized single-layer structures of effective parameters.

Figure 1 shows the free-body diagram of a three-layer beam. Introducing the parameter

$$d = \frac{h_1 + h_2}{2} \tag{1}$$

and applying conservation of angular momentum to all three layers gives the relations

$$M_{1,x} - Q_1 - T_1 \frac{h_2}{2} - N_{1,x}d = 0,$$

$$M_{2,x} - Q_2 + (T_1 + T_2)\frac{h_2}{2} = 0, \tag{2}$$

$$M_{3,x} - Q_3 - T_2 \frac{h_2}{2} + N_{3,x}d = 0,$$

where T_1 and T_2 denote the interlaminar shear forces per unit length, and Q_i is the transverse shear force in the i-th layer. The bending moments M_i refer to the individual layer axes. Conservation of momentum in the axial and transverse directions renders

$$N_{1,x} + T_1 = 0, \qquad N_{2,x} - T_1 + T_2 = 0, \qquad N_{3,x} - T_2 = 0, \tag{3}$$

$$\sum_{i=1}^{3} Q_{i,x} + p = \mu \ddot{w}, \qquad \mu = \sum_{i=1}^{3} \rho_i A_i. \tag{4}$$

Summation of (2) yields the global, classical conservation of momentum:

$$M_{,x} - Q = 0, \tag{5}$$

where

$$M = \sum_{i=1}^{3} M_i - (N_1 - N_3)d, \qquad Q = \sum_{i=1}^{3} Q_i. \tag{6}$$
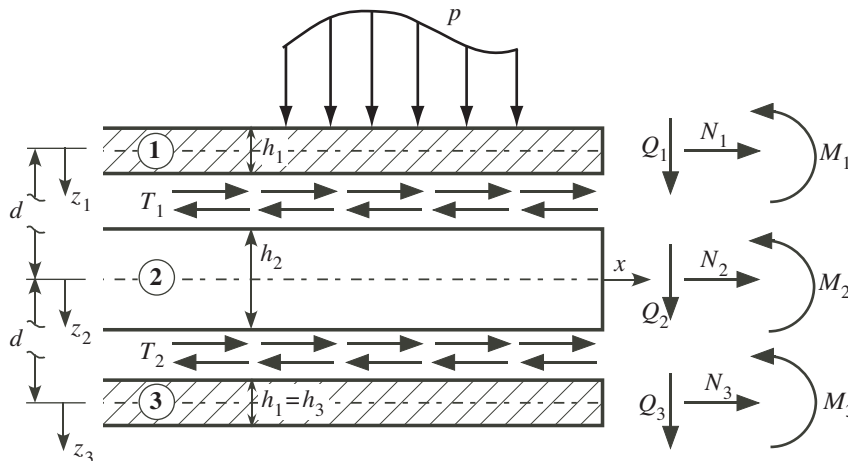

Figure 1. Geometry and stress resultants of a laterally loaded symmetric three-layer beam.

Considering the kinematic assumptions according to the first-order shear-deformation theory applied to each layer, the displacement field in the i-th layer is of the form

$$\begin{bmatrix} u_i \\ w_i \end{bmatrix} = \begin{bmatrix} u_i^{(0)} + z_i \psi_i \\ w \end{bmatrix}, \quad i = 1, 2, 3, \tag{7}$$

where $u_i^{(0)}$ and ψ_i are respectively the axial deformation and cross-sectional rotation of the i-th layer. Applying the linearized strain-displacement relations together with a linear viscoelastic constitutive law, and denoting by ϑ the retardation time of the single parameter viscous model,

$$\begin{pmatrix} \sigma_{xx}(x, z; t) \\ \tau_{xz}(x, z; t) \end{pmatrix} = \begin{pmatrix} E(z_i)[\epsilon_{xx}(x, z_i; t) + \vartheta \dot{\epsilon}_{xx}(x, z_i; t) - \epsilon_{xx}^*(x, z_i; t)] \\ G(z_i)[\gamma_{xz}(x, z_i; t) + \vartheta \dot{\gamma}_{xz}(x, z_i; t)] \end{pmatrix}, \tag{8}$$

the stress resultants of each layer are determined from

$$\begin{bmatrix} N_i \\ M_i \\ Q_i \end{bmatrix} = \begin{bmatrix} D_i & 0 & 0 \\ 0 & B_i & 0 \\ 0 & 0 & S_i \end{bmatrix} \begin{bmatrix} u_{i,x}^{(0)} + \vartheta \dot{u}_{i,x}^{(0)} - e_i^* \\ \psi_{i,x} + \vartheta \dot{\psi}_{i,x} - \kappa_i^* \\ \psi_i + w_{,x} + \vartheta(\dot{\psi}_i + \dot{w}_{,x}) \end{bmatrix}, \quad i = 1, 2, 3, \tag{9}$$

where

$$e_i^* = \frac{1}{A_i} \int_{A_i} \epsilon_{xx}^* \, dA, \qquad \kappa_i^* = \frac{1}{J_i} \int_{A_i} \epsilon_{xx}^* z \, dA, \tag{10}$$

and

$$D_i = E_i A_i = E_i b h_i, \quad B_i = E_i J_i = E_i \frac{b h_i^3}{12}, \quad S_i = \kappa^2 G_i A_i = \kappa^2 G_i b h_i; \tag{11}$$

ϵ_{xx}^*, e_i^*, and κ_i^* are the imposed strain, the imposed mean strain, and the curvature (of thermal nature or resulting from electromagnetic fields applied to smart layers), respectively. The dimensions b and h_i denote a constantly assumed width and the individual layer thickness, respectively; κ^2 stands for a shear factor.

For symmetrically three-layer beams with perfect bonds the following assumptions are made:

(1) The thin faces of high strength material are rigid in shear:

$$\begin{aligned} \psi_1 &= \psi_3 = -w_{,x}, \\ u_1^{(0)} &= u_2^{(0)} + \tfrac{1}{2}(h_1 w_{,x} - h_2 \psi_2), \\ u_3^{(0)} &= u_2^{(0)} - \tfrac{1}{2}(h_1 w_{,x} - h_2 \psi_2). \end{aligned} \tag{12}$$

(2) The individual bending stiffness of the faces are not neglected:

$$B_i = E_i \frac{b h_i^3}{12} \neq 0, \quad i = 1, 3. \tag{13}$$

(3) The bending stiffness of the core (its material is of low density and strength) is neglected:

$$B_2 = 0 \implies M_2 = B_2(\psi_{2,x} + \vartheta \dot{\psi}_{2,x} - \kappa_2^*) = 0. \tag{14}$$

Alternatively, for sandwich beams with viscoelastic interlayer slip [Adam et al. 2000], the classical assumption of all three layers being rigid in shear is made, with the shear traction in the physical interfaces of vanishing thickness being proportional to the displacement jumps with a viscoelastic interface stiffness understood. Hence, contrary to perfectly bonded laminates, elastic interlayer slips, Δu_1 between upper face and core and Δu_2 between lower face and core, are considered. The displacement field in the i-th layer is of the form

$$\begin{bmatrix} u_i \\ w_i \end{bmatrix} = \begin{bmatrix} u_i^{(0)} - z_i w_{,x} \\ w \end{bmatrix}, \quad i = 1, 2, 3, \tag{15}$$

where the axial deformations $u_i^{(0)}$ can be written as

$$u_1^{(0)} = u_2^{(0)} + (dw_{,x} - \Delta u_1), \qquad u_3^{(0)} = u_2^{(0)} - (dw_{,x} - \Delta u_2). \tag{16}$$

The relative horizontal displacement between two layers causes linear viscoelastic shear tractions at the interfaces,

$$T_i = k(\Delta u_i + \vartheta \Delta \dot{u}_i), \tag{17}$$

where the parameter k denotes the slip modulus and

$$\Delta u_1 = dw_{,x} + u_2^{(0)} - u_1^{(0)}, \qquad \Delta u_2 = dw_{,x} - u_2^{(0)} + u_3^{(0)}. \tag{18}$$

Differentiating (17) with respect to the axial coordinate x by taking into account (18) and substituting (3) and (9) renders

$$N_{1,xx} - N_{3,xx} - \frac{k}{D_1}(N_1 - N_3) + 2kd(w_{,xx} + \vartheta \dot{w}_{,xx}) - k(e_1^* - e_3^*) = 0. \tag{19}$$

Furthermore, by means of $(6)_1$ and (9), the difference of the axial forces can be expressed by

$$N_1 - N_3 = -\frac{1}{d}[B_0(w_{,xx} + \vartheta \dot{w}_{,xx} + \kappa^{*(0)}) + M], \qquad B_0 = \sum_{i=1}^{3} B_i, \qquad \kappa^{*(0)} = \frac{1}{B_0} \sum_{i=1}^{3} B_i \kappa_i^*. \tag{20}$$

Inserting (20) into (19) and using (4) and (5) to eliminate the bending moment M, the equation of motion takes on the desired form, namely of an exact homogenization of the viscoelastic composite beam,

$$(w_{,xxxxxx} + \vartheta \dot{w}_{,xxxxxx}) - \lambda^2(w_{,xxxx} + \vartheta \dot{w}_{,xxxx}) + \frac{\mu}{B_0}\ddot{w}_{,xx} - \lambda^2 \frac{\mu}{B_\infty}\ddot{w}$$
$$= -\frac{\lambda^2}{B_\infty}p + \frac{1}{B_0}p_{,xx} + \lambda^2 \kappa^*_{,xx} - \kappa^{*(0)}_{,xxxx}. \tag{21}$$

In (21), p denotes the lateral load per unit of length. In smart layers, we may impose eigenstrains ϵ_x^*, for example, piezoelectric strain, for control purpose. Consequently, the eigencurvature terms have been considered in (21) for the sake of completeness in the case of nonstationary thermal loads and further

investigations of vibration control and smart piezoelectric layers,

$$\kappa^{*(0)} = \frac{1}{B_0} \sum_{i=1}^{3} \frac{12B_i}{h_i^3} \int_{h_i} \epsilon_x^* z_i \, dz,$$

$$\kappa^* = \frac{1}{B_\infty} \left[2D_1 d \left(\frac{1}{h_3} \int_{h_3} \epsilon_x^* \, dz - \frac{1}{h_1} \int_{h_1} \epsilon_x^* \, dz \right) + \sum_{i=1}^{3} \frac{12B_i}{h_i^3} \int_{h_i} \epsilon_x^* z_i \, dz \right].$$

(22)

In addition, with the self-explanatory effective parameters — see Figure 1 and Equation (1) for the notation — we have

$$\mu = 2\rho_1 h_1 + \rho_2 h_2, \qquad D_1 = D_3 = E_1 b h_1, \qquad B_1 = B_3 = E_1 \frac{b h_1^3}{12},$$

$$B_0 = B_1 + B_3, \qquad B_\infty = B_0 + 2D_1 d^2, \qquad B_0/B_\infty = [1 + 3(1 + h_2/h_1)^2]^{-1} \le \tfrac{1}{4}.$$

(23)

The shear coefficient in (21) is either proportional to the core's shear modulus G_2 in the case of perfectly bonded interfaces, i.e., (see [Heuer 2004])

$$\lambda^2 = (\kappa^2 G_2) \frac{2b}{h_2} \frac{B_\infty}{D_1 B_0},$$

(24)

or, for the symmetric three-layer beam with elastic interlayer slip, it becomes proportional to the elastic stiffness k when common to both physical interfaces:

$$\lambda^2 = k \frac{B_\infty}{D_1 B_0}.$$

(25)

Thus, (24) and (25) when substituted in (21) render qualitatively one and the same result with, for example, hard-hinged supports of a single-span beam understood.

For completeness we note also the gross bending moment and the shear force related to the deflection w and its derivatives:

$$M = -B_\infty(w_{,xx} + \vartheta \dot{w}_{,xx} + \kappa^*) + \frac{B_\infty}{\lambda^2} \left(w_{,xxxx} + \vartheta \dot{w}_{,xxxx} + \frac{\mu}{B_0} \ddot{w} + \kappa^{*(0)}_{,xx} \right), \qquad Q = M_{,x}.$$

(26)

Since (26) contains high-order spatial derivatives, numerical calculations require further considerations: the deflection should be partitioned into its quasistatic part and the (modally expanded and truncated) complementary dynamic response. The quasistatic solution is either evaluated in exact closed form or determined by means of the method of influence functions. For details of the partitioning procedure see again [Adam et al. 2000].

2.2. Two-layer beams with interlayer slip.

In the following section the governing homogenized equation (21) given above for symmetric three-layer sandwich beams is modified for asymmetric two-layer elastic beams exhibiting the important defect of viscoelastic interlayer slip. Such a model refers to the practically very important case of a single-span compound bridge consisting, for example, of a steel girder connected (elastically) to the concrete deck; for details see, for example, [Girhammar and Pan 1993]. Figure 2 shows the free body diagram of such a two-layer beam with marked centroids, S_1 and S_2, of the individual
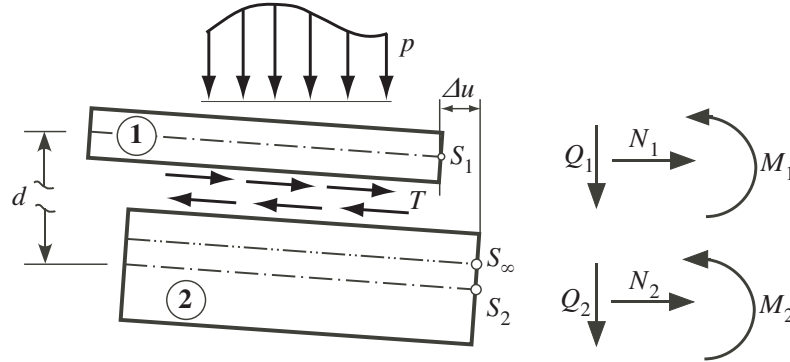
Figure 2. Geometry and stress resultants of a laterally loaded two-layer beam with linear (visco)elastic interlayer slip.

cross-sections, and, as well, global elastic center, S_∞, of the gross composite cross-section. The partial cross-sectional areas are denoted by A_1 and A_2; thus, the lateral coordinate of S_∞ is determined by

$$z_\infty = \frac{D_2}{D_0}d, \qquad D_0 = D_1 + D_2, \qquad D_i = E_i A_i, \quad i = 1, 2. \tag{27}$$

Neglecting the effect of rotational inertia, the conservation of the angular momentum about the lateral y-axis yields the classical relationship

$$M_{,x} = Q, \qquad M = M_1 + M_2 - N_1 d, \tag{28}$$

where M is the gross bending moment (see again Figure 2). In the case of no external axial forces the equilibrium condition becomes simply

$$N_1 + N_2 = 0. \tag{29}$$

Conservation of momentum in the axial direction, when applied in the free body diagram to both individual layers, gives

$$N_{1,x} = -T_1, \qquad N_{2,x} = T_2, \tag{30}$$

and the continuity condition of the interfacial shear force T renders

$$T = T_1 = T_2. \tag{31}$$

T is the shear force per unit of length continuously transmitted through the interface between the two layers.

Due to the interlayer slip to be considered between the layers (see again Figure 2), the Bernoulli–Euler hypothesis is not applicable for the cross-section as a whole. However, the assumption that plane sections remain plane after deformation is still valid for either layer, respectively. Considering geometrically linearized conditions for both, the lateral and the axial deformations yield

$$w = w_1 = w_2, \qquad \Delta u = u_2^{(0)} - u_1^{(0)} + dw_{,x}. \tag{32}$$

The relative horizontal displacement between the two layers causes shear tractions at the interface, where again a linear viscoelastic relation is assigned, see Equation (17),

$$T = k(\Delta u + \vartheta \, \Delta \dot{u}), \tag{33}$$

where k is the elastic slip modulus [Hoischen 1954; Girhammar and Pan 1993]. Differentiating (33) with respect to the axial coordinate x and substituting (29), (30), and $(32)_2$ lead to

$$N_{1,xx} - kN_1 \left(\frac{1}{D_1} + \frac{1}{D_2} \right) + kd(w_{,xx} + \vartheta \dot{w}_{,xx}) - k(e_1^* - e_3^*) = 0. \tag{34}$$

The axial force is expressed by means of $(28)_2$ with cross-sectional resultants of the constitutive relations inserted,

$$N_1 = -\frac{1}{d}[B_0(w_{,xx} + \vartheta \dot{w}_{,xx} + \kappa^{*(0)}) + M], \qquad B_0 = \sum_{i=1}^{2} B_i, \qquad \kappa^{*(0)} = \frac{1}{B_0} \sum_{i=1}^{2} B_i \kappa_i^*. \tag{35}$$

Finally, the exactly homogenized equation of motion results qualitatively in the form of (21); however the definitions (22) and (23) must be properly adapted to account for the beam with two layers — see (38) below — and essentially, this new parameter is to be substituted in (21):

$$\lambda^2 \to \lambda_2^2 = k \left(\frac{1}{D_1} + \frac{1}{D_2} + \frac{d^2}{B_0} \right). \tag{36}$$

3. The linear elastic layered beam with fuzzy stiffness

To explore the effects of fuzzy elastic material parameters and to make a modal analysis possible, the viscous effects in the homogenized equation (21) are neglected. Fuzziness of the light damping forces can be based on (21) but needs a separate investigation and thus is left for future consideration. Since no attempts are made within this paper to control the vibrations and no thermally driven vibrations are analyzed, the terms in (21) referring to imposed strains are no longer taken into account. Consequently, (21) takes on the simpler form, see also [Heuer et al. 2003], however, generalized by its explicit dependence on a fuzzy parameter,

$$w_{,xxxxxx} - \lambda^2 w_{,xxxx} + \frac{\mu}{B_0} \ddot{w}_{,xx} - \lambda^2 \frac{\mu}{B_\infty} \ddot{w} = -\frac{\lambda^2}{B_\infty} p + \frac{1}{B_0} p_{,xx}. \tag{37}$$

Thus, the shear coefficient in (37) is considered proportional to the fuzzy core shear modulus G_2 either in the case of perfectly bonded interfaces of a symmetric three-layered beam, when (24) applies, or in the practically even more important case of a two-layer beam with fuzzy elastic interface slip, with (36) substituted. In addition, the latter case requires a change in the expression of the limiting flexural stiffness given in (23); it becomes

$$B_\infty = B_0 + \frac{D_1 D_2}{D_1 + D_2} d^2. \tag{38}$$

Fuzziness in (37) is thus simply introduced by inserting, for example, a triangular membership function either for the core's shear modulus in (24) or for the physical interface stiffness in (36). Consequently the analysis of the illustrative example of a simply supported beam, as mentioned above, is based on

the interval representation, that is, the set of α-cuts. In both cases, the analysis can be based on the structurally relevant worst-case assumption of the lower bound, that is, of a maximum width of the fuzzy set by the cut at $\alpha = 0$, to include either the classical sandwich beam of vanishing core stiffness or no bond at the interface between two layers (fully delaminated), thereby avoiding the use of unreliable data on the variability of the material parameters.

Subsequently, we consider the fuzzy core stiffness of the three-layer beam in some detail.

3.1. *Modal analysis of the elastic three-layer beam, hard-hinged support.*

For the single-span beam with hard hinged support, the homogeneous equation (37) yields the simple orthonormalized mode shapes of free vibrations:

$$\phi_n(x) = A_n \sin \beta_{1n} x, \qquad \beta_{1n} = n\pi/l, \qquad A_n = \left[\frac{\mu l}{2} \left(\frac{\lambda^2}{B_\infty} + \frac{\beta_{1n}^2}{B_0} \right) \right]^{-\frac{1}{2}}, \tag{39}$$

with λ^2 in that case substituted from (24).

The corresponding (undamped) circular natural frequencies are

$$\left(\frac{\omega_n(\alpha)}{\omega_{n\infty}} \right)^2 = \frac{(B_0/B_\infty) + \gamma_{2,n} k_2(\alpha)}{1 + \gamma_{2,n} k_2(\alpha)}, \tag{40}$$

when referred to

$$\omega_{n\infty}^2 = \beta_{1n}^4 B_\infty / \mu \tag{41}$$

with the fuzzy nondimensional elastic material constant of the core (typically made of a low-density, low-cost matrix material) substituted:

$$k_2(\alpha) = [\kappa_2^2 G_2(\alpha)]/(\kappa_2^2 G_2)_0. \tag{42}$$

The assigned constants are referred to the mean shear rigidity of the core (see again Figure 1):

$$(\kappa_2^2 G_2)_0 = [\kappa_2^2 G_2(\alpha = 1)], \qquad \gamma_{2,n} = (\kappa_2^2 G_2)_0 \frac{2b}{\beta_{1n}^2 D_1 h_2}. \tag{43}$$

In this relation, the shear coefficient κ_2 must not be numerically specified. The monotonic behavior of the eigenfrequencies of positive-definite mechanical systems was proven by [Mullen and Modares 2009]; see (40).

The normalizing factor in (39) becomes fuzzy too. The largest level set at $\alpha = 0$ may include the worst case of a core with vanishing stiffness (or a delaminated composite). Due to the lack of reliable core material data, such a choice is recommended for the analysis. Putting constraints on the fuzziness of the dynamic parameters inversely results in a more narrow specification of the allowable largest parameter interval of the core material (see Section 4.2 for details):

$$\Delta k_{2,0} = \Delta(\kappa_2^2 G_2)_0/(\kappa_2^2 G_2)_0 \leq 1. \tag{44}$$

The linear functions plotted in Figure 3, both intersecting at $\alpha = 1$, arise when, for the sake of simplicity, an isosceles triangular membership function is assumed:

$$\min k_2(\alpha) = 1 - \Delta k_{2,0}(1 - \alpha), \quad \max k_2(\alpha) = 1 + \Delta k_{2,0}(1 - \alpha). \tag{45}$$
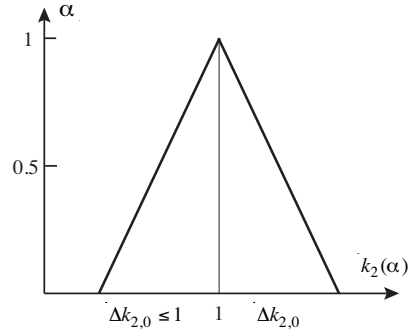
Figure 3. Three-layer beam: uncertainty of the core shear stiffness, under the assumption of an isosceles membership function.

These two equations are substituted into (40), when considering the fuzzy set in Figure 3, giving respectively the lower and upper branches of the membership functions of the natural frequencies.

Since light modal damping is commonly considered in the steady-state response of the modal oscillators to time-harmonic forcing (thus neglecting any modal coupling resulting from the viscous deformations) the dynamic magnification factor and the phase angle for every assigned forcing frequency $(\omega/\omega_{n\infty})$ become uncertain functions with the parameter α. For the sake of simplicity, constant light modal damping coefficients are assumed, $\xi = \xi_n \ll 1$, $n = 1, 2, 3, \ldots$. The dynamic magnification factor can be delineated in the proper form (see, for example, [Ziegler 1998]),

$$\chi_n = \left\{ 1 - 2(1 - 2\xi^2)\left(\frac{\omega}{\omega_n(\alpha)}\right)^2 + \left(\frac{\omega}{\omega_n(\alpha)}\right)^4 \right\}^{-1/2}. \tag{46}$$

However, clearly any set of light modal damping coefficients $\zeta_1 \le \zeta_2 \le \cdots \le \zeta_n \le \zeta_n + 1$ can be accounted for in (46) when estimated according to a design class of the beam.

The phase angle is given for every assigned nondimensional forcing frequency $(\omega/\omega_{n\infty})$ as a function of α, as follows (see [Ziegler 1998]):

$$\tan \phi_n = 2\xi \frac{\omega}{\omega_{n\infty}} \left[\omega_n(\alpha)/\omega_{n\infty} - \left(\frac{\omega}{\omega_{n\infty}}\right)^2 \frac{\omega_{n\infty}}{\omega_n(\alpha)} \right]^{-1}. \tag{47}$$

When (40) is substituted in (46) and (47), these equations become explicitly dependent on the fuzzy variable, (42); in particular, we get

$$\tan \phi_n = 2\xi \frac{\omega}{\omega_{n\infty}} \left[\sqrt{\frac{(B_0/B_\infty) + \gamma_{2,n}k_2(\alpha)}{1 + \gamma_{2,n}k_2(\alpha)}} - \left(\frac{\omega}{\omega_{n\infty}}\right)^2 \sqrt{\frac{1 + \gamma_{2,n}k_2(\alpha)}{(B_0/B_\infty) + \gamma_{2,n}k_2(\alpha)}} \right]^{-1}. \tag{48}$$

Again, considering the fuzzy set in Figure 3, (45) is respectively substituted in (40), (46), and (47), to yield explicitly the envelope functions of the fuzzy dynamic magnification factor and of the fuzzy phase angle as well.

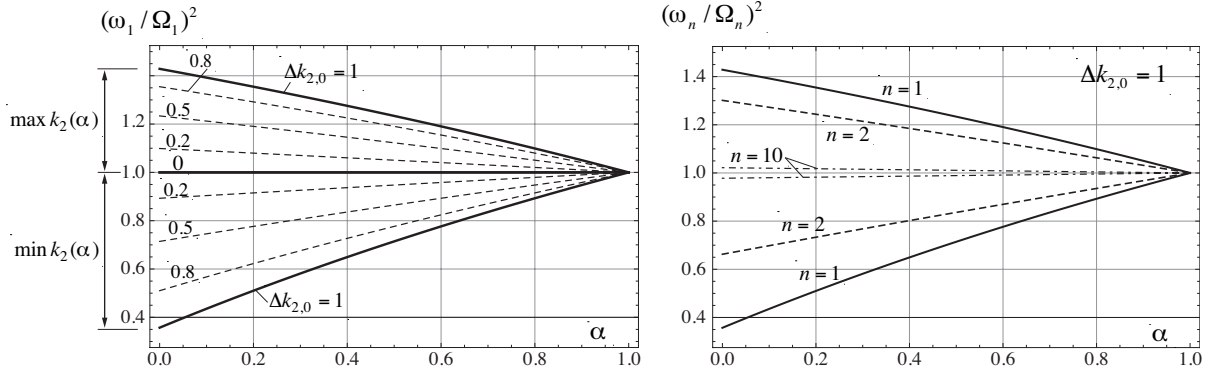## 4. Numerical results: fuzzy sandwich beam

Figure 4. Three-layer beam. Left: uncertainty of the fundamental frequency ratio for various widths of fuzzy core stiffness. Right: uncertainty of the natural frequencies relative to the natural frequencies of the same order for $\alpha = 1$, for largest width of fuzzy core stiffness.

4.1. *Isosceles uncertainty.* (See Figure 3.) Two basic parameters are assigned for the elastic three-layer beam of Section 3, with constant light modal damping $\xi_n = \xi = 0.05$ understood throughout:

$$B_0/B_\infty = 0.1 < \tfrac{1}{4}, \qquad \gamma_{2,n=1} = 0.25. \tag{49}$$

In the course of numerical analyses, it was found to be most illustrative to refer the frequencies to the deterministic natural circular frequencies at $\alpha = 1$. With $\Omega_n = \omega_n(1)$, Equation (40) reads

$$\left(\frac{\Omega_n}{\omega_{n\infty}} \right)^2 = \frac{(B_0/B_\infty) + \gamma_{2,n}}{1 + \gamma_{2,n}}. \tag{50}$$

This equation is evaluated first to explore the influence of the width of uncertainty on the fundamental frequency; see Figure 4, left. Considering the largest structurally possible uncertainty in the isosceles membership function of Figure 3, the intervals of uncertainty of the natural frequencies of higher order, $n \leq 10$, are depicted in Figure 4, right, relative to their values at $\alpha = 1$, (50). These membership functions become more informative when referred to the assigned fundamental frequency Ω_1: see Figure 5.

The envelope function of the dynamic magnification factor of the basic mode, that is, putting $n = 1$ in (46), is plotted in Figure 6 varying the width of uncertainty according to Figure 3 in the $\alpha = 0.2$ cut; the forcing frequency is referred to Ω_1. Considering maximum uncertainty (the worst-case scenario) but taking into account the whole range of parameter α gives the envelope surfaces in Figure 7. The $\alpha = 0.2$ cut yields the envelope functions of the first three modes as plotted in Figure 8.

Complementary to Figure 6, the variations of the range of uncertainty of the phase angle of the basic mode with the width of uncertainty according to Figure 3 in the $\alpha = 0.2$ cut are shown in Figure 9; the forcing frequency is again referred to the assigned fundamental frequency Ω_1. To complement Figure 8, the ranges of uncertainty of the first three phase angles are plotted in Figure 10.

4.2. *Constraints affecting the uncertain natural frequencies.* Constraints on the design uncertainty, say on the uncertainty of the core shear stiffness of the three-layered beam, are often based on limiting the maximum allowable variability of the natural frequencies in a given frequency window, see [Massa

et al. 2008]. By inspection of Figure 5 it is easily recognized that such constraints can be expressed in assigning fractions of the frequency intervals (where the maximum structural width of uncertainty has been considered) that are maximum at $\alpha = 0$. Thus, a 50% reduction of that maximum result interval is chosen and, we refer to the assigned fundamental frequency Ω_1,

$$\frac{\omega_{n,\max}(\alpha) - \omega_{n,\min}(\alpha)}{\Omega_1} \leq [\omega_{n,\max}(\alpha = 0) - \omega_{n,\min}(\alpha = 0)]/2\Omega_1, \quad n = 1, 2, \ldots, 5. \tag{51}$$

The frequency window includes and is limited by the mode number 5 for some practical reasons. Hence, the resulting α-cut is determined by solving the equation

$$\frac{\omega_{5,\max}(\alpha) - \omega_{5,\min}(\alpha)}{\Omega_1} = [\omega_{5,\max}(\alpha = 0) - \omega_{5,\min}(\alpha = 0)]/2\Omega_1. \tag{52}$$
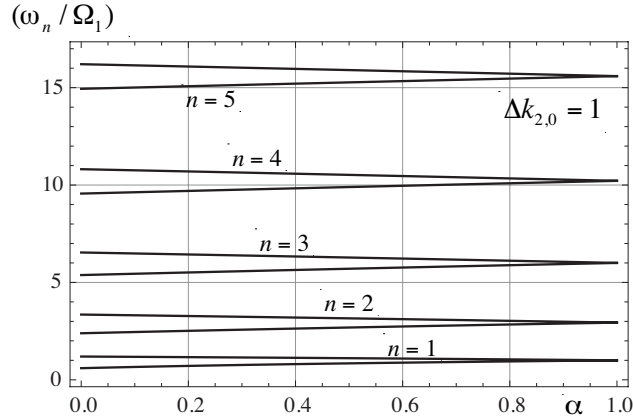
with (45) substituted.

Figure 5. Three-layer beam: uncertainty of the first five natural frequencies relative to the fundamental frequency for $\alpha = 1$; increasing fuzziness observed.
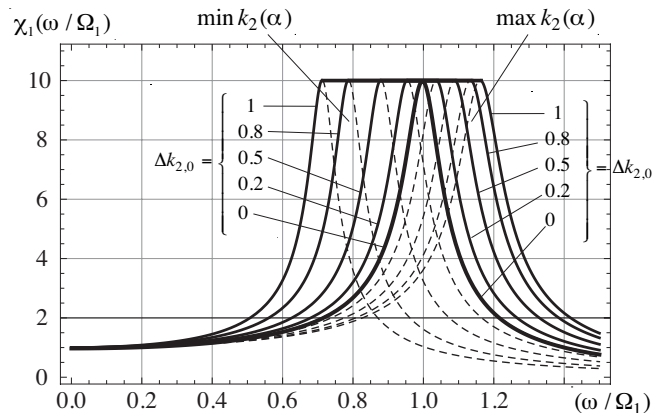

Figure 6. Three-layer beam: uncertainty of the first dynamic magnification factor for the single $\alpha = 0.2$ cut, varying the maximum interval of the fuzzy core stiffness.
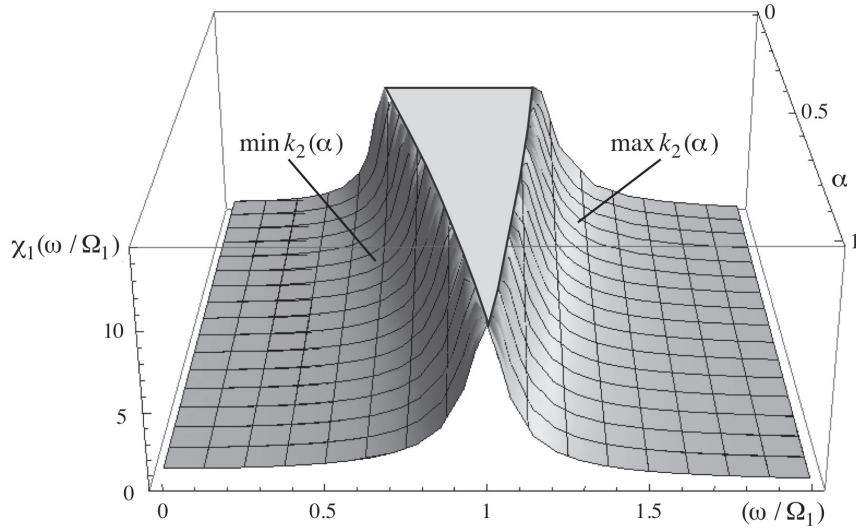
Figure 7. Three-layer beam: uncertainty surfaces of the dynamic magnification factor of the fundamental mode, light modal damping assigned.
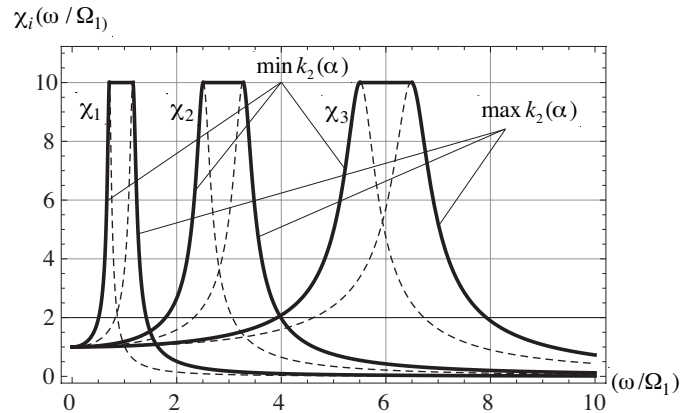

Figure 8. Three-layer beam: envelopes of the first three dynamic magnification factors for the single $\alpha = 0.2$ cut; increasing fuzziness observed.

Alternatively, it might fit the needs of precision design to limit the uncertainty relatively to the sure natural frequency of the same order of the mode (see Figure 6). We thus refer to Ω_n:

$$\frac{\omega_{n,\max}(\alpha) - \omega_{n,\min}(\alpha)}{\Omega_n} \leq [\omega_{n,\max}(\alpha = 0) - \omega_{n,\min}(\alpha = 0)]/2\Omega_n, \qquad n = 1, 2, 3, \ldots. \qquad (53)$$

In that case of relative uncertainty, the fundamental frequency interval gives the maximum tolerable uncertainty in the core shear stiffness if we solve for α the equation

$$\frac{\omega_{1,\max}(\alpha) - \omega_{1,\min}(\alpha)}{\Omega_1} = [\omega_{1,\max}(\alpha = 0) - \omega_{1,\min}(\alpha = 0)]/2\Omega_1. \qquad (54)$$
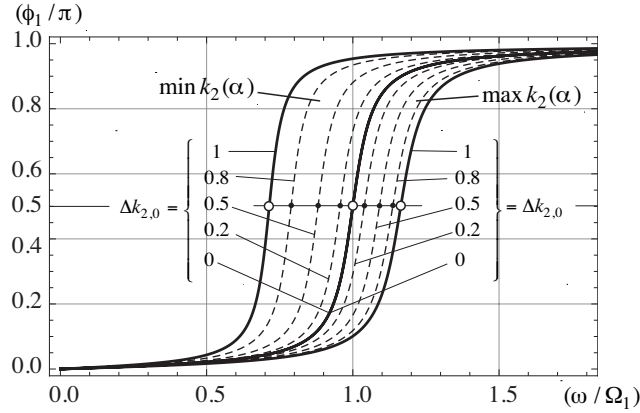
Figure 9. Three-layer beam: uncertainty of the first phase angle for the single $\alpha = 0.2$ cut, maximum interval of the fuzzy core stiffness varied; complement of Figure 6.
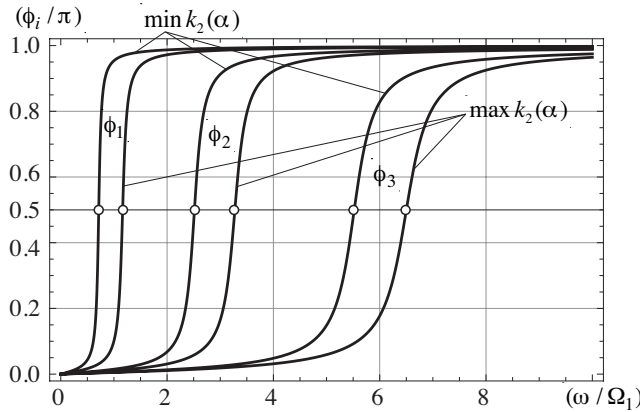

Figure 10. Three-layer beam: the envelopes of three phase angles for the single $\alpha = 0.2$ cut: complementing Figure 8; increasing fuzziness observed.

Since the largest width of uncertainty of the core shear stiffness is considered in Figures 5 and 6, the constraints allow the definition of the allowable α-cuts by solving either (52), to render $\alpha(n = 5) \doteq 0.50$, or alternatively, (54), to yield the less stringent condition at $\alpha(n = 1) \doteq 0.44$. The maximum bases of the allowable isosceles membership functions of the uncertainty of the core shear stiffness in the more precise designs are plotted in Figure 11.

4.3. *Some effects of nonsymmetric uncertainty.* We consider the extreme (worst) case of $\Delta k_{2,0} = 1$, that is, (45) when generalized reduces to $\min k_2(\alpha) = \alpha$, $\max k_2(\alpha) = 1 + \Lambda(1 - \alpha)$, where $\Lambda > 1$ renders the maximum core stiffness enlarged in Figure 3. The extremes of upper and lower bounds of

the neighboring natural frequencies thus are obtained for the $\alpha = 0$ cut, (42) is properly adapted,

$$\max\left(\frac{\omega_n}{\omega_{n,\infty}}\right)^2_{\alpha=0} = \frac{(B_0/B_\infty) + \gamma_{2,1}(1+\Lambda)/n^2}{1 + \gamma_{2,1}(1+\Lambda)/n^2}, \qquad \omega^2_{n,\infty} = n^4\beta^4_{1,1}(B_\infty/\mu),$$

$$\min\left(\frac{\omega_{n+1}}{\omega_{n+1,\infty}}\right)^2_{\alpha=0} = (B_0/B_\infty) \le 1/4, \qquad \omega^2_{n+1,\infty} = (n+1)^4\beta^4_{1,1}(B_\infty/\mu).$$

(55)

A first effect of asymmetric uncertainty in an ensemble of fuzzy beams is observed by putting $\max \omega^2_n = \min \omega^2_{n+1}$ rendering the coefficient of asymmetry,

$$\Lambda = \Lambda(n) = \frac{(B_0/B_\infty)}{\gamma_{2,1}} \frac{n^2\left[\left(\frac{n+1}{n}\right)^4 - 1\right]}{\left[1 - (B_0/B_\infty)\left(\frac{n+1}{n}\right)^4\right]} - 1 > 0.$$

(56)

The parameters assumed in Section 4.1, $(B_0/B_\infty) = 0.1$ and $\gamma_{2,1} = 0.25$, exclude a solution of (56) for $n = 1$ and yield the coefficients of asymmetry: $\Lambda(n = 2) = 12.16$, $\Lambda(n = 3) = 10.37$, $\Lambda(n = 4) = 11.20$, $\Lambda(n = 5) = 12.54$, $\Lambda(n = 6) = 14.07$, $\Lambda(n = 7) = 15.68$, ... Equation (56), when virtually considered for continuous order n exhibits a singularity at $n = 1.28$. It is moved to $n = 1$ for the smaller ratio of the flexural stiffness $B_0/B_\infty = 1/16 \approx 0.063$. Since max $B_0/B_\infty = 0.25$, reported in (23), the singularity at $n = 2$ is still possible for the flexural stiffness ratio $B_0/B_\infty = 0.198$, but no effect on higher modes is observed for $n \ge 3$. Consequently, for these two ratios of the flexural stiffness, a violation of the assumption of a triangular membership function of uncertainty is observed and consequently the solution for the coefficient of asymmetry becomes invalid.

5. Conclusions

For symmetric three-layer slender beam, thin plate and thin shell structures, and for a two-layer composite an exact homogenization exists, which is worked out in detail for the hard hinged supported
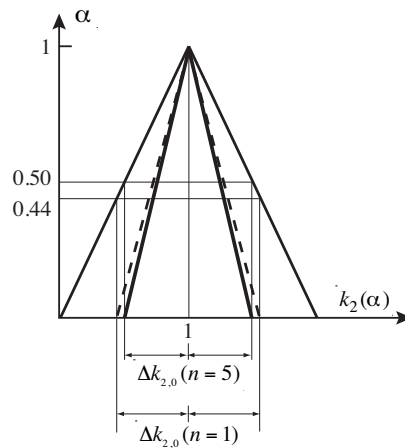

Figure 11. Three-layer beam: reduction of the intervals of uncertainty under the conditions of 50% constraints of either uncertainty of the first five natural frequency intervals, $n = 5$, or of the relative natural frequency intervals, $n = 1$.

beam. Consequently, interval mathematics becomes applicable to the solution of the sixth-order resulting homogenized equation with a fuzzy parameter either of a fuzzy core shear stiffness and/or fuzzy stiffness of the physical interface between layers to define the intervals of the dynamic response. A detailed study of a three-layer simply hard supported beam with fuzzy core of an isosceles membership function is performed, giving lower and upper bounds of the natural frequencies and providing deep insights into level-sets of the dynamic magnification factors and of the phase angles in forced time harmonic vibrations. Fuzziness of dynamic parameters like natural frequencies, dynamic magnification factors, and modal phase angles increases with frequency as expressed by the modal order. This efficient analysis can be performed without reliable knowledge of the uncertainty of the material parameters by considering structurally inherent worst-case scenarios. It yields exact and robust results, preserves the problem's physics, and obtaining bounds does not require expensive stochastic procedures such as Monte Carlo simulations. Putting constraints on the variability, say of the natural frequencies, leads to the maximum tolerable uncertainty in the core shear stiffness and/or that of the stiffness of the physical interface. Effects within an ensemble of such beams, for example, overlapping of the intervals of their natural frequencies, however, under the assumption of a nonsymmetric triangular membership function of the core stiffness uncertainty, are shown to be limited to the first and second modes. Either fuzzy modal superposition or fuzzy control of vibrations can be based on the results of this paper. However, since the normalizing factor of the orthogonal mode shapes turns out to be fuzzy too, the complexity of modal superposition is increased. Thus, modal superposition of forced vibrations becomes fuzzy in both time records and amplitude response. Superposition of the modal maxima by considering the square root of sum of squares yields the (approximate) total displacement response even in this more complex case. The analysis of a fuzzy retardation time of viscoelastic layers (such a fuzzy parameter is included in the homogenized equation) leads to coupled modal equations even in a Ritz–Galerkin approximation and is out of the scope of this paper. It is left to future investigations.

6. Acknowledgement

References

[Adam et al. 2000] C. Adam, R. Heuer, A. Raue, and F. Ziegler, "Thermally induced vibrations of composite beams with interlayer slip", *J. Therm. Stresses* **23**:8 (2000), 747–772.

[Altenbach et al. 2004] H. Altenbach, J. Altenbach, and W. Kissing, *Mechanics of composite structural elements*, Springer, Berlin, 2004.

[Backström and Nilsson 2005] D. Backström and A. Nilsson, "Flexural vibrations of a three-layer sandwich beam: using ordinary fourth order beam theory in combination with frequency dependent parameters to predict the flexural dynamics of a sandwich beam", pp. 567–576 in *Sandwich structures 7: advancing with sandwich structures and materials* (Aalborg, 2005), edited by O. T. Thomsen et al., Springer, Dordrecht, 2005. Part 5.

[Chonan 1982] S. Chonan, "Vibration and stability of sandwich beams with elastic bonding", *J. Sound Vib.* **85**:4 (1982), 525–537.

[Dasgupta 2008] G. Dasgupta, "Stochastic shape functions and stochastic strain-displacement matrix for a stochastic finite element stiffness matrix", *Acta Mech.* **195** (2008), 379–395.

[Dubois and Prade 1997] D. Dubois and H. Prade, "The three semantics of fuzzy sets", *Fuzzy Sets Syst.* **90**:2 (1997), 141–150.

[Girhammar and Pan 1993] U. A. Girhammar and D. Pan, "Dynamic analysis of composite members with interlayer slip", *Int. J. Solids Struct.* **30**:6 (1993), 797–823.

[Goodman and Popov 1968] J. R. Goodman and E. P. Popov, "Layered beam systems with interlayer slip", *J. Struct. Div.* (*ASCE*) **94** (1968), 2535–2547.

[Hansen et al. 2005] J. S. Hansen, G. Kennedy, and S. F. M. de Almeida, "A homogenization based theory for laminated and sandwich beams", pp. 221–230 in *Sandwich structures 7: advancing with sandwich structures and materials* (Aalborg, 2005), edited by O. T. Thomsen et al., Springer, Dordrecht, 2005. Part 3.

[Hanss and Willner 2000] M. Hanss and K. Willner, "A fuzzy arithmetical approach to the solution of finite element problems with uncertain parameters", *Mech. Res. Commun.* **27**:3 (2000), 257–272.

[Heuer 2004] R. Heuer, "Equivalence of the analyses of sandwich beams with or without interlayer slip", *Mech. Adv. Mater. Struct.* **11**:4–5 (2004), 425–432.

[Heuer 2007] R. Heuer, "Equivalences in the analysis of thermally induced vibrations of sandwich structures", *J. Therm. Stresses* **30**:6 (2007), 605–621.

[Heuer et al. 1992] R. Heuer, H. Irschik, and F. Ziegler, "Thermally forced vibrations of moderately thick polygonal plates", *J. Therm. Stresses* **15**:2 (1992), 203–210.

[Heuer et al. 2003] R. Heuer, C. Adam, and F. Ziegler, "Sandwich panels with interlayer slip subjected to thermal loads", *J. Therm. Stresses* **26**:11–12 (2003), 1185–1192.

[Hoischen 1954] A. Hoischen, "Verbundträger mit elastischer und unterbrochener Verdübelung", *Bauingenieur* **29** (1954), 241–244.

[Irschik 1993] H. Irschik, "On vibrations of layered beams and plates", *Z. Angew. Math. Mech.* **73**:4-5 (1993), T34–T45.

[Irschik et al. 2000] H. Irschik, R. Heuer, and F. Ziegler, "Statics and dynamics of simply supported polygonal Reissner–Mindlin plates by analogy", *Arch. Appl. Mech.* **70**:4 (2000), 231–244.

[Massa et al. 2008] F. Massa, K. Ruffin, T. Tison, and B. Lallemand, "A complete method for efficient fuzzy modal analysis", *J. Sound Vib.* **309**:1–2 (2008), 63–85.

[Modares et al. 2010] M. Modares, I. Olabarri, and R. L. Mullen, "Reliable dynamic analysis of an uncertain shear beam", pp. 173–185 in *REC 2010: proceedings of the 4th International Workshop on Reliable Engineering Computing* (Singapore, 2010), edited by M. Beer et al., Research Publication Service, Singapore, 2010.

[Möller and Beer 2004] B. Möller and M. Beer, *Fuzzy randomness*, Springer, Berlin, 2004.

[Mullen and Modares 2009] R. L. Mullen and M. Modares, "Free vibration of structures with interval uncertainty", in *Proceedings of the 9th ASCE Speciality Conference on Probabilistic Mechanics and Structural Reliability* (*PMC2004*) (Albuquerque, NM, 2004), Curran, Red Hook, NY, 2009.

[Mura 1991] T. Mura, *Micromechanics of defects in solids*, 2nd ed., Kluwer, Dordrecht, 1991.

[Murakami 1984] H. Murakami, "A laminated beam theory with interlayer slip", *J. Appl. Mech.* (*ASME*) **51**:3 (1984), 551–559.

[Plantema 1966] F. J. Plantema, *Sandwich construction*, Wiley, New York, 1966.

[Reddy 1993] J. N. Reddy, "An evaluation of equivalent-single-layer and layerwise theories of composite laminates", *Compos. Struct.* **25** (1993), 21–35.

[Stamm and Witte 1974] K. Stamm and H. Witte, *Sandwichkonstruktionen*, Springer, New York, 1974.

[Viertl and Hareter 2006] R. Viertl and D. Hareter, *Beschreibung und Analyse unscharfer Information: Statistische Methoden für unscharfe Daten*, Springer, Berlin, 2006.

[Zadeh 1965] L. A. Zadeh, "Fuzzy sets", *Information and control* **8** (1965), 338–353.

[Ziegler 1998] F. Ziegler, *Mechanics of solids and fluids*, 2nd corrected ed., Springer, New York, 1998.

[Ziegler 2005] F. Ziegler, "Computational aspects of structural shape control", *Comp. Struct.* **83**:15–16 (2005), 1191–1204.

RUDOLF HEUER: rudolf.heuer@tuwien.ac.at
Civil Engineering Department, Vienna University of Technology, Karlsplatz 13/E2063, A-1040 Vienna, Austria

FRANZ ZIEGLER: franz.ziegler@tuwien.ac.at
Civil Engineering Department, Vienna University of Technology, Karlsplatz 13/E2063, A-1040 Vienna, Austria

PATH-INDEPENDENT H-INTEGRAL FOR INTERFACE CORNERS UNDER THERMAL LOADINGS

CHYANBIN HWU, TAI-LIANG KUO AND CHUN-CHIH HUANG

It is well known that the path-independent H-integral is an appropriate tool for calculating the mixed mode stress intensity factors for the interface corners between dissimilar elastic materials. To extend the applicability of the H-integral from the mechanical loading condition to the thermal loading condition, a modified H-integral is proposed in this paper. This modified H-integral possesses an extra domain integral which needs the input of temperature field. Moreover, this domain integral contains singular functions that come from the strain components of the auxiliary system, and a special treatment should be made for the accurate computation of stress intensity factors. The near-tip solutions and auxiliary solutions of displacements, stresses, and temperature required in the calculation of H-integral are all provided in this paper. The validity and versatility of the proposed approach are then shown by carrying out several numerical examples such as cracks under mixed-mode thermal loadings, interface cracks/corners under uniform heat flow or uniform temperature change, and an electronic package, in which the chip has a heat generation rate, placed at a constant temperature ambiance.

1. Introduction

Many engineering objects, for example electronic packages, engines of power vehicles, solar panels, and so on, are operated in thermal environments. Temperature changes, heat flux on the object surface, and heat generation in the interior can deform the object and induce stress when restrictions on deformation are imposed, such as a clamped boundary condition or a perfect-bonded condition along an interface between dissimilar materials. Interface corners commonly appear in these engineering objects and failures initiate from these critical regions due to discontinuities of geometry and material properties. Hence, methods of fracture analysis for estimating the potential of failure and the mode of fracture of interface corners in elastic materials subjected to thermal loading are of great importance. Orders of stress singularity and stress intensity factors are two commonly used parameters when we perform fracture analyses within the category of linear elastic fracture mechanics.

This paper provides an accurate, efficient, and systematic solution technique to calculate these two parameters for interface corners between dissimilar elastic materials subjected to thermal loading.

Studies of fracture analysis of interface cracks subjected to thermal loadings include [Erdogan 1965; Brown and Erdogan 1968; Hwu 1990; 1992; Ikeda and Sun 2001; Banks-Sills and Dolev 2004; Nagai et al. 2007]. Relatively few studies have dealt with interface corners; they include [Munz and Yang 1992; 1993; Banks-Sills and Ishbir 2004; Hwu and Lee 2004; Nomura et al. 2009]. To understand the

mechanical behavior of anisotropic elastic materials under thermal environments, Stroh formalism [Stroh 1958; Ting 1996; Hwu 2010] for two-dimensional linear anisotropic thermoelasticity has been employed in many works [Clements 1973; Hwu 1990]. By this formalism, analytical closed-form solutions for the orders of heat flux/stress singularity and near-tip solutions of multimaterial anisotropic wedges under thermal loadings have been obtained in [Hwu and Lee 2004]. To understand the fracture behavior of interface corners, a unified definition of stress intensity factors for connecting cracks, corners, interface cracks, and interface corners was proposed in [Hwu and Kuo 2007]. In that work, in order to avoid the oscillatory singular problems of interface corners a path-independent H-integral [Bueckner 1973; Stern 1976] was suggested to calculate the stress intensity factors. Based on these works, in this paper the H-integral is further modified to be suitable for the thermal loading condition.

The modified H-integral contains an additional domain integral that is not included in the H-integral for pure mechanical loading. The integrand in this domain integral contains singular functions that come from the strain components in the auxiliary system. Considerable numerical error will be induced if we perform this domain integral via normal numerical methods, for example, Simpson's rule and Gaussian quadrature. To deal with this problem, the domain integral is separated into two parts. One is the near-tip part, to be integrated analytically, and the other is the adjacent part, to be calculated numerically. A similar approach has been proposed by [Banks-Sills and Ishbir 2004; Nomura et al. 2009].

Several numerical examples are analyzed for the purpose of verification: a center crack under mixed-mode thermal loading, a center interface crack under uniform heat flow, edge interface cracks under uniform temperature change, and interface corners under uniform temperature change. In addition, an example about electronic packages is analyzed to show the feasibility and practicability of the modified H-integral.

2. Thermoelastic analysis of interface corners

In a fixed rectangular coordinate system $x_i, i = 1, 2, 3$, let u_i, σ_{ij}, ϵ_{ij}, T, and h_i be, respectively, the displacement, stress, strain, temperature, and heat flux. The heat conduction, energy equation, strain-displacement relation, constitutive law, and equilibrium equations for the uncoupled steady state thermoelastic problems can be written as [Nowacki 1962]

$$h_i = -k_{ij}T_{,j}, \ h_{i,i} = 0, \ \epsilon_{ij} = \tfrac{1}{2}(u_{i,j} + u_{j,i}), \ \sigma_{ij} = C_{ijks}\epsilon_{ks} - \beta_{ij}T, \ \sigma_{ij,j} = 0, \quad i, j, k, s = 1, 2, 3, \quad (1)$$

where repeated indices imply summation, a comma stands for differentiation, and C_{ijks}, k_{ij}, and β_{ij} are, respectively, the elastic constants, heat conduction coefficients, and thermal moduli. C_{ijks} are assumed to be fully symmetric, that is, $C_{ijks} = C_{jiks} = C_{ijsk} = C_{ksij}$, and are required to be positive definite due to the positiveness of the strain energy. β_{ij} and k_{ij} are also assumed to be symmetric, that is, $\beta_{ij} = \beta_{ji}$ and $k_{ij} = k_{ji}$.

Consider an interface corner between two dissimilar anisotropic elastic materials (Figure 1) in which a local polar coordinate system (r, θ) is specified at the corner tip. Assume a perfect bond along the interface. The displacement, traction, temperature, and heat flux continuity across the interface $\theta = 0$ can be written as [Hwu and Lee 2004]

$$\boldsymbol{u}_1(0) = \boldsymbol{u}_2(0), \quad \boldsymbol{\phi}_1(0) = \boldsymbol{\phi}_2(0), \quad T_1(0) = T_2(0), \quad h_1^*(0) = h_2^*(0), \tag{2a}$$
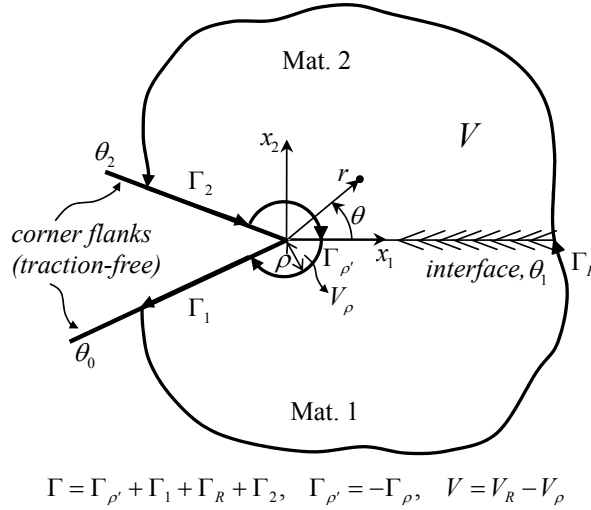
$$\Gamma = \Gamma_{\rho'} + \Gamma_1 + \Gamma_R + \Gamma_2, \quad \Gamma_{\rho'} = -\Gamma_\rho, \quad V = V_R - V_\rho$$

Figure 1. A closed contour Γ and its domain V for the H-integral of an interface corner.

where $\boldsymbol{u}$ and $\boldsymbol{\phi}$ are, respectively, the displacement vector and stress function vector; h^* is the heat flux in the direction normal to the interface; the subscripts 1 and 2 stand for the values related to materials 1 and 2; and the argument 0 denotes the value on the interface. The corner flanks are both assumed to be traction-free which can be expressed by the stress function as

$$\boldsymbol{\phi}_1(\theta_0) = \boldsymbol{\phi}_2(\theta_2) = \boldsymbol{0}. \tag{2b}$$

Four different thermal conditions on the corner flanks are considered in this paper:

$$\begin{aligned}
\text{isothermal-isothermal:} \quad & T_{1,r}(\theta_0) = T_{2,r}(\theta_2) = 0, \\
\text{insulated-insulated:} \quad & h_1^*(\theta_0) = h_2^*(\theta_2) = 0, \\
\text{insulated-isothermal:} \quad & h_1^*(\theta_0) = T_{2,r}(\theta_2) = 0, \\
\text{isothermal-insulated:} \quad & T_{1,r}(\theta_0) = h_2^*(\theta_2) = 0.
\end{aligned} \tag{2c}$$

Note that the components of stress function vector, ϕ_i, $i = 1, 2, 3$, are related to the stresses σ_{ij} and the surface traction t_i by

$$\sigma_{i1} = -\phi_{i,2}, \quad \sigma_{i2} = \phi_{i,1}, \quad t_i = \sigma_{ij} n_j = \frac{d\phi_i}{ds}, \tag{3}$$

where n_i is the normal of the surface and s is the tangential.

3. Near-tip solutions

The near-tip solutions satisfying all the basic equations (1) and boundary conditions (2) were obtained in our previous study [Hwu and Lee 2004] as

$$\boldsymbol{v}(r, \theta) = -\delta(1-\delta) r^{-1-\delta} \boldsymbol{\Gamma}(\theta) \boldsymbol{v}_0, \quad \boldsymbol{w}(r, \theta) = r^{1-\delta} \{ -\delta(1-\delta) \boldsymbol{F}(\theta) \boldsymbol{v}_0 + \boldsymbol{E}(\theta) \boldsymbol{w}_0 \}, \tag{4a}$$

where

$$\boldsymbol{w}(r,\theta) = \begin{Bmatrix} \boldsymbol{u}(r,\theta) \\ \boldsymbol{\phi}(r,\theta) \end{Bmatrix}, \quad \boldsymbol{v}(r,\theta) = \begin{Bmatrix} T_{,r}(r,\theta) \\ h^*(r,\theta) \end{Bmatrix}, \quad \boldsymbol{w}_0 = \begin{Bmatrix} \boldsymbol{u}_0 \\ \boldsymbol{\phi}_0 \end{Bmatrix}, \quad \boldsymbol{v}_0 = \begin{Bmatrix} T_{0,r} \\ h_0^* \end{Bmatrix}. \tag{4b}$$

In (4), δ is the singular order and $\boldsymbol{\Gamma}(\theta)$, $\boldsymbol{F}(\theta)$, and $\boldsymbol{E}(\theta)$ are the matrices related to the material properties and corner angles and have different expressions for materials 1 and 2 (given in the Appendix). The eigenvectors associated with the singular order δ are $\boldsymbol{w}_0$ and $\boldsymbol{v}_0$, and can be determined by

$$\boldsymbol{K}_e^{(3)}\boldsymbol{u}_0 = \boldsymbol{0}, \quad \boldsymbol{\phi}_0 = \boldsymbol{0}, \quad \boldsymbol{v}_0 = \boldsymbol{0}, \tag{5a}$$

where $\boldsymbol{K}_e^{(3)}$ is a 3×3 submatrix of $\boldsymbol{K}_e$ defined by

$$\boldsymbol{K}_e = \begin{bmatrix} \boldsymbol{K}_e^{(1)} & \boldsymbol{K}_e^{(2)} \\ \boldsymbol{K}_e^{(3)} & \boldsymbol{K}_e^{(4)} \end{bmatrix}, \quad \boldsymbol{K}_e = \hat{\boldsymbol{N}}_2^{1-\delta}(\theta_2,\theta_1)\hat{\boldsymbol{N}}_1^{1-\delta}(\theta_1,\theta_0), \tag{5b}$$

and $\hat{\boldsymbol{N}}$ is the key matrix introduced in [Hwu et al. 2003]. In (4) and hereafter, the subscript k denoting the values related to the k-th material is dropped to lighten the notation.

In (5a), the equalities $\boldsymbol{K}_e^{(3)}\boldsymbol{u}_0 = \boldsymbol{0}$ and $\boldsymbol{\phi}_0 = \boldsymbol{0}$ come from the traction-free boundary condition set in (2b), and $\boldsymbol{v}_0 = \boldsymbol{0}$ comes from the requirement that the temperature is not allowed to be singular near the corner tip. From (4a) we see that if $\boldsymbol{v}_0 \neq \boldsymbol{0}$, both temperature and heat flux will be singular if the stresses are singular. If the stresses are singular and the strain energy cannot be unbounded, only the singular orders located in $0 < \mathrm{Re}(\delta) < 1$ are considered in this paper.

With $\boldsymbol{v}_0 = \boldsymbol{0}$, the near-tip solutions (4a) become

$$\boldsymbol{v}(r,\theta) = \boldsymbol{0}, \quad \boldsymbol{w}(r,\theta) = r^{1-\delta}\boldsymbol{E}(\theta)\boldsymbol{w}_0, \tag{6}$$

which are the solutions without considering thermal effects. Since the singular order δ may be distinct or repeated, real or complex, combination of all the possible solutions associated with the most critical singular order whose real part δ_R is maximal leads, as in [Hwu and Kuo 2007; Hwu and Ikeda 2008], from (6) to

$$T(r,\theta) = 0, \qquad h_i(r,\theta) = 0,$$
$$\boldsymbol{u}(r,\theta) = \frac{1}{\sqrt{2\pi}}r^{1-\delta_R}\boldsymbol{V}(\theta)\langle(1-\delta_R+i\epsilon_\alpha)^{-1}(r/l)^{i\epsilon_\alpha}\rangle\boldsymbol{\Lambda}^{-1}\boldsymbol{k}, \tag{7}$$
$$\boldsymbol{\phi}(r,\theta) = \frac{1}{\sqrt{2\pi}}r^{1-\delta_R}\boldsymbol{\Lambda}(\theta)\langle(1-\delta_R+i\epsilon_\alpha)^{-1}(r/l)^{i\epsilon_\alpha}\rangle\boldsymbol{\Lambda}^{-1}\boldsymbol{k}.$$

In (7) the angular brackets $\langle\ \rangle$ stand for a diagonal matrix in which each component is varied according to the subscript α, for example, $\langle z_\alpha\rangle = \mathrm{diag}[z_1, z_2, z_3]$; δ_R and ϵ_α are, respectively, the real and imaginary parts of the *most critical singular order* δ_c determined by (5a) with $0 < \mathrm{Re}(\delta) < 1$; l is a length parameter which may be chosen arbitrarily as long as it is held fixed when specimens of a given material pair are compared; $\boldsymbol{V}(\theta)$ and $\boldsymbol{\Lambda}(\theta)$ are *eigenfunction matrices* related to $\boldsymbol{E}(\theta)\boldsymbol{w}_0$; $\boldsymbol{\Lambda} = \boldsymbol{\Lambda}(0)$, detailed expressions for which can be found in [Hwu and Kuo 2007; Kuo and Hwu 2010]; and $\boldsymbol{k}$ is a vector containing different

modes of *stress intensity factors* and is defined by

$$
\boldsymbol{k} = \left\{ \begin{array}{c} K_{\mathrm{II}} \\ K_{\mathrm{I}} \\ K_{\mathrm{III}} \end{array} \right\} = \lim_{\substack{r \to 0 \\ \theta = 0}} \sqrt{2\pi} r^{\delta_R} \boldsymbol{\Lambda} \langle (r/l)^{-i\epsilon_\alpha} \rangle \boldsymbol{\Lambda}^{-1} \left\{ \begin{array}{c} \sigma_{r\theta} \\ \sigma_{\theta\theta} \\ \sigma_{3\theta} \end{array} \right\}.
\tag{8}
$$

4. Path-independent H-integral for thermoelastic problems

According to the definition of stress intensity factors (8), to calculate their values we need to know the stresses near the tips of interface corners. However, due to the singular and possibly oscillatory behaviors of the near-tip stresses, it is not easy to get convergent values for the stress intensity factors directly from the definition (8). To overcome this problem, several path-independent integrals have been proposed for crack problems such as the J-integral [Rice 1968], the L-integral [Choi and Earmme 1992], the M-integral [Im and Kim 2000], and the H-integral [Bueckner 1973; Stern 1976; Sinclair et al. 1984; Chen 1985]. For corner problems that are usually in the status of mixed-mode intensity, the H-integral was suggested by [Hwu and Kuo 2007] for two-dimensional interface corners, and modified by [Kuo and Hwu 2010] for three-dimensional interface corners, which are valid for pure mechanical loading conditions. For interface corners under thermal loadings, the H-integral was modified by [Banks-Sills and Ishbir 2004; Nomura et al. 2009]. However, some important details that should be clarified were not interpreted in their works, such as the near-tip temperature field and the reason why the thermal effect disappears in near-tip solutions of displacements and stresses. Their H-integral cannot calculate the mixed-mode stress intensity factors via one expression. To have a complete picture of the H-integral for thermoelastic problems, in this section we first prove the path-independence property for the proposed modified H-integral, then provide a special treatment for the extra domain integral added in the modified H-integral.

4.1. *Path-independence of the modified H-integral.* If an elastic body is subjected to two different thermal loading systems (indicated by a hat or its absence), the constitutive laws shown in (1) give

$$
\int_V \hat{\sigma}_{ij} \epsilon_{ij} \, \mathrm{d}V = \int_V (C_{ijks} \hat{\epsilon}_{ks} \epsilon_{ij} - \beta_{ij} \hat{T} \epsilon_{ij}) \, \mathrm{d}V,
\tag{9a}
$$

$$
\int_V \sigma_{ij} \hat{\epsilon}_{ij} \, \mathrm{d}V = \int_V (C_{ijks} \epsilon_{ks} \hat{\epsilon}_{ij} - \beta_{ij} T \hat{\epsilon}_{ij}) \, \mathrm{d}V.
\tag{9b}
$$

Subtracting (9a) from (9b) and using the symmetry property of elastic constants, we get a conservative integral for thermoelastic problems,

$$
\int_V (\hat{\sigma}_{ij} \epsilon_{ij} - \sigma_{ij} \hat{\epsilon}_{ij}) \, \mathrm{d}V + \int_V \beta_{ij} (\hat{T} \epsilon_{ij} - T \hat{\epsilon}_{ij}) \, \mathrm{d}V = 0.
\tag{10a}
$$

If one prefers the use of stresses instead of strains, (10a) can be rewritten as

$$
\int_V (\hat{\sigma}_{ij} \epsilon_{ij} - \sigma_{ij} \hat{\epsilon}_{ij}) \, \mathrm{d}V + \int_V \alpha_{ij} (\hat{T} \sigma_{ij} - T \hat{\sigma}_{ij}) \, \mathrm{d}V = 0,
\tag{10b}
$$

where α_{ij} are the thermal expansion coefficients which are related to the thermal moduli β_{ij} by

$$
\beta_{ij} = C_{ijks} \alpha_{ks}.
\tag{11}
$$

By using the strain-displacement relation $(1)_3$, the symmetric property of stresses, Cauchy's formula $t_i = \sigma_{ij} n_j$, and Gauss's divergence theorem $\int_V (\ldots)_{,j} \, dV = \int_\Gamma (\ldots) n_j \, d\Gamma$, the first domain integral in (10a) can be further reduced to a path integral, and consequently the following relation can be obtained

$$\int_\Gamma (\hat{t}_i u_i - t_i \hat{u}_i) \, d\Gamma + \int_V \beta_{ij} (\hat{T} \epsilon_{ij} - T \hat{\epsilon}_{ij}) \, dV = 0. \tag{12}$$

In (12), Γ can be any positively oriented closed contour in a simply connected region. If Γ is selected to be $-\Gamma_\rho + \Gamma_1 + \Gamma_R + \Gamma_2$ and $V \, (= V_R - V_\rho)$ is the domain enclosed by Γ (see Figure 1), due to the traction-free condition on the corner flanks, that is, $t_i = \hat{t}_i = 0$ along Γ_1 and Γ_2, (12) now leads to

$$\int_{\Gamma_\rho} (\hat{t}_i u_i - t_i \hat{u}_i) \, d\Gamma + \int_{V_\rho} \beta_{ij} (\hat{T} \epsilon_{ij} - T \hat{\epsilon}_{ij}) \, dV = \int_{\Gamma_R} (\hat{t}_i u_i - t_i \hat{u}_i) \, d\Gamma + \int_{V_R} \beta_{ij} (\hat{T} \epsilon_{ij} - T \hat{\epsilon}_{ij}) \, dV, \tag{13}$$

which means that the following modified H-integral is path-independent:

$$H = \int_{\Gamma_R} (\boldsymbol{u}^T \hat{\boldsymbol{t}} - \hat{\boldsymbol{u}}^T \boldsymbol{t}) \, d\Gamma + \int_{V_R} \beta_{ij} (\hat{T} \epsilon_{ij} - T \hat{\epsilon}_{ij}) \, dV. \tag{14}$$

The superscript T denotes a transpose; Γ_R is a counterclockwise integral path with arbitrary shape which emanates from the lower corner flank ($\theta = \theta_0$) and terminates on the upper flank ($\theta = \theta_2$); $\boldsymbol{u}$ and $\boldsymbol{t}$ are the displacement and traction vectors of the actual system, which can be obtained through any method, for example, finite element analyses, boundary element analyses, or even experimental measurement, while $\hat{\boldsymbol{u}}$ and $\hat{\boldsymbol{t}}$ are those of the auxiliary system with singular order $2 - \delta$.

In order to use the path-independence property of the modified H-integral to calculate the stress intensity factors, we can first select Γ_R to be a circular path which passes the region dominated by the singular field. Along the circular path, we have

$$\boldsymbol{t} = \boldsymbol{\phi}_{,\theta}/r \quad \text{and} \quad d\Gamma = r \, d\theta;$$

and hence (14) can be rewritten as

$$H = \int_{\theta_0}^{\theta_2} (\boldsymbol{u}^T \hat{\boldsymbol{\phi}}_{,\theta} - \hat{\boldsymbol{u}}^T \boldsymbol{\phi}_{,\theta}) \, d\theta + \int_V \beta_{ij} (\hat{T} \epsilon_{ij} - T \hat{\epsilon}_{ij}) \, dV, \tag{15}$$

in which $\boldsymbol{u}$, $\boldsymbol{\phi}$, ϵ_{ij}, and T are the near-tip solutions given in (7), and $\hat{\boldsymbol{u}}$, $\hat{\boldsymbol{\phi}}$, $\hat{\epsilon}_{ij}$, and $\hat{T}$ are the auxiliary solutions, which can be obtained by

$$\hat{\boldsymbol{u}}(r, \theta) = r^{\delta_R - 1} \hat{\boldsymbol{V}}(\theta) \langle r^{-i\epsilon_\alpha} \rangle \hat{\boldsymbol{c}}, \quad \hat{\boldsymbol{\phi}}(r, \theta) = r^{\delta_R - 1} \hat{\boldsymbol{\Lambda}}(\theta) \langle r^{-i\epsilon_\alpha} \rangle \hat{\boldsymbol{c}}, \quad \hat{T}(r, \theta) = 0. \tag{16}$$

Since the temperature fields in both the near-tip and auxiliary solutions are zero, the H-integral passing through the singular field will be exactly the same as that for the pure mechanical loading problems. Since the relation between the stress intensity factors $\boldsymbol{k}$ and the H-integral is obtained from the results of the H-integral passing through the singular field, it can now be written by referring to the relation obtained for the pure mechanical loading problem [Hwu and Kuo 2007], that is,

$$\boldsymbol{k} = \sqrt{2\pi} \boldsymbol{\Lambda} \langle (1 - \delta_R + i\epsilon_\alpha) l^{i\epsilon_\alpha} \rangle \boldsymbol{H}^{*-1} \boldsymbol{h}, \tag{17a}$$

where

$$\boldsymbol{H}^* = \int_{\theta_0}^{\theta_2} [\hat{\boldsymbol{\Lambda}}'^T(\theta)\boldsymbol{V}(\theta) - \hat{\boldsymbol{V}}^T(\theta)\boldsymbol{\Lambda}'^{(\theta)}]\mathrm{d}\theta, \quad \boldsymbol{h} = \begin{Bmatrix} H_1 \\ H_2 \\ H_3 \end{Bmatrix}. \tag{17b}$$

In (17b), the prime $'$ denotes differentiation with respect to θ; H_i, $i = 1, 2, 3$, are the values of H calculated from (14) using the auxiliary solutions given in (16) with $\hat{c}_i = 1$ and $\hat{c}_j = 0$, $i \neq j$. Since the path-independence property has been proved through (13), the integral path calculating H_i can be chosen arbitrarily and $\boldsymbol{u}$, $\boldsymbol{t}$, and T of the actual system can be provided through any method, such as finite element analysis.

4.2. *Special treatment of the extra domain integral.*

The difference between the modified H-integral (14) and the H-integral for the pure mechanical loading problem is the additional domain integral in the second term of (14). By selecting the auxiliary temperature field $\hat{T} = 0$, the domain integral becomes

$$-\int_{V_R} T\beta_{ij}\hat{\epsilon}_{ij}\,\mathrm{d}V. \tag{18}$$

From the auxiliary solutions given in (16), we see that the auxiliary strain $\hat{\epsilon}_{ij}$ has a strong singularity as $r^{\delta-2}$, where $0 < \mathrm{Re}(\delta) < 1$. This term will cause tremendous numerical error and hence should be treated with special attention. Banks-Sills and Ishbir [2004] proposed that the domain V_R can be separated into two parts: one is close to the corner tip, V_ρ in Figure 1, which can be integrated analytically, and the other adjacent part, $V = V_R - V_\rho$ in Figure 1, can be calculated numerically. However, in their study no analytical solution has been provided for the near-tip solution of temperature field, and hence no further analytical solution was provided for the integration. In [Nomura et al. 2009], this domain integral was calculated analytically for a circular integral path in which the circular sector domain is divided into several small elements whose temperature is assumed to be constant in each element. In the present study, the analytical integration is further simplified by using the near-tip solution of the temperature field.

Based upon the analytical solutions given in (4a) we see that the near-tip solution of the temperature field can be obtained by integrating (4a)$_1$ with respect to r, which will lead to

$$T(r, \theta) = (1 - \delta_h)r^{-\delta_h}\gamma(\theta)c_1 + c_2, \tag{19}$$

where $\gamma(\theta)$ is a function related to $\Gamma(\theta)$ of (4a); c_1 and c_2 are the coefficients to be determined via the actual temperature field which can come from analytical solution or numerical analysis. Here, δ_h is the *singular order of heat flux*, which is located in the region of $-1 < \mathrm{Re}(\delta_h) < 0$ and will not induce a singularity in temperature or stresses.

According to the thermal conditions on the corner flanks (2c), the singular orders of heat flux have been obtained from the following relations [Hwu and Lee 2004]:

$$\begin{aligned} &\text{isothermal-isothermal:} \quad K_T^{(2)} = 0, &\qquad &\text{insulated-insulated:} \quad K_T^{(3)} = 0, \\ &\text{insulated-isothermal:} \quad K_T^{(1)} = 0, &\qquad &\text{isothermal-insulated:} \quad K_T^{(4)} = 0, \end{aligned} \tag{20}$$

where $K_T^{(i)}$, $i = 1, 2, 3, 4$, are the components of $\boldsymbol{K}_T$, which is a 2×2 matrix related to the heat conduction coefficients and corner angles.

To evaluate (18) analytically for the part of V_ρ, the auxiliary strain $\hat{\epsilon}_{ij}$ obtained from (16) can be expressed as

$$
\begin{aligned}
&\hat{\epsilon}_{11} = \boldsymbol{i}_1^T \hat{\boldsymbol{u}}_{,1}, && \hat{\epsilon}_{22} = \boldsymbol{i}_2^T \hat{\boldsymbol{u}}_{,2}, && \hat{\epsilon}_{33} = 0, \\
&\hat{\epsilon}_{12} = \tfrac{1}{2}(\boldsymbol{i}_2^T \hat{\boldsymbol{u}}_{,1} + \boldsymbol{i}_1^T \hat{\boldsymbol{u}}_{,2}), && \hat{\epsilon}_{23} = \tfrac{1}{2}\boldsymbol{i}_3^T \hat{\boldsymbol{u}}_{,2}, && \hat{\epsilon}_{13} = \tfrac{1}{2}\boldsymbol{i}_3^T \hat{\boldsymbol{u}}_{,1},
\end{aligned}
\tag{21a}
$$

where

$$
\hat{\boldsymbol{u}}_{,1} = r^{\delta_R-2}(\cos\theta \boldsymbol{e}_1 - \sin\theta \boldsymbol{e}_2), \quad \hat{\boldsymbol{u}}_{,2} = r^{\delta_R-2}(\sin\theta \boldsymbol{e}_1 + \cos\theta \boldsymbol{e}_2),
$$
$$
\boldsymbol{i}_1^T = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}, \quad \boldsymbol{i}_2^T = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}, \quad \boldsymbol{i}_3^T = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix},
\tag{21b}
$$

and

$$
\boldsymbol{e}_1 = \hat{\boldsymbol{V}}(\theta)\langle(\delta_R - 1 - i\epsilon_\alpha)r^{-i\epsilon_\alpha}\rangle\hat{\boldsymbol{c}}, \quad \boldsymbol{e}_2 = \hat{\boldsymbol{V}}'(\theta)\langle r^{-i\epsilon_\alpha}\rangle\hat{\boldsymbol{c}}.
\tag{21c}
$$

Substituting (19) and (21) into (18), and letting $dV = r\,dr\,d\theta$ for V_ρ, we get

$$
\int_{V_R} T\beta_{ij}\hat{\epsilon}_{ij}\,dV = I_\rho + \int_{V_R - V_\rho} T\beta_{ij}\hat{\epsilon}_{ij}\,dV,
\tag{22a}
$$

where I_ρ is the integral that has been integrated analytically with respect to r in the near-tip domain V_ρ whose result is

$$
I_\rho = \rho^{\delta_R}\int_{\theta_0}^{\theta_2} e_0(\rho, \theta)\,d\theta,
\tag{22b}
$$

in which

$$
e_0(\rho, \theta) = \boldsymbol{\beta}_1^{*T}(\theta)\hat{\boldsymbol{V}}(\theta)\langle(\delta_R - 1 - i\epsilon_\alpha)g_\alpha(\rho, \theta)\rangle\hat{\boldsymbol{c}} + \boldsymbol{\beta}_2^{*T}(\theta)\hat{\boldsymbol{V}}'(\theta)\langle g_\alpha(\rho, \theta)\rangle\hat{\boldsymbol{c}},
\tag{22c}
$$

and

$$
g_\alpha(\rho, \theta) = \left\{\frac{c_1(1 - \delta_h)\rho^{-\delta_h}}{\delta_R - \delta_h - i\epsilon_\alpha}\gamma(\theta) + \frac{c_2}{\delta_R - i\epsilon_\alpha}\right\}\rho^{-i\epsilon_\alpha},
\tag{22d}
$$

$$
\boldsymbol{\beta}_1^*(\theta) = \cos\theta\,\boldsymbol{\beta}_1 + \sin\theta\,\boldsymbol{\beta}_2, \quad \boldsymbol{\beta}_2^*(\theta) = -\sin\theta\,\boldsymbol{\beta}_1 + \cos\theta\,\boldsymbol{\beta}_2, \quad \boldsymbol{\beta}_1 = \begin{Bmatrix} \beta_{11} \\ \beta_{12} \\ \beta_{13} \end{Bmatrix}, \ \boldsymbol{\beta}_2 = \begin{Bmatrix} \beta_{21} \\ \beta_{22} \\ \beta_{23} \end{Bmatrix}.
\tag{22e}
$$

In (22a)–(22c), ρ is the radius of a small circle ahead of the corner tip. Since $0 < \delta_R < 1$ and $-1 < \mathrm{Re}(\delta_h) < 0$, from the results of (22a)–(22e) we see that the singular problem of (18) has been solved through the analytical integration I_ρ. To have a proper choice of ρ, the convergent test about ρ should be done in a numerical calculation, which will be illustrated by an example shown in the next section. Since the singular problem in the near-tip domain occurs from approaching zero distance, that is, $r \to 0$, whether to obtain the analytical integration with respect to θ is not the main concern of our study. Therefore, due to the complexity of $\gamma(\theta)$ and $\hat{\boldsymbol{V}}(\theta)$ in $e_0(\rho, \theta)$, the analytical integration of I_ρ shown in (22b) is only for the variable r not including θ.

5. Numerical examples

To provide a stable and efficient computing approach for the general mixed-mode stress intensity factors under thermal loadings, the path-independent H-integral proposed in the literature [Hwu and Kuo 2007] has been modified by adding an extra domain integral as shown in (14). This integral is applicable to

cracks, interface cracks, corners, and interface corners, and the materials containing the cracks/corners can be any kinds of anisotropic materials, including degenerate materials such as isotropic materials. The stress intensity factors calculated through the H-integral include the pure mode and mixed mode, and also the factors associated with lower singular orders [Kuo and Hwu 2010]. The main feature of the present approach is that one unified H-integral can deal with several different kinds of thermal problems which are generally discussed separately. Thus, to provide an enhanced comparison, several different kinds of materials and cracks/corners considered in the literature are shown in the following examples. In order to show that the modified H-integral is path-independent, the data shown below will be presented by stress intensity factors with different radius of integral path.

All the examples presented in this section consider the state of plane strain. The physical quantities of the actual system, u, t, and T, needed for the calculation of the H-integral in (14) are obtained from the commercial finite element software ANSYS. A two-dimensional 4-node thermal element PLANE55 is adopted to perform the thermal analyses, and then the temperature results are read into a two-dimensional 4-node structural solid element PLANE42 and treated as the body force to proceed with the stress analyses. Since the numerical output will depend on element meshes, the convergent test needs to be done before performing all the following examples. In our numerical implementation, the number of elements for the most refined mesh is 29574 for modeling the electronic package, and 7484 for modeling the interface crack. For convenience, the H-integral path doesn't need to pass through nodal points, while the integration points can be arbitrary points whose numerical data are produced by extrapolating the results of their surrounding nodal points [Lancaster and Salkauskas 1981; Nomura et al. 2009] and then integrated by Gaussian quadrature. Note that although the path-independence property of the H-integral has been proved theoretically in Section 4, when using the H-integral to calculate the stress intensity factors we still have to avoid taking the numerical results overly close to the corner tip due to the incorrect stress information in the neighborhood of the corner tip provided by finite element analysis. Also note that although the solution techniques proposed in this paper are applicable to the most general two-dimensional problems, such as the generalized plane strain and generalized plane stress, due to the limitation of two-dimensional elements provided by the finite element software ANSYS only the plane strain condition is considered in our examples.

5.1. *Comparison with existing solutions.* In order to prove the path-independence property numerically, to verify the correctness of the stress intensity factors calculated by the proposed H-integral and to show the versatility of the present unified approach, six different kinds of cracks/corners under thermal loadings are implemented and compared with the existing solutions presented in the literature. They are:

Case 1: A center crack under mode I thermal loading (Figure 2, left).

Case 2: A center crack under mixed-mode thermal loading (Figure 2, right).

Case 3: A center interface crack under uniform heat flow (Figure 3, left).

Case 4: Edge interface cracks under uniform temperature change (Figure 3, right).

Case 5: Edge interface corners under uniform temperature change (Figure 4, left).

Case 6: An interface corner under uniform temperature change (Figure 4, right).

Point A in Figures 2–4 stands for the corner or crack tip we are concerned with in these problems. The geometry, loading, and material properties of these problems are collected in Table 1. The results of
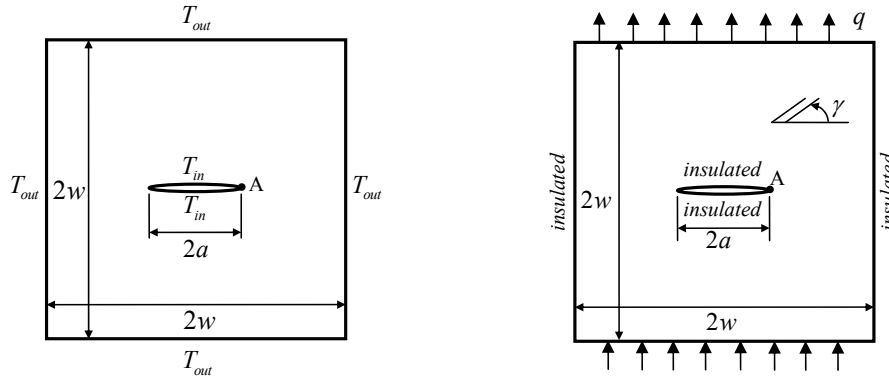
Figure 2. A center crack in a square plate under mode I (left, Case 1) and mixed-mode (right, Case 2) thermal loadings.
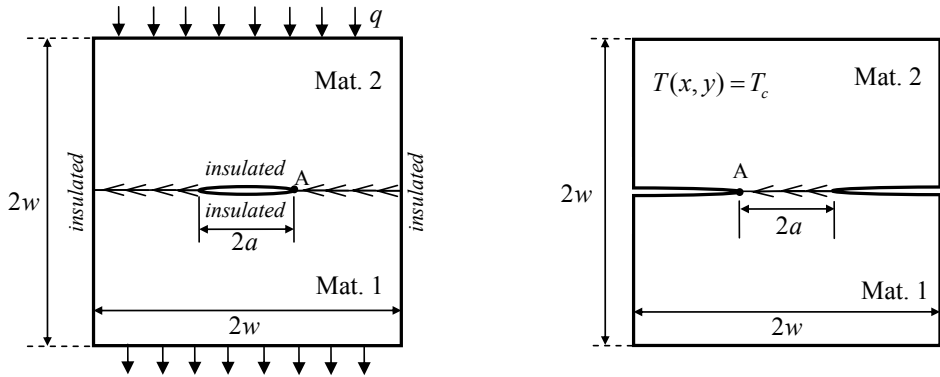

Figure 3. A center interface crack (left, Case 3) and edge interface cracks (right, Case 4) in a square bimaterial plate.
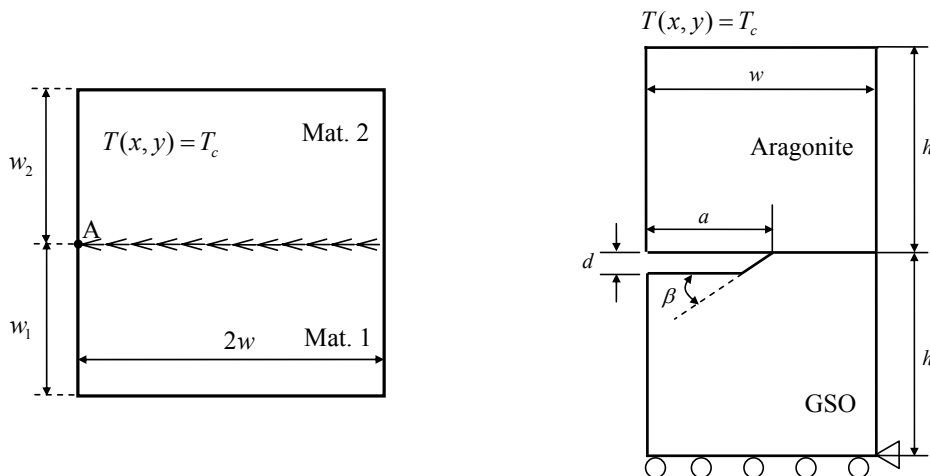

Figure 4. Edge interface corners (left, Case 5) and an interface corner (right, Case 6) in a bimaterial plate.

Case 1 (Figure 2, left) Center crack under mode I thermal loading, isotropic material			
$a = 4\,\text{mm}$	$w = 10a$	$T_{\text{in}} = 0\,°\text{C}$	$T_{\text{out}} = 100\,°\text{C}$
$E = 1\,\text{MPa}$	$\nu = 0.3$	$\alpha = 10^{-4}\,°\text{C}^{-1}$	$k = 1\,\text{W/m}\,°\text{C}$

Case 2 (Figure 2, right) Center crack under mixed-modes thermal loading, orthotropic material ($\gamma = 0°$), anisotropic material ($\gamma = 30°$)			
$a = 1\,\text{mm}$	$w = 30a$	$q = 10^3\,\text{W/m}^2$	
$E_{11} = 144.23\,\text{GPa}$	$E_{22} = E_{33} = 9.65\,\text{GPa}$	$G_{12} = 4.14\,\text{GPa}$	$\alpha_{11} = 0.88 \times 10^{-6}\,°\text{C}^{-1}$
$\nu_{12} = \nu_{13} = 0.301$	$\nu_{23} = 0.49$	$G_{13} = 4.14\,\text{GPa}$	$\alpha_{22} = 31 \times 10^{-6}\,°\text{C}^{-1}$
$k_{11} = 4.48\,\text{W/m}\,°\text{C}$	$k_{22} = k_{33} = 3.21\,\text{W/m°C}$	$G_{23} = 3.45\,\text{GPa}$	$\alpha_{33} = 31 \times 10^{-6}\,°\text{C}^{-1}$

Case 3 (Figure 3, left) Center interface crack under uniform heat flow, isotropic bimaterials			
$a = 4\,\text{mm}$	$w = 20a$	$q = 10^5\,\text{W/m}^2$	
$E_1 = 1000\,\text{GPa}$	$\nu_1 = 0.3$	$\alpha_1 = 10^{-6}\,°\text{C}^{-1}$	$k_1 = 100\,\text{W/m}\,°\text{C}$
$E_2 = 100\,\text{GPa}$	$\nu_2 = 0.3$	$\alpha_2 = 10^{-7}\,°\text{C}^{-1}$	$k_2 = 100\,\text{W/m°C}$

Case 4 (Figure 3, right) Edge interface crack under uniform temperature change, isotropic bimaterials			
$a = 1\,\text{mm}$	$w = 100a$	$T_c = 100\,°\text{C}$	
$E_1 = 1000\text{GPa}$	$\nu_1 = 0.3$	$\alpha_1 = 10^{-6}\,°\text{C}^{-1}$	
$E_2 = 100\text{GPa}$	$\nu_2 = 0.3$	$\alpha_2 = 10^{-7}\,°\text{C}^{-1}$	

Case 5 (Figure 4, left) Edge interface corner under uniform temperature change, isotropic bimaterials			
$w = 1000\,\text{mm}$	$w_1 = 461\,\text{mm}$	$w_2 = 359\,\text{mm}$	$T_c = -100\,°\text{C}$
$E_1 = 72\,\text{GPa}$	$\nu_1 = 0.3$	$\alpha_1 = 18.95 \times 10^{-6}\,°\text{C}^{-1}$	
$E_2 = 280\,\text{GPa}$	$\nu_2 = 0.26$	$\alpha_2 = 2.5 \times 10^{-6}\,°\text{C}^{-1}$	

Case 6 (Figure 4, right) Interface corner under uniform temperature change, anisotropic bimaterials

$a = 1.6\,\text{mm}, \quad d = 0.1\,\text{mm}, \quad h = 7.5\,\text{mm}, \quad w = 3\,\text{mm}, \quad \beta = 20°, \quad T_c = -20\,°\text{C},$

$$C_{\text{GSO}} = \begin{bmatrix} 223 & 108 & 98.5 & 0 & 84 & 0 \\ & 150 & 102 & 0 & 33.3 & 0 \\ & & 251 & 0 & -6 & 0 \\ & & & 78.8 & 0 & 6.6 \\ & \text{sym.} & & & 68.8 & 0 \\ & & & & & 82.7 \end{bmatrix} \text{[GPa]}, \qquad C_{\text{aragonite}} = \begin{bmatrix} 87.8 & 26.3 & 36.6 & 0 & 18.75 & 0 \\ & 87 & 26.3 & 0 & 10.35 & 0 \\ & & 87.8 & 0 & 18.75 & 0 \\ & & & 42 & 0 & 0.7 \\ & \text{sym.} & & & 60.27 & 0 \\ & & & & & 42 \end{bmatrix} \text{[GPa]},$$

$$\boldsymbol{\alpha}_{\text{GSO}} = \begin{bmatrix} 4.4 & 0 & 0 \\ 0 & 14 & 0 \\ 0 & 0 & 6.8 \end{bmatrix} [10^{-6}\,°\text{C}^{-1}], \qquad \boldsymbol{\alpha}_{\text{aragonite}} = \begin{bmatrix} 22.5 & 0 & 0 \\ 0 & 17 & 0 \\ 0 & 0 & 22.5 \end{bmatrix} [10^{-6}\,°\text{C}^{-1}],$$

Table 1. Geometries, loading, and material properties of numerical examples.

the order of stress singularity δ, the order of heat flux singularity δ_h, the stress intensity factors K_{I} and K_{II} versus the radius of path integral r/a, and the reference solutions are all shown in Table 2.

From this table we see that the results calculated by the present approach are not only path-independent but also agree well with those presented in the literature for all different cases, for example, Case 1 [Sumi and Katayama 1980; Maiti 1992; Mukhopadhyay et al. 1999], Case 2 [Hwu 1990], Case 3 [Banks-Sills and Dolev 2004], Case 4 [Erdogan 1965], and Case 5 [Banks-Sills and Ishbir 2004]. In Case 6, due to the limitations of two-dimensional elements of ANSYS only the plane strain condition is considered, and hence in our example the thermal expansion coefficient α_{13} considered in the reference paper [Nomura et al. 2009], which may induce deformation in the third direction, is neglected. With this neglect, as shown in Table 2 our results are slightly different from those presented in [Nomura et al. 2009]. Table 3
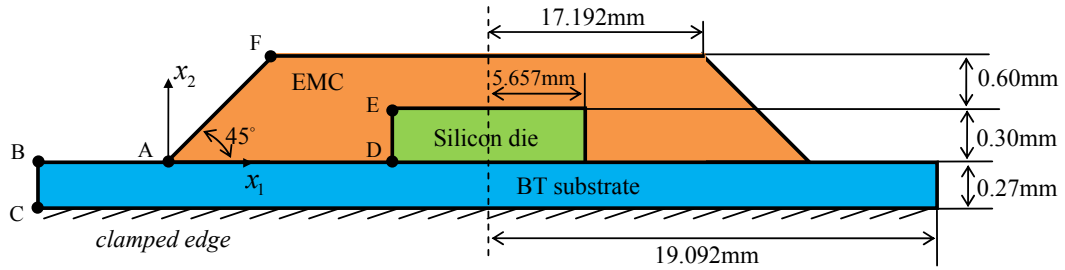
Figure 5. A sample electronic package.

| Case | Singular order | | r/a | K_{I} (MPa $\times$ mm$^{\delta_R}$) | | K_{II} (MPa $\times$ mm$^{\delta_R}$) | |
	δ	δ_h		Present	Reference	Present	Reference
1	0.5	−0.5	0.3	0.01722	0.01739[1]	0	
			0.6	0.01722	0.01695[2]	0	0
			0.8	0.01722	0.01722[3]	0	
2	0.5	−0.5		(Orthotropic material)			
			0.3	0		−0.02699	
			0.6	0	0	−0.02702	−0.02693[4]
			0.8	0		−0.02700	
				(Anisotropic material)			
			0.3	0.01147		−0.03375	
			0.6	0.01148	0.01150[4]	−0.03375	−0.03357[4]
			0.8	0.01148		−0.03376	
3*	$0.5+0.076i$	−0.5	0.3	-1.611×10^5		-8.890×10^3	
			0.6	-1.615×10^5	-1.636×10^{5} [5]	-8.850×10^3	-8.908×10^{3} [5]
			0.8	-1.615×10^5		-8.817×10^3	
4*	$0.5+0.076i$	−0.5	0.3	−1.559		10.65	
			0.6	−1.577	−1.526[6]	10.69	10.07[6]
			0.9	−1.537		10.63	
5**	0.111	−1	0.03	1.867×10^3		−317.6	
			0.06	1.869×10^3	1.848×10^{3} [7]	−317.9	–
			0.09	1.867×10^3		−317.5	
6***	$0.482+0.041i$	–	0.6	3.540	3.278[8]	16.67	22.90[8]
			0.8	3.572		16.47	

δ: order of stress singularity; δ_h: order of heat flux singularity; δ_R: real part of stress singular order;
*: reference length l is selected to be $2a$; **: normalized factor a is replaced by $w = 1000$ mm in Figure 4, left;
***: reference length l is selected to be 10 μm and normalized factor a is replaced by $d = 0.1$ mm in Figure 4,
right; [1] [Sumi and Katayama 1980]; [2] [Maiti 1992]; [3] [Mukhopadhyay et al. 1999]; [4] [Hwu 1990];
[5] [Banks-Sills and Dolev 2004]; [6] [Erdogan 1965]; [7] [Banks-Sills and Ishbir 2004]; [8] [Nomura et al. 2009].

Table 2. Comparison of stress intensity factors.

ρ/a	$(r/a = 0.3)$		$(r/a = 0.6)$	
	$K_{\mathrm{I}}(\mathrm{GPa} \times \mathrm{mm}^{0.5})$	$K_{\mathrm{II}}(\mathrm{GPa} \times \mathrm{mm}^{0.5})$	$K_{\mathrm{I}}(\mathrm{GPa} \times \mathrm{mm}^{0.5})$	$K_{\mathrm{II}}(\mathrm{GPa} \times \mathrm{mm}^{0.5})$
0.0001	-161.2	-9.093	-161.4	-8.942
0.001	-161.3	-9.075	-161.5	-8.925
0.005	-161.1	-8.910	-161.5	-8.880
0.01	-161.1	-8.890	-161.5	-8.850
0.05	-161.2	-8.878	-161.5	-8.813
0.1	-161.4	-8.952	-161.6	-8.801
0.2	-161.6	-8.861	-161.8	-8.710
0.3	-161.9	-8.672		
0.4			-162.3	-8.236
0.5			-162.5	-7.856
0.6			-162.7	-7.381

Table 3. Effects of radius ρ on the stress intensity factors for Case 3.

shows the effects of the radius ρ of the small circle chosen for the analytical area integral of (22b). From this table, we see that the effect of ρ is very trifling on the results of the stress intensity factors when $\rho/a \leq 0.2$, and this phenomenon is consistent with the results presented in [Banks-Sills and Ishbir 2004]. Note that the bigger ρ is, the more mesh and computation time we can save, and this vindicates to the use of analytical integration in (22b).

5.2. *Application to electronic packages.* A typical example of electronic packages is shown in Figure 5. This package consists of three different parts: silicon die, epoxy molding compound (EMC), and bis-maleimide triazine (BT) substrate. Their mechanical properties are shown in Table 4. Due to the discontinuity of geometries and/or material properties, stress concentration usually occurs at the corners or interface corners, such as points A, B, C, D, E, and F shown in Figure 5. To know which corner is the most critical corner, we first calculate the orders of stress/heat flux singularity. Table 5 shows the results of singular orders of these corners, in which the values of points A, B, and F are calculated from (5a)

Material	E [GPa]	ν	α [$10^{-6}\,^{\circ}\mathrm{C}^{-1}$]	k [W/m$\,^{\circ}$C]
Silicon die	131	0.3	2.8	300
EMC	16	0.25	8	14
BT substrate	26	0.11	15	0.95

Table 4. Material properties of the sample electronic package.

Singular order	Location					
	A	B	C	D	E	F
δ	0.280	0	0.143	0.277	0.253	0
δ_h	-0.979	-1	-1	-0.894	-0.699	-1

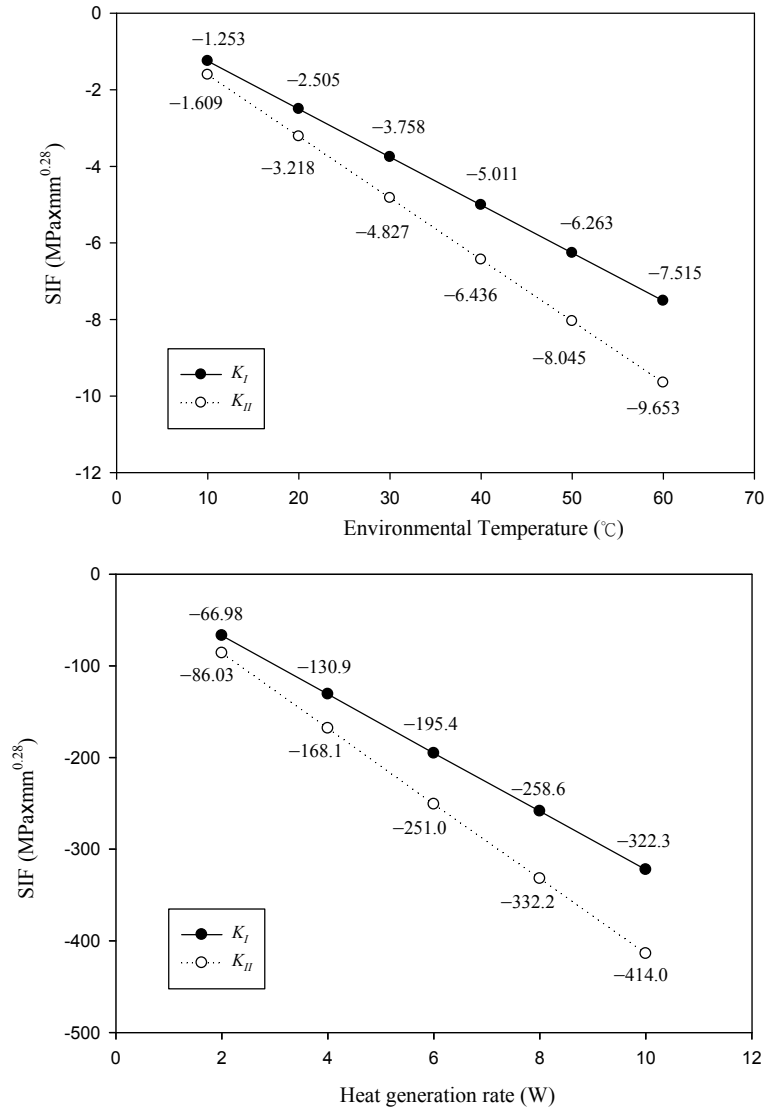**Table 5.** Orders of stress/heat flux singularity of the sample electronic package.

Figure 6. Stress intensity factors versus environmental temperature (top) and versus heat generation rate (bottom).

and (20), for stress and heat flux singularity respectively. For points C, D, and E, whose corner edges are not traction-free, related formulae can be found in [Hwu and Lee 2004]. From Table 5, we see that point A is the most critical point.

To study the effects of thermal environment on the stress intensity factors, we now consider two different thermal conditions: (1) the package is placed within an environmental temperature maintained at a constant temperature T_c, and the chip (silicon die) has a heat generation rate of 10 W; (2) the package is placed amid an environment with reference temperature $T_{ref} = 25 \, °C$ before the chip begins to generate heat, convection with heat transfer coefficient $10 \, W/m^2 °C$ is imposed on all the outer edges, and the chip (silicon die) has a heat generation rate Q. It is assumed that this package doesn't deform at $0 \, °C$. By using

the modified H-integral proposed in this paper, the stress intensity factors associated with the singular order $\delta = 0.28$ of point A are calculated by varying T_c for the first case and varying Q for the second case. Figure 6 shows that both of the mode I and mode II stress intensity factors vary linearly with environmental temperature T_c and heat generation rate Q.

6. Conclusions

In this investigation, the modified H-integral is proposed to calculate the stress intensity factors of interface corners subjected to thermal loadings, and its required near-tip solutions and auxiliary solutions of displacements, stresses, and temperature are all provided. Through several different material types, corner types, and thermal loading types used in the numerical examples, the H-integral proves its validity and versatility in thermal problems. Moreover, the path-independence property of the modified H-integral in thermal problems has been proved both theoretically and numerically. A special treatment for the strongly singular function in the domain integral of the modified H-integral saves us a lot of computational time and also raises the accuracy for the calculation of stress intensity factors.

Appendix: Explicit expressions of the near-tip solutions

Under thermal effects, the field solutions near the tip of multimaterial wedges has been shown in Equation (53) of [Hwu and Lee 2004]. To explicitly show the r-dependent relation of the near-tip solutions by using Equation (27) of the same reference, we may let the solutions along the wedge surface $\theta = \theta_0$ be

$$v_1(\theta_0) = -\delta(1-\delta)r^{-1-\delta}v_0, \quad w_1(\theta_0) = r^{1-\delta}w_0. \tag{A-1}$$

Substituting (A-1) into [Hwu and Lee 2004, (53)], the near-tip solutions can be expressed as those shown in (4a) and (4b), in which

$$\mathbf{\Gamma}(\theta) = \begin{cases} \mathbf{\Gamma}_1^*(\theta) & \text{when } \theta_0 \leq \theta \leq \theta_1, \\ \mathbf{\Gamma}_2^*(\theta)(\mathbf{K}_T)_1 & \text{when } \theta_1 \leq \theta \leq \theta_2, \end{cases}$$

$$\mathbf{E}(\theta) = \begin{cases} \mathbf{E}_1^*(\theta) & \text{when } \theta_0 \leq \theta \leq \theta_1, \\ \mathbf{E}_2^*(\theta)(\mathbf{K}_e)_1 & \text{when } \theta_1 \leq \theta \leq \theta_2, \end{cases} \tag{A-2}$$

$$\mathbf{F}(\theta) = \begin{cases} \mathbf{F}_1^*(\theta) & \text{when } \theta_0 \leq \theta \leq \theta_1, \\ \mathbf{F}_2^*(\theta)(\mathbf{K}_T)_1 + \mathbf{E}_2^*(\theta)(\mathbf{K}_c)_1 & \text{when } \theta_1 \leq \theta \leq \theta_2, \end{cases}$$

where

$$\mathbf{\Gamma}_k^*(\theta) = \mathbf{\Theta}_k \langle \hat{\tau}_\beta^{-\delta}(\theta, \theta_{k-1}) \rangle_k \mathbf{\Theta}_k^{-1}, \qquad \mathbf{E}_k^*(\theta) = \hat{N}_k^{1-\delta}(\theta, \theta_{k-1}),$$

$$\mathbf{F}_k^*(\theta) = \frac{1}{\delta(1-\delta)}\{\hat{N}_k^{1-\delta}(\theta, \theta_{k-1})\mathbf{U}_k - \mathbf{U}_k \langle \hat{\tau}_\beta^{1-\delta}(\theta, \theta_{k-1}) \rangle_k\} \langle \hat{\tau}_\beta^{1-\delta}(\theta_{k-1}, 0) \rangle_k \mathbf{\Theta}_k^{-1}, \quad k = 1, 2, \tag{A-3}$$

and

$$(\mathbf{K}_T)_1 = \mathbf{\Gamma}_1^*(\theta_1), \quad (\mathbf{K}_e)_1 = \mathbf{E}_1^*(\theta_1), \quad (\mathbf{K}_c)_1 = \mathbf{F}_1^*(\theta_1),$$

$$\mathbf{\Theta}_k = \begin{bmatrix} 1 & 1 \\ -ik_0 & ik_0 \end{bmatrix}_k, \quad \mathbf{U}_k = \begin{bmatrix} \mathbf{c}_k & \bar{\mathbf{c}}_k \\ \mathbf{d}_k & \bar{\mathbf{d}}_k \end{bmatrix},$$

$$\hat{\mathbf{N}}_k^{1-\delta}(\theta, \theta') = \begin{bmatrix} \mathbf{A}_k & \bar{\mathbf{A}}_k \\ \mathbf{B}_k & \bar{\mathbf{B}}_k \end{bmatrix} \begin{bmatrix} \langle \hat{\mu}_\alpha^{1-\delta}(\theta, \theta') \rangle_k & \mathbf{0} \\ \mathbf{0} & \overline{\langle \hat{\mu}_\alpha^{1-\delta}(\theta, \theta') \rangle}_k \end{bmatrix} \begin{bmatrix} \mathbf{B}_k^T & \mathbf{A}_k^T \\ \bar{\mathbf{B}}_k^T & \bar{\mathbf{A}}_k^T \end{bmatrix}, \qquad \text{(A-4)}$$

$$\hat{\mu}_\alpha(\theta, \theta') = \cos(\theta - \theta') + \sin(\theta - \theta')\mu_\alpha(\theta'), \quad \mu_\alpha(\theta') = \frac{\mu_\alpha \cos\theta' - \sin\theta'}{\mu_\alpha \sin\theta' + \cos\theta'}, \quad \alpha = 1, 2, 3,$$

$$\hat{\tau}_\beta(\theta, \theta') = \cos(\theta - \theta') + \sin(\theta - \theta')\tau_\beta(\theta'), \quad \tau_\beta(\theta') = \frac{\tau_\beta \cos\theta' - \sin\theta'}{\tau_\beta \sin\theta' + \cos\theta'}, \quad \beta = 1, 2,$$

In (A-2)–(A-4), subscript k (taking the values 1, 2) denotes the quantities related to the k-th wedge, whereas subscript α and β denote the diagonal components of the diagonal matrix. $i = \sqrt{-1}$ is an imaginary unit; a bar above a letter denotes complex conjugation; k_0 is a real constant related to the heat conduction coefficients k_{ij} by

$$k_0 = \sqrt{k_{11}k_{22} - k_{12}^2}; \qquad \text{(A-5)}$$

μ_α and τ_β are the elastic eigenvalues and thermal eigenvalues; and $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{c}$, $\mathbf{d}$ are the elastic eigenvector matrices and thermal eigenvectors of the Stroh formalism of anisotropic elasticity; see [Ting 1996; Hwu 2010].

References

[Banks-Sills and Dolev 2004] L. Banks-Sills and O. Dolev, "The conservative M-integral for thermal-elastic problems", *Int. J. Fract.* **125**:1–2 (2004), 149–170.

[Banks-Sills and Ishbir 2004] L. Banks-Sills and C. Ishbir, "A conservative integral for bimaterial notches subjected to thermal stresses", *Int. J. Numer. Methods Eng.* **60**:6 (2004), 1075–1102.

[Brown and Erdogan 1968] E. J. Brown and F. Erdogan, "Thermal stresses in bonded materials containing cuts on the interface", *Int. J. Eng. Sci.* **6**:9 (1968), 517–529.

[Bueckner 1973] H. F. Bueckner, "Field singularities and related integral representations", pp. 239–314 in *Methods of analysis and solutions of crack problems: recent developments in fracture mechanics: theory and methods of solving crack problems*, edited by G. C. Sih, Mechanics of Fracture **1**, Noordhoff, Leyden, 1973.

[Chen 1985] Y. Z. Chen, "New path independent integrals in linear elastic fracture mechanics", *Eng. Fract. Mech.* **22**:4 (1985), 673–686.

[Choi and Earmme 1992] N. Y. Choi and Y. Y. Earmme, "Evaluation of stress intensity factors in circular arc-shaped interfacial crack using L integral", *Mech. Mater.* **14**:2 (1992), 141–153.

[Clements 1973] D. L. Clements, "Thermal stress in an anisotropic elastic half-space", *SIAM J. Appl. Math.* **24**:3 (1973), 332–337.

[Erdogan 1965] F. Erdogan, "Stress distribution in bonded dissimilar materials with cracks", *J. Appl. Mech.* (*ASME*) **32** (1965), 403–410.

[Hwu 1990] C. Hwu, "Thermal stresses in an anisotropic plate disturbed by an insulated elliptic hole or crack", *J. Appl. Mech.* (*ASME*) **57**:4 (1990), 916–922.

[Hwu 1992] C. B. Hwu, "Thermoelastic interface crack problems in dissimilar anisotropic media", *Int. J. Solids Struct.* **29**:16 (1992), 2077–2090.

[Hwu 2010] C. Hwu, *Anisotropic elastic plates*, Springer, New York, 2010.

[Hwu and Ikeda 2008] C. Hwu and T. Ikeda, "Electromechanical fracture analysis for corners and cracks in piezoelectric materials", *Int. J. Solids Struct.* **45**:22–23 (2008), 5744–5764.

[Hwu and Kuo 2007] C. Hwu and T. L. Kuo, "A unified definition for stress intensity factors of interface corners and cracks", *Int. J. Solids Struct.* **44**:18–19 (2007), 6340–6459.

[Hwu and Lee 2004] C. Hwu and W.-J. Lee, "Thermal effect on the singular behavior of multibonded anisotropic wedges", *J. Therm. Stresses* **27**:2 (2004), 111–136.

[Hwu et al. 2003] C. Hwu, M. Omiya, and K. Kishimoto, "A key matrix *N* for the stress singularity of the anisotropic elastic composite wedges", *JSME Int. J. A Mech. M.* **46**:1 (2003), 40–50.

[Ikeda and Sun 2001] T. Ikeda and C. T. Sun, "Stress intensity factor analysis for an interface crack between dissimilar isotropic materials under thermal stress", *Int. J. Fract.* **111**:3 (2001), 229–249.

[Im and Kim 2000] S. Im and K.-S. Kim, "An application of two-state *M*-integral for computing the intensity of the singular near-tip field for a generic wedge", *J. Mech. Phys. Solids* **48**:1 (2000), 129–151.

[Kuo and Hwu 2010] T.-L. Kuo and C. Hwu, "Multi-order stress intensity factors along three-dimensional interface corners", *J. Appl. Mech. (ASME)* **77**:3 (2010), 031020.

[Lancaster and Salkauskas 1981] P. Lancaster and K. Salkauskas, "Surfaces generated by moving least squares methods", *Math. Comput.* **37** (1981), 141–158.

[Maiti 1992] S. K. Maiti, "A finite element for variable order singularities based on the displacement formulation", *Int. J. Numer. Methods Eng.* **33**:9 (1992), 1955–1974.

[Mukhopadhyay et al. 1999] N. K. Mukhopadhyay, S. K. Maiti, and A. Kakodkar, "Modified crack closure integral based computation of stress intensity factors for 2-D thermoelastic problems through boundary element method", *Nucl. Eng. Des.* **187**:3 (1999), 277–290.

[Munz and Yang 1992] D. Munz and Y. Y. Yang, "Stress singularities at the interface in bonded dissimilar materials under mechanical and thermal loading", *J. Appl. Mech. (ASME)* **59**:4 (1992), 857–861.

[Munz and Yang 1993] D. Munz and Y. Y. Yang, "Stresses near the edge of bonded dissimilar materials described by two stress intensity factors", *Int. J. Fract.* **60**:2 (1993), 169–177.

[Nagai et al. 2007] M. Nagai, T. Ikeda, and N. Miyazaki, "Stress intensity factor analysis of an interface crack between dissimilar anisotropic materials under thermal stress using the finite element analysis", *Int. J. Fract.* **146**:4 (2007), 233–248.

[Nomura et al. 2009] Y. Nomura, T. Ikeda, and N. Miyazaki, "Stress intensity factor analysis at an interfacial corner between anisotropic bimaterials under thermal stress", *Eng. Fract. Mech.* **76**:2 (2009), 221–235.

[Nowacki 1962] W. Nowacki, *Thermoelasticity*, Addison-Wesley, Reading, MA, 1962.

[Rice 1968] J. R. Rice, "A path independent integral and the approximate analysis of strain concentration by notches and cracks", *J. Appl. Mech. (ASME)* **35** (1968), 379–386.

[Sinclair et al. 1984] G. B. Sinclair, M. Okajima, and J. H. Griffen, "Path independent integrals for computing stress intensity factors at sharp notches in elastic plates", *Int. J. Numer. Methods Eng.* **20**:6 (1984), 999–1008.

[Stern 1976] M. Stern, "A boundary integral representation for stress intensity factors", pp. 125–132 in *Recent advances in engineering science* (Raleigh, NC, 1973), vol. 7, edited by T. S. Chang, Scientific Publishers, Boston, 1976.

[Stroh 1958] A. N. Stroh, "Dislocations and cracks in anisotropic elasticity", *Philos. Mag.* **3**:30 (1958), 625–646.

[Sumi and Katayama 1980] N. Sumi and T. Katayama, "Thermal stress singularities at tips of a Griffith crack in a finite rectangular plate", *Nucl. Eng. Des.* **60**:3 (1980), 389–394.

[Ting 1996] T. C. T. Ting, *Anisotropic elasticity: theory and applications*, Oxford Engineering Science Series **45**, Oxford University Press, New York, 1996.

CHYANBIN HWU: chwu@mail.ncku.edu.tw
Institute of Aeronautics and Astronautics, National Cheng Kung University, No.1, University Road, Tainan 70101, Taiwan

TAI-LIANG KUO: kelvenkiss@yahoo.com.tw
Institute of Aeronautics and Astronautics, National Cheng Kung University, No.1, University Road, Tainan 70101, Taiwan

CHUN-CHIH HUANG: orange648@gmail.com
Institute of Aeronautics and Astronautics, National Cheng Kung University, No.1, University Road, Tainan 70101, Taiwan

msp

THREE-DIMENSIONAL ISOFIELD MICROMECHANICS MODEL FOR EFFECTIVE ELECTROTHERMOELASTIC PROPERTIES OF PIEZOELECTRIC COMPOSITES

SANTOSH KAPURIA AND POONAM KUMARI

A fully coupled three-dimensional micromechanics model based on the isofield method is developed for the effective electrothermoelastic properties of piezoelectric fiber-reinforced composite (PFRC) materials with poling and an electric field applied normal to the fiber direction. In the isofield method, the strain and electric field components parallel to the plane connecting two phases are assumed to be uniform across both phases, and likewise for the stress and electric displacement components normal to the connecting plane. The model employs the isofield assumptions for two possible connectivities, which are then combined so as to yield transverse isotropy of the effective properties when both constituents are transversely isotropic. The assumption of uniform electric field across two phases made by some existing theories can be achieved as a special case of the present formulation when the dielectric constants of the fiber and matrix phases are equal. The effects of the fiber volume fraction and dielectric ratio on the effective properties are studied for two PFRC systems, PZT-7A/epoxy and PZT-5H/epoxy. The results are compared with those available in the literature based on uniform electric field assumptions. It is found that the dielectric ratio has a very significant effect on the electromechanical and electrothermal coupling constants of PFRCs.

1. Introduction

Piezoelectric materials are being increasingly used as distributed sensors and actuators in structural health monitoring [Park et al. 2010] and control [Dong et al. 2006] applications. Their advantages over other available smart materials include easy commercial availability, efficient conversion of energy, relatively linear electromechanical behavior (at low fields), and large useful bandwidth [Chopra 2002]. However, for large-scale structural control applications such as in aerospace, automotive, and ship structures, monolithic piezoelectric actuators and sensors suffer from shortcomings with regard to tailorable anisotropic actuation, that is, directional actuation, robustness against damage during use and handling, ability to cover the entire structure for distributed actuation and sensing, and conformability to curved shell-type structural members. To address these concerns, piezoelectric fiber-reinforced composites (PFRCs) have been developed recently by embedding piezoceramic fibers in a resin matrix system, which, in addition to overcoming all the above-mentioned shortcomings, also possess higher specific stiffness, toughness, operating voltage range (from -1500 to $+2800\,\mathrm{V}$), and lifespan (200 million cycles) than the bulk material [Uchino 2000]. For designing such PFRC sensors and actuators as well as smart laminates integrated with these, it is necessary to have micromechanics models capable of estimating the effective electrothermoelastic properties of a unidirectional PFRC layer from the properties of its constituents.

A number of micromechanics models have been proposed for piezoelectric composites in which the piezoelectric fibers are oriented along the thickness direction (Figure 1a). In this case, the poling and electric field directions are parallel to the fiber axis, causing a d_{33} effect, which is useful in ultrasonic transducer applications. These micromechanics models have been developed based on the Voigt-type isofield method [Chan and Unsworth 1989; Smith and Auld 1991], the Mori–Tanaka method [Dunn and Taya 1993], the self-consistent method [Dunn and Taya 1993; Levin et al. 2000], the generalized method of cells [Aboudi 1998], and the asymptotic homogenization method [Sabina et al. 2001; Levin et al. 2008].

Commonly used piezoelectric materials such as PZT and PVDF are transversely isotropic about their poling axes (class mm6 symmetry). Thus, when the piezoelectric fibers are aligned along the poling direction and the matrix is also transversely isotropic about the fiber direction, the transverse isotropy is retained in the composite system. All the above micromechanics models are thus concerned with transversely isotropic effective properties. In [Kar-Gupta and Venkatesh 2005], a unit cell-based finite element model was employed to obtain the electromechanical effective properties of a 1-3 piezoelectric composite system. In this system, the fibers are oriented along the thickness direction, while the fiber and matrix phases are poled along different directions (parallel or normal to the fiber direction) to generate a wide range of specific acoustic impedances.

For structural applications, however, the stiff piezoceramic fibers must be oriented in the plane of the structures (Figure 1b) and the poling as well as the electric field directions are perpendicular to the fiber axis, resulting in a d_{31} effect. In this case, the composite system is no longer transversely isotropic about the fiber axis, particularly with regard to the piezoelectric properties. Very few studies have been reported on the micromechanics of PFRC laminas of this type. The first such model was presented in [Bent 1994] using the isofield method for computing effective electroelastic properties. In that work, even though a general methodology was briefly outlined for calculation of effective material properties for the three-dimensional (3D) stress field considering two possible connectivity planes of the constituent phases, the detailed closed-form solutions and the results for effective material properties were presented by considering the uniaxial stress field only. The results for the 3D stress field based on this method have been presented only recently [Kapuria and Kumari 2010]. In [Mallik and Ray 2003;
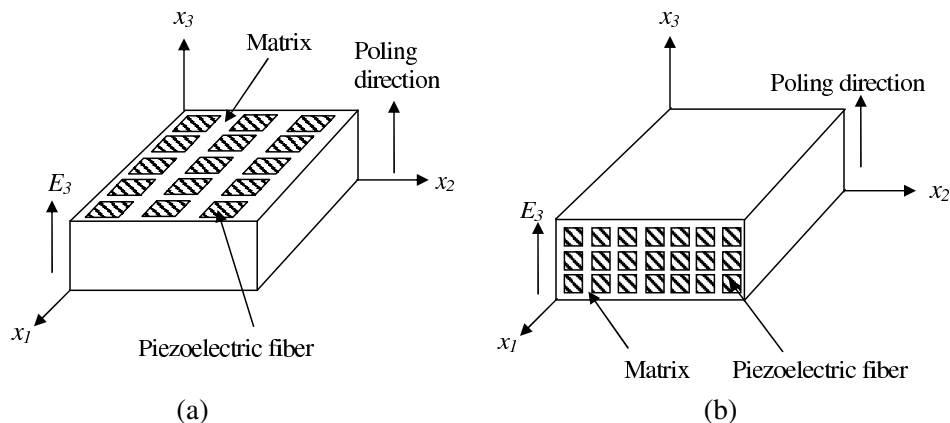

Figure 1. Schematic representation of PFRC with poling and electric field directions (a) parallel and (b) normal to fiber.

Ray 2006] a simpler model was presented using the uniform fields concept with a single connectivity plane (parallel to the fiber axis) for computing the effective electroelastic properties of PFRCs, which has been recently extended to the thermoelectroelastic case in [Kumar and Chakraborty 2009]. In this formulation, however, the electric field is assumed to be uniform across both the piezoelectric fiber and elastic matrix phases, which is not achievable when the electric field is applied across the thickness of the lamina, due to large difference in the dielectric constants of the two phases. The assumption is valid only when the two phases have the same dielectric constants, which is not the case for the materials commonly used for the purpose. Their formulation thus gives unrealistically high values (upper bounds) of the effective piezoelectric constants.

This paper presents a coupled 3D isofield model for estimating the effective electrothermoelastic properties of a unidirectional PFRC lamina with in-plane fibers and poling and electric field applied along the thickness direction. The effective properties are obtained for representative volume elements (RVEs) with two possible connectivity planes for the piezoelectric fiber and matrix phases, namely, parallel and normal to the fiber plane. The two models are combined in a way which maintains the transverse isotropy in the effective properties when both constituents are transversely isotropic about the fiber axis. Results are presented for two PFRC systems comprising, respectively, PZT-7A and PZT-5H fibers with an epoxy matrix. The results are compared with those of [Ray 2006; Kumar and Chakraborty 2009] obtained based on the uniform electric field assumption. The effect of the ratio of transverse dielectric constants of the fiber and matrix phases (hereafter called the dielectric ratio) on the effective piezoelectric and pyroelectric constants is illustrated. It is revealed that both effective piezoelectric and pyroelectric constants are maximal when the dielectric ratio is unity, and reduce drastically as the ratio increases. The effective thermoelastic properties are also compared with simplified models such as the rule of mixtures (ROM) and the modified rule of mixtures (MROM) [Gibson 2007] so as to ascertain their validity.

2. 3D isofield micromechanics model

2.1. *Constitutive relations.* The effective thermoelectroelastic constants of PFRC materials are determined from the properties of individual phases (fiber and matrix) by generalizing the 3D isofield approach of [Bent 1994; Kapuria and Kumari 2010] for the electrothermomechanical field. In order to have a unified treatment, both fiber and matrix are assumed to be piezoelectric materials, which are of orthotropic class mm2 symmetry, with principal material axes x_1, x_2, and x_3, and are polarized along the thickness direction x_3. The 3D linear constitutive equations of such a piezoelectric continuum are given by [Auld 1973]

$$\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \gamma_{23} \\ \gamma_{31} \\ \gamma_{12} \\ D_1 \\ D_2 \\ D_3 \end{bmatrix} = \begin{bmatrix} s_{11} & s_{12} & s_{13} & 0 & 0 & 0 & 0 & 0 & d_{31} \\ s_{12} & s_{22} & s_{23} & 0 & 0 & 0 & 0 & 0 & d_{32} \\ s_{13} & s_{23} & s_{33} & 0 & 0 & 0 & 0 & 0 & d_{33} \\ 0 & 0 & 0 & s_{44} & 0 & 0 & 0 & d_{24} & 0 \\ 0 & 0 & 0 & 0 & s_{55} & 0 & d_{15} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & s_{66} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & d_{15} & 0 & \epsilon_{11} & 0 & 0 \\ 0 & 0 & 0 & d_{24} & 0 & 0 & 0 & \epsilon_{22} & 0 \\ d_{31} & d_{32} & d_{33} & 0 & 0 & 0 & 0 & 0 & \epsilon_{33} \end{bmatrix} \begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \tau_{23} \\ \tau_{31} \\ \tau_{12} \\ E_1 \\ E_2 \\ E_3 \end{bmatrix} + \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ q_3 \end{bmatrix} T, \tag{1}$$

where ε_i and γ_{ij} denote the normal and shearing strain components, σ_i and τ_{ij} denote the normal and shear stress components, D_i denotes the electric displacements, E_i denotes the electric field components in the principal material axis system, and T denotes the temperature change over the reference stress-free temperature. Constants s_{ij}, d_{ij}, ϵ_{ij}, α_i, and q_3 are the elastic compliances, piezoelectric strain constants, dielectric constants, thermal expansion coefficients and pyroelectric constant, respectively.

The physical background of the isofield micromechanics model including its assumptions is described in Appendix A. In the generalized 3D isofield approach, the effective properties are first obtained for RVEs of two possible connectivities for the piezoelectric fiber and matrix phases: models A and B with material connectivity on the x_1-x_2 and x_1-x_3 planes respectively, as shown in Figure 2. The strain and electric field components parallel to the connecting plane of the two phases in a given RVE are assumed to be uniform across both phases (isofield condition), while isostress and isoelectric displacement conditions are assumed to exist along the direction normal to the connecting plane. The computations for the effective properties for the two models A and B are described below followed by the procedure of combining the two.

2.2. _Model A._ In model A, the connecting plane is x_1-x_2, and hence the strain components (ε_1, ε_2, γ_{12}) and electric field components (E_1, E_2) which are parallel to the connecting plane are assumed to be uniform across both phases, and isostress and isoelectric displacement conditions are assumed to exist
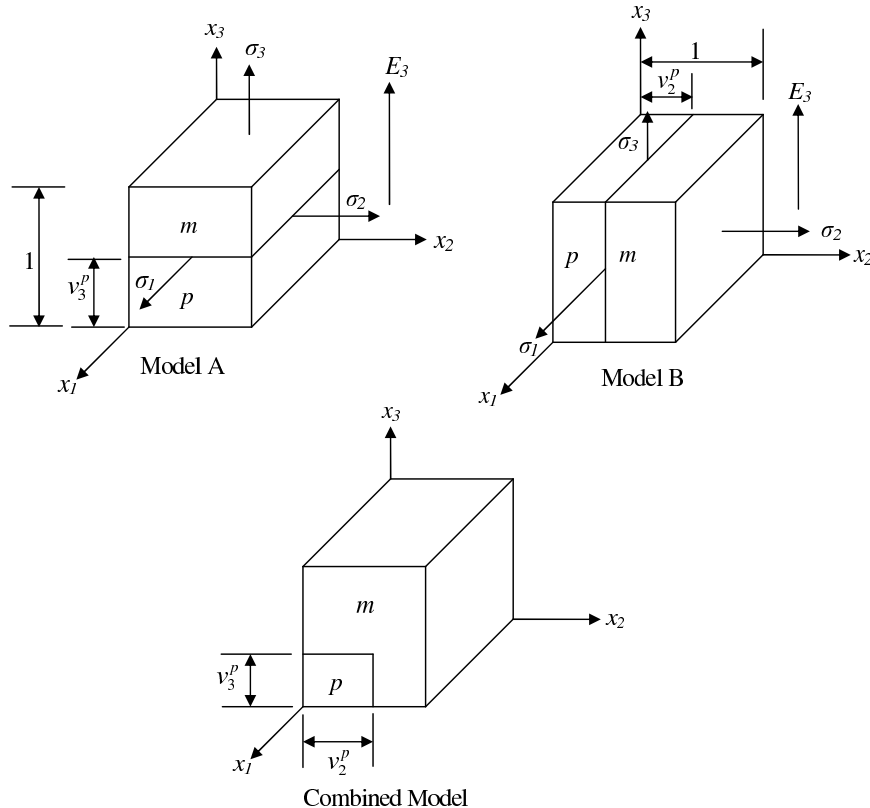

Figure 2. RVEs for isofield model (p: piezoceramic, m: matrix).

for $(\sigma_3, \tau_{23}, \tau_{13})$ and D_3, which are normal to the connecting plane. The electric field is considered to be applied along the thickness direction only, and the in-plane electric field components that may be used due to direct piezoelectric effect is neglected being small in comparison to E_3, that is, $E_1 \simeq E_2 \simeq 0$. It follows from (1) that, for this case, γ_{31} and γ_{12} become electromechanically uncoupled and hence the expressions for the effective shear compliances s_{55} and s_{66} (both normal to the fiber axis) will be similar. Hence, only one strain components γ_{12} is included in the following derivation, without any loss in generality. Thus, the isofield conditions for this case can be written as

$$\bar{U}^a = U_p^a = U_m^a \quad \text{with} \quad U^a = \left[\varepsilon_1, \varepsilon_2, \sigma_3, \tau_{23}, \gamma_{12}, D_3\right]^{\mathrm{T}}. \tag{2}$$

The superscript a corresponds to model A. An overbar denotes the average value, and subscripts p and m (also used as superscripts elsewhere) denote piezoelectric fiber and matrix phases, respectively. Temperature change is assumed to be uniform over fiber and matrix phases, that is, $\bar{T} = T_p = T_m$. The average value of the complimentary field components $\sigma_1, \sigma_2, \varepsilon_3, \gamma_{23}, \tau_{12}$, and E_3 will have contributions from each phase in proportion to their volume fractions v_3^p and $v_3^m(= 1 - v_3^p)$ measured along the x_3 direction:

$$\bar{V}^a = v_3^p V_p^a + v_3^m V_m^a \quad \text{with} \quad V^a = \left[\sigma_1, \sigma_2, \varepsilon_3, \gamma_{23}, \tau_{12}, E_3\right]^{\mathrm{T}}. \tag{3}$$

The dependent field variables V^a can be expressed in terms of the independent variables U^a using the constitutive equation (1) as

$$V_\gamma^a = A^\gamma U_\gamma^a + A_t^\gamma \bar{T} \quad \text{with} \quad A^\gamma = \begin{bmatrix} a_{11} & -a_{12} & a_{13} & 0 & 0 & a_{16} \\ -a_{12} & a_{22} & a_{23} & 0 & 0 & a_{26} \\ -a_{13} & -a_{23} & a_{33} & 0 & 0 & a_{36} \\ 0 & 0 & 0 & a_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & a_{55} & 0 \\ a_{16} & a_{26} & -a_{36} & 0 & 0 & a_{66} \end{bmatrix}^\gamma, \quad A_t^\gamma = \begin{bmatrix} a_{t1} \\ a_{t2} \\ a_{t3} \\ 0 \\ 0 \\ a_{t6} \end{bmatrix}^\gamma, \tag{4}$$

where $\gamma = p, m$ and the constants a_{ij}^γ and a_{tj}^γ are given by

$$a_{11}^\gamma = \frac{s_{22}'}{\delta}, \quad a_{12}^\gamma = \frac{s_{12}'}{\delta}, \quad a_{13}^\gamma = p_2^s, \quad a_{16}^\gamma = p_2^d, \quad a_{22}^\gamma = \frac{s_{11}'}{\delta}, \quad a_{23}^\gamma = p_1^s, \quad a_{26}^\gamma = p_1^d,$$

$$a_{33}^\gamma = p_2^s s_{13}' + p_1^s s_{23}' + s_{33} - d_{33} d_{33}', \quad a_{36}^\gamma = p_2^d s_{13}' + p_1^d s_{23}' + d_{33}', \quad a_{44}^\gamma = s_{44}, \quad a_{55}^\gamma = \frac{1}{s_{66}},$$

$$a_{66}^\gamma = -p_2^d d_{31}' - p_1^d d_{32}' + \frac{1}{\epsilon_{33}}, \quad d_{3j}' = \frac{d_{3j}}{\epsilon_{33}}, \quad s_{ij}' = s_{ij} - d_{3i} d_{3j}', \quad \delta = s_{11}' s_{22}' - s_{12}'^2,$$

$$a_{t1}^\gamma = p_2^t, \quad a_{t2}^\gamma = p_1^t, \quad a_{t3}^\gamma = p_3^t, \quad a_{t6}^\gamma = p_6^t, \quad \alpha_j' = \alpha_j - d_{3j}' q_3,$$

$$p_3^t = s_{13}' p_2^t + s_{23}' p_1^t + \alpha_3', \quad p_6^t = -d_{31}' p_2^t - d_{32}' p_1^t - \frac{q_3}{\epsilon_{33}}, \quad p_i^t = (s_{i1}' \alpha_2' - s_{2i}' \alpha_1') \frac{(-1)^i}{\delta},$$

$$p_i^s = (s_{i1}' s_{23}' - s_{2i}' s_{13}') \frac{(-1)^i}{\delta}, \quad p_i^d = (s_{i1}' d_{32}' - s_{2i}' d_{31}') \frac{(-1)^i}{\delta}, \quad i = 1, 2, \quad j = 1, 2, 3. \tag{5}$$

Substituting (4) for $\gamma = p$ and m into (3) and applying (2) yields

$$\bar{V}^a = (v_3^p A^p + v_3^m A^m) \bar{U}^a + (v_3^p A_t^p + v_3^m A_t^m) \bar{T} = A \bar{U}^a + A_t \bar{T}, \tag{6}$$

and therefore

$$\bar{U}^a = \hat{A}\bar{V}^a + \hat{A}_t\bar{T}, \qquad \hat{A} = A^{-1}, \qquad \hat{A}_t = -A^{-1}A_t. \tag{7}$$

Rearranging (7) in the form of (1), the effective constitutive equation and hence the effective material properties of PFRC for model A are obtained as

$$
\begin{bmatrix} \bar{\varepsilon}_1 \\ \bar{\varepsilon}_2 \\ \bar{\varepsilon}_3 \\ \bar{\gamma}_{23} \\ \bar{\gamma}_{12} \\ \bar{D}_3 \end{bmatrix}
=
\begin{bmatrix}
s_{11} & s_{12} & s_{13} & 0 & 0 & d_{31} \\
s_{12} & s_{22} & s_{23} & 0 & 0 & d_{32} \\
s_{13} & s_{23} & s_{33} & 0 & 0 & d_{33} \\
0 & 0 & 0 & s_{44} & 0 & 0 \\
0 & 0 & 0 & 0 & s_{66} & 0 \\
d_{31} & d_{32} & d_{33} & 0 & 0 & \epsilon_{33}
\end{bmatrix}^A
\begin{bmatrix} \bar{\sigma}_1 \\ \bar{\sigma}_2 \\ \bar{\sigma}_3 \\ \bar{\tau}_{23} \\ \bar{\tau}_{12} \\ \bar{E}_3 \end{bmatrix}
+
\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ 0 \\ 0 \\ q_3 \end{bmatrix}^A
\bar{T}, \tag{8}
$$

where s_{ij}^A, d_{ij}^A, ϵ_{33}^A, α_i^A, and q_3^A are the effective material properties for model A, given by

$$
\begin{aligned}
& s_{i3}^A = \hat{a}_{i3}/\hat{a}_{33}, && s_{33}^A = 1/\hat{a}_{33}, && d_{33}^A = \hat{a}_{36}/\hat{a}_{33}, \\
& s_{ij}^A = \hat{a}_{ij} - s_{i3}^A s_{j3}^A/s_{33}^A, && && d_{3j}^A = \hat{a}_{6j} - d_{33}^A s_{j3}^A/s_{33}^A, \quad \text{for } i, j = 1, 2, \\
& s_{44}^A = 1/\hat{a}_{44}, && s_{66}^A = \hat{a}_{55}, && \epsilon_{33}^A = \hat{a}_{66} - (d_{33}^A)^2/s_{33}^A, \\
& \alpha_1^A = \hat{a}_{t1} - \hat{a}_{13}\hat{a}_{t3}/\hat{a}_{33}, && && \alpha_2^A = \hat{a}_{t2} - \hat{a}_{23}\hat{a}_{t3}/\hat{a}_{33}, \\
& \alpha_3^A = -\hat{a}_{t3}/\hat{a}_{33}, && && q_3^A = \hat{a}_{t6} - \hat{a}_{63}\hat{a}_{t3}/\hat{a}_{33}.
\end{aligned} \tag{9}
$$

2.3. Model B. In this case, the connecting plane being $x_1 - x_3$, the isofield condition over both phases is assumed to exist for ε_1, σ_2, ε_3, τ_{23}, γ_{31}, τ_{12}, and E_3, and their complimentary counterparts, σ_1, ε_2, σ_3, γ_{23}, τ_{23}, γ_{12}, and D_3 are averaged over the two phases

$$
\bar{U}^b = U_p^b = U_m^b, \qquad\qquad U^b = \begin{bmatrix} \varepsilon_1, \sigma_2, \varepsilon_3, \tau_{23}, \tau_{12}, E_3 \end{bmatrix}^{\mathrm{T}}, \tag{10}
$$

$$
\bar{V}^b = v_2^p V_p^b + v_2^m V_m^b, \qquad\qquad V^b = \begin{bmatrix} \sigma_1, \varepsilon_2, \sigma_3, \gamma_{23}, \gamma_{12}, D_3 \end{bmatrix}^{\mathrm{T}}, \tag{11}
$$

where v_2^p is the volume fraction of the fiber phase measured in model B along the x_2 direction (see Figure 2) and $v_2^m = 1 - v_2^p$. Using the constitutive equation (1), the dependent variables V_γ^b can be expressed in terms of the independent variables U_γ^b as

$$
V_\gamma^b = B^\gamma U_\gamma^b + B_t^\gamma \bar{T} \quad \text{with} \quad
B^\gamma =
\begin{bmatrix}
b_{11} & -b_{12} & b_{13} & 0 & 0 & -b_{16} \\
b_{12} & b_{22} & b_{23} & 0 & 0 & b_{26} \\
b_{13} & -b_{23} & b_{33} & 0 & 0 & -b_{36} \\
0 & 0 & 0 & b_{44} & 0 & 0 \\
0 & 0 & 0 & 0 & b_{55} & 0 \\
b_{16} & b_{26} & b_{36} & 0 & 0 & b_{66}
\end{bmatrix}^\gamma,
\quad
B_t^\gamma =
\begin{bmatrix} b_{t1} \\ b_{t2} \\ b_{t3} \\ 0 \\ 0 \\ b_{t6} \end{bmatrix}^\gamma, \tag{12}
$$

where $\gamma = p, m$ and the elements b_{ij}^{γ} and b_{tj}^{γ} are given by

$$
\begin{aligned}
&b_{11}^{\gamma} = s_{33}/\delta_1, &&b_{12}^{\gamma} = (s_{33}s_{12} - s_{13}s_{23})/\delta_1, &&b_{22}^{\gamma} = s_{22} - s_{21}b_{12} - s_{23}b_{23},\\
&b_{13}^{\gamma} = -s_{13}/\delta_1, &&b_{16}^{\gamma} = (s_{33}d_{31} - s_{13}d_{33})/\delta_1, &&b_{26}^{\gamma} = d_{32} - s_{21}b_{16} - s_{23}b_{36},\\
&b_{33}^{\gamma} = s_{11}/\delta_1, &&b_{23}^{\gamma} = (s_{11}s_{23} - s_{13}s_{12})/\delta_1, &&b_{66}^{\gamma} = \epsilon_{33} - d_{31}b_{16} - d_{33}b_{36},\\
&b_{44}^{\gamma} = s_{44}, &&b_{36}^{\gamma} = (s_{11}d_{33} - s_{13}d_{31})/\delta_1, &&b_{t1}^{\gamma} = (-s_{33}\alpha_1 + s_{13}\alpha_3)/\delta_1,\\
&b_{55}^{\gamma} = s_{66}, &&b_{t2}^{\gamma} = (s_{21}b_{t1} + s_{23}b_{t3} + \alpha_2), &&b_{t3}^{\gamma} = (-s_{11}\alpha_3 + s_{13}\alpha_1)/\delta_1,\\
&\delta_1 = s_{33}s_{11} - s_{13}^2, &&b_{t6}^{\gamma} = d_{31}b_{t1} + d_{33}b_{t3} + q_3. &&
\end{aligned}
\tag{13}
$$

Substituting (13) for $\gamma = p$ and m into (11) and applying (10) yields

$$
\bar{V}^b = (v_2^p B^p + v_2^m B^m)\bar{U}^b + (v_2^p B_t^p + v_2^m B_t^m)\bar{T} = \bar{B}\bar{U}^b + \bar{B}_t\bar{T}.
\tag{14}
$$

Rewriting (14) in the original form of (1) yields the effective constitutive equation for model B as

$$
\begin{bmatrix} \bar{\varepsilon}_1 \\ \bar{\varepsilon}_2 \\ \bar{\varepsilon}_3 \\ \bar{\gamma}_{23} \\ \bar{\gamma}_{12} \\ \bar{D}_3 \end{bmatrix} = \begin{bmatrix} s_{11} & s_{12} & s_{13} & 0 & 0 & d_{31} \\ s_{12} & s_{22} & s_{23} & 0 & 0 & d_{32} \\ s_{13} & s_{23} & s_{33} & 0 & 0 & d_{33} \\ 0 & 0 & 0 & s_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & s_{66} & 0 \\ d_{31} & d_{32} & d_{33} & 0 & 0 & \epsilon_{33} \end{bmatrix}^B \begin{bmatrix} \bar{\sigma}_1 \\ \bar{\sigma}_2 \\ \bar{\sigma}_3 \\ \bar{\tau}_{23} \\ \bar{\tau}_{12} \\ \bar{E}_3 \end{bmatrix} + \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ 0 \\ 0 \\ q_3 \end{bmatrix}^B \bar{T},
\tag{15}
$$

where the effective coefficients for model B are obtained as

$$
\begin{aligned}
&s_{11}^B = \bar{b}_{33}/\delta_2, &&s_{44}^B = \bar{b}_{44}, &&s_{12}^B = (\bar{b}_{33}\bar{b}_{12} - \bar{b}_{13}\bar{b}_{23})/\delta_2, &&s_{13}^B = -\bar{b}_{13}/\delta_2,\\
&s_{33}^B = \bar{b}_{11}/\delta_2, &&s_{66}^B = \bar{b}_{55}, &&s_{22}^B = \bar{b}_{22} + s_{23}^B\bar{b}_{23} + s_{12}^B\bar{b}_{12}, &&s_{23}^B = (-\bar{b}_{12}\bar{b}_{13} + \bar{b}_{23}\bar{b}_{11})/\delta_2,\\
&d_{31}^B = (-\bar{b}_{36}\bar{b}_{13} + \bar{b}_{16}\bar{b}_{33})/\delta_2, &&d_{32}^B = \bar{b}_{26} + d_{31}^B\bar{b}_{12} + d_{33}^B\bar{b}_{23}, &&d_{33}^B = (-\bar{b}_{16}\bar{b}_{13} + \bar{b}_{36}\bar{b}_{11})/\delta_2,\\
&\delta_2 = \bar{b}_{33}\bar{b}_{11} - \bar{b}_{13}^2, &&q_3^B = \bar{b}_{16}\alpha_1^B + \bar{b}_{36}\alpha_3^B + \bar{b}_{t6}, &&\epsilon_{33}^B = \bar{b}_{66} + d_{31}^B\bar{b}_{16} + d_{33}^B\bar{b}_{36},\\
&\alpha_1^B = (-\bar{b}_{33}\bar{b}_{t1} + \bar{b}_{13}\bar{b}_{t3})/\delta_2, &&\alpha_2^B = \bar{b}_{21}\alpha_1^B + \bar{b}_{23}\alpha_3^B + \bar{b}_{t2}, &&\alpha_3^B = (\bar{b}_{13}\bar{b}_{t1} - \bar{b}_{11}\bar{b}_{t3})/\delta_2,
\end{aligned}
\tag{16}
$$

2.4. Combined model. For the case of the combined model AB, the material properties of the piezoelectric fiber phase in model B are replaced with the effective properties from model A. Since the shear stresses are uncoupled, it is readily possible to obtain the closed-form expressions for the effective s_{44} and s_{66} from model AB using (9) and (16) as

$$
s_{44}^{AB} = v_f s_{44}^p + v_m s_{44}^m, \qquad s_{66}^{AB} = \frac{v_2^p s_{66}^p s_{66}^m + v_2^m s_{66}^m (v_3^p s_{66}^m + v_3^m s_{66}^p)}{(v_3^p s_{66}^m + v_3^m s_{66}^p)},
\tag{17}
$$

where $v_f = v_2^p v_3^p$, $v_m = 1 - v_f$. Similarly, s_{55}^{AB} can be obtained as

$$
s_{55}^{AB} = \frac{(v_3^p s_{55}^p + v_3^m s_{55}^m)s_{55}^m}{v_2^p s_{55}^m + v_2^m (v_2^p s_{55}^m + v_2^m s_{55}^p)}.
\tag{18}
$$

Usually v_2^p is taken as equal to v_3^p, in which case $v_2^p = v_3^p = \sqrt{v_f}$, where v_f is the overall volume fraction of fiber.

For fiber and matrix phases which are transversely isotropic about the fiber direction x_1, the effective material properties of the combined model should exhibit transverse isotropy. However, it can be seen from (17) and (18) that s_{66}^{AB} and s_{55}^{AB} will not be equal for this case. A similar inequality is observed among the other pairs of effective constants such as s_{12} and s_{13}, and s_{22} and s_{33}, which should also be equal for the transversely isotropic case. This discrepancy exists in the electromechanical micromechanics model of [Bent 1994]. To eliminate it, the effective properties of the combined model BA are obtained by replacing the properties of the fiber phase of model A with the effective properties from model B. The final effective properties P_i^e are then obtained by averaging those of models AB and BA:

$$P_i^e = (P_i^{AB} + P_i^{BA})/2. \tag{19}$$

While the expression for s_{44}^{BA} is the same as for model AB given in (17), the expressions for s_{55}^{BA} and s_{66}^{BA} are obtained by interchanging their expressions for model AB given by (17) and (18).

2.5. *Thermal conductivity.* The 3D heat conduction, according to Fourier's law, is governed by

$$Q_i = -k_i T_{,i} \quad \text{for} \quad i = 1, 2, 3, \tag{20}$$

where k_i, Q_i, and $T_{,i}$ denote respectively the thermal conductivities, heat flux, and temperature gradient along the x_i direction. To obtain the effective thermal conductivities using the isofield method, the temperature gradients parallel to the connecting plane of two phases and the heat flux along the normal to the connecting plane are assumed to be uniform over the two phases. The remaining complementary field variables are averaged over the two phases. For models A and B, this yields:

Model A:

$$\begin{aligned}
\bar{H}^a &= H_p^a = H_m^a, & H^a &= \left[T_{,1}, T_{,2}, Q_3\right]^{\mathrm{T}}, \\
\bar{M}^a &= v_3^p M_p^a + v_3^m M_m^a, & M^a &= \left[Q_1, Q_2, T_{,3}\right]^{\mathrm{T}}.
\end{aligned} \tag{21}$$

Model B:

$$\begin{aligned}
\bar{H}^b &= H_p^b = H_m^b, & H^b &= \left[T_{,1}, Q_2, T_{,3}\right]^{\mathrm{T}}, \\
\bar{M}^b &= v_3^p M_p^b + v_3^m M_m^b, & M^b &= \left[Q_1, T_{,2}, Q_3\right]^{\mathrm{T}}.
\end{aligned} \tag{22}$$

Using (20), the dependent variables M_γ^a and M_γ^b ($\gamma = p, m$) are expressed, respectively, in terms of the independent variables H_γ^a and H_γ^b, and the resulting equations are arranged in the form of (20), to yield the effective thermal conductivities for models A and B. As before, the effective thermal conductivities k_i^{AB} of the combined model AB are then obtained by using the effective thermal conductivities obtained from model A in model B as the conductivities of its fiber phase, which yields

$$k_1^{AB} = v_f k_1^p + v_m k_1^m, \quad k_2^{AB} = \frac{v_3^p k_2^p k_2^m + v_3^m (k_2^m)^2}{v_2^p k_2^m + v_2^m (v_3^p k_2^p + v_3^m k_2^m)}, \quad k_3^{AB} = \frac{v_2^p k_3^p k_3^m + v_2^m k_3^m (k_3^m v_3^p + v_3^m k_3^p)}{v_3^p k_3^m + v_3^m k_3^p}. \tag{23}$$

In the combined model BA, the expression for the effective k_1 is the same as for model AB, and the expressions for the effective k_2 and k_3 get interchanged. The final effective k_i^e are obtained using (19).

3. Results and discussion

Numerical results for the effective thermoelectroelastic properties are presented for two PFRC systems made, respectively, of PZT-7A and PZT-5H fibers and an epoxy matrix. The material properties of the fibers and the matrix are listed in Table 1. The following nondimensional parameters are introduced to compare the effective properties of the PFRC materials with the corresponding piezoelectric fibers:

$$R^e_{3i} = \frac{e^e_{3i}}{e^p_{3i}}, \qquad R^d_{3i} = \frac{d^e_{3i}}{d^p_{3i}}, \qquad \bar{\alpha}_i = \frac{\alpha^e_i}{\alpha^p_i}, \qquad \bar{k}_i = \frac{k^e_i}{k^p_i} \quad \text{for } i = 1, 2, 3;$$

$$\bar{s}_{ij} = \frac{s^e_{ij}}{s^p_{ij}} \quad \text{for } i = 1, 2, 6; \qquad R^\epsilon_{33} = \frac{\epsilon^e_{33}}{\epsilon^p_{33}}, \qquad R^q_3 = \frac{q^e_3}{q^p_3}.$$

The effective elastic stiffness constants c^e_{ij} of the PZT-7A/epoxy system are plotted in Figure 3 against the fiber volume fraction v_f and compared with those predicted by [Ray 2006]. Due to two-way electromechanical coupling, the effective stiffness is affected by the piezoelectric coupling constants d_{ij}. In order to ascertain this effect, the elastic constants c^e_{ij} computed considering $d_{ij} = 0$ are also compared in Figure 3. It is observed that the electromechanical coupling has a stiffening effect on PFRC resulting in greater values for constants c^e_{11}, c^e_{33}, and c^e_{12}. While the c^e_{11} predicted by [Ray 2006] match closely with the present estimate with $d_{ij} = 0$, there are appreciable differences between the two results for c^e_{33}, c^e_{12}, and c^e_{23}. The difference increases for all constants, when the d_{ij} are not considered zero in the present model.

The variations of nondimensional effective piezoelectric stress constants R^e_{3i} with fiber volume fraction are plotted in Figure 4 for the same PFRC system for different values of the dielectric ratio (DR $= \epsilon^p_{33}/\epsilon^m_{33}$). The DR was varied by varying the matrix property (ϵ^m_{33}) keeping the fiber property fixed. The case of DR $= 1$ leads to uniform electric field E_3 across both fiber and matrix phases, a condition assumed by

Material	c_{11}	c_{22}	c_{33}	c_{12}	c_{23}	c_{31}	c_{44}	c_{55}	c_{66}
PZT-7A[1]	148	148	131	76.2	74.2	74.2	25.4	25.4	35.9
PZT-5H[2]	126	126	117	79.5	84.1	84.1	23	23	23.25
epoxy[1]	3.86	3.86	3.86	2.57	2.57	2.57	0.645	0.645	0.645

	e_{31}	e_{32}	e_{33}	e_{14}	e_{24}			ϵ_{33}	
PZT-7A[1]	–2.1	–2.1	9.5	9.2	9.2			2.07	
PZT-5H[2]	–6.5	–6.5	23.3	17	17			30.42	
epoxy[1]	0	0	0	0	0			0.079	

	α_1	α_2	α_3	k_{11}	k_{22}	k_{33}		p_3	
PZT-7A[1]	1	1	1	–	–	–		2	
PZT-5H[3]	9.64	9.64	3.96	50	50	75		5.483	
epoxy[1,4]	24	24	24	0.18	0.18	0.18		0.0	

Table 1. Material properties: c_{ij} in GPa, e_{ij} in $\mathrm{C\,m^{-2}}$, ϵ_{33} in $10^{-9}\,\mathrm{C\,V^{-1}\,m^{-1}}$, α_i in $10^{-6}\,\mathrm{K^{-1}}$, k_{ij} in $\mathrm{W\,K^{-1}\,m^{-1}}$, p_i in $10^{-5}\,\mathrm{C\,m^2\,K^{-1}}$.
[1] [Kumar and Chakraborty 2009]   [2] [Kapuria and Hagedorn 2007]   [3] [Chen 2006]
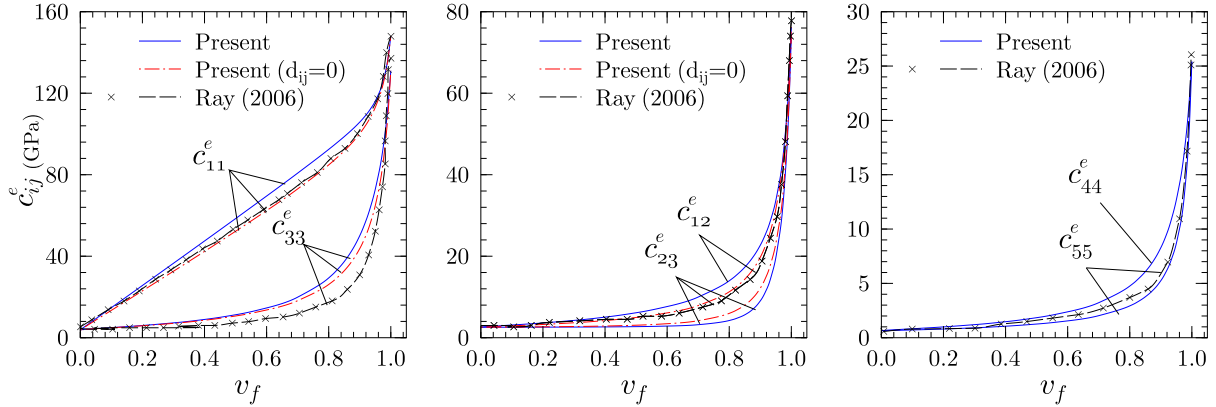[4] [Gibson 2007, Table 3.2, p. 106].

Figure 3. Variation of effective stiffness constants of PZT-7A/epoxy with fiber volume fraction.

[Ray 2006]. While the value of R^e_{31} predicted by the present model for DR $= 1$ is in close agreement with the reference solution, values of R^e_{32} and R^e_{33} predicted by the later model are much lower than the present solution. In the present model, ε_2 and σ_3 are uniform over two phases in model A, while σ_2 and ε_3 are uniform in model B, according to the respective connecting planes, which are consistent with the actual boundary conditions of the RVE. However, in the model of [Ray 2006], both σ_2 and σ_3 are assumed to have a uniform variation over the two phases, simultaneously. This explains the difference between the two results even though both correspond to the uniform electric field case. Figure 4 also reveals that the effective piezoelectric constants of PFRC are highly sensitive to the DR and reduce drastically with its increase. This is because the ratio of electric field across the fiber phase to that applied across the PFRC thickness decreases with the increase in DR. Even though both constituents exhibit transverse isotropy about axis x_3, effective values of e_{31} and e_{32} differ, since the connectivity between the fiber and matrix phases does not follow symmetry about axis x_3. This difference between e_{31} and e_{32} in PFRC enables directional (anisotropic) in-plane actuation, which is desirable in many control applications. The effective thermal stress coefficients β_1 and β_3 for the PZT-7A/epoxy system for varying v_f, computed with and without making $d_{ij} = 0$, are compared with those of [Kumar and Chakraborty 2009] in Figure 5. Once again, the present results match with the reference solution, when computed with $d_{ij} = 0$, but otherwise differ considerably for intermediate values of v_f.

In Figures 6 and 7, nondimensionalized compliances ($\bar{s}_{11}, \bar{s}_{12}, \bar{s}_{66}$) and coefficients of thermal expansion ($\bar{\alpha}_1, \bar{\alpha}_2$) are compared with the simple rule of mixtures (ROM)/inverse rule of mixtures (IROM) and the modified rule of mixtures (MROM) given in Appendix B. The effective values of the thermoelastic constants predicted by the present model are close to those predicted by the ROM for longitudinal constants s^e_{11} and α^e_1 and by the MROM for the transverse constants s^e_{66} and α^e_2, but are not so for the transverse compliance constants s^e_{22} and s^e_{33}. The nondimensionalized effective thermal conductivities $\bar{k}^e_1$ and $\bar{k}^e_3$ are also plotted in Figure 7.

The variations of the ratios R^d_{3i} ($i = 1, 2, 3$) of effective values of piezoelectric strain constants d_{3i} to the corresponding values for the bulk PZT are plotted in Figure 8 for both PZT-7A/epoxy and PZT-5H/epoxy systems for different values for DR ranging from 1 to 100. Similar to constants e_{3i}, the effective values of d_{3i} are maximum for DR $= 1$, when the electric field is uniform across piezoelectric and matrix phases, and decrease sharply with the increase in DR, as happens for commonly used matrix materials. Even
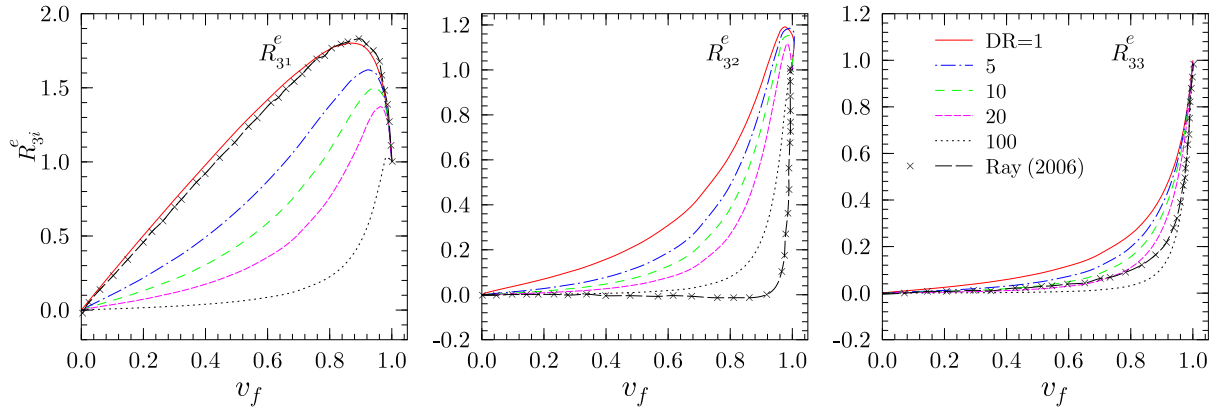
Figure 4. Variation of effective piezoelectric stress constant ratios of PZT-7A/epoxy with fiber volume fraction.
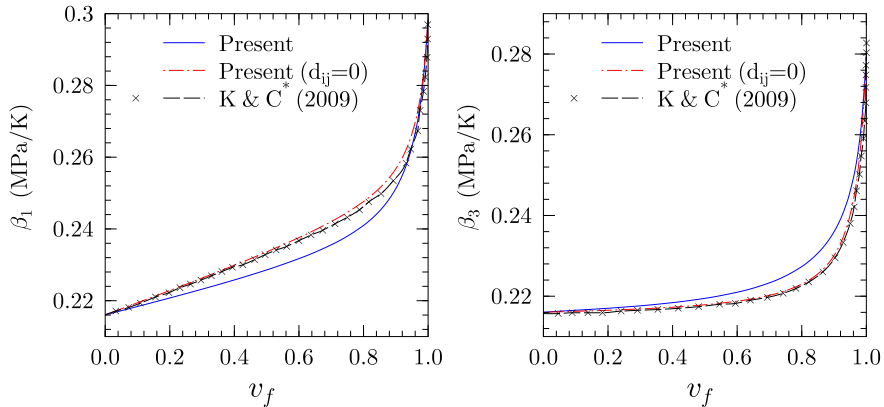

Figure 5. Variation of effective thermal stress coefficients of PZT-7A/epoxy with fiber volume fraction.
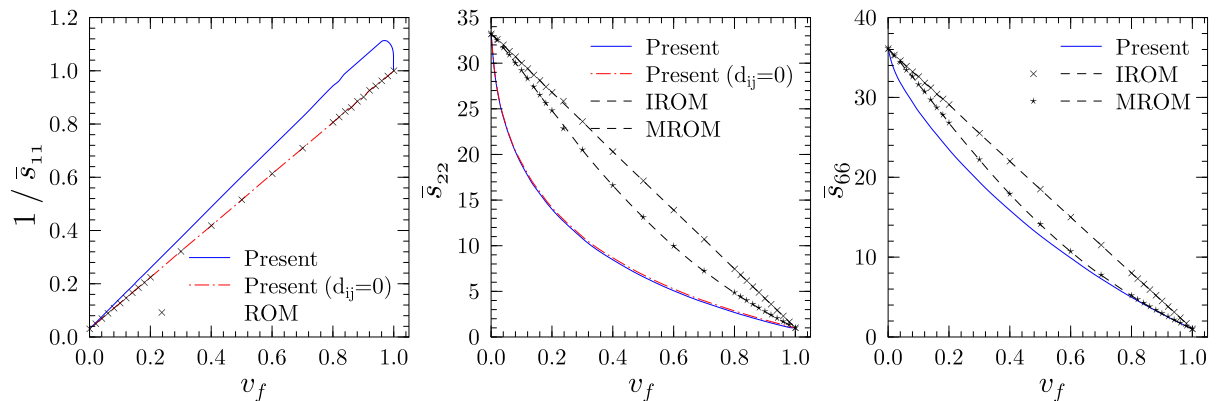
* [Kumar and Chakraborty 2009].

Figure 6. Variation of effective compliance coefficients of PZT-5H/epoxy with fiber volume fraction.
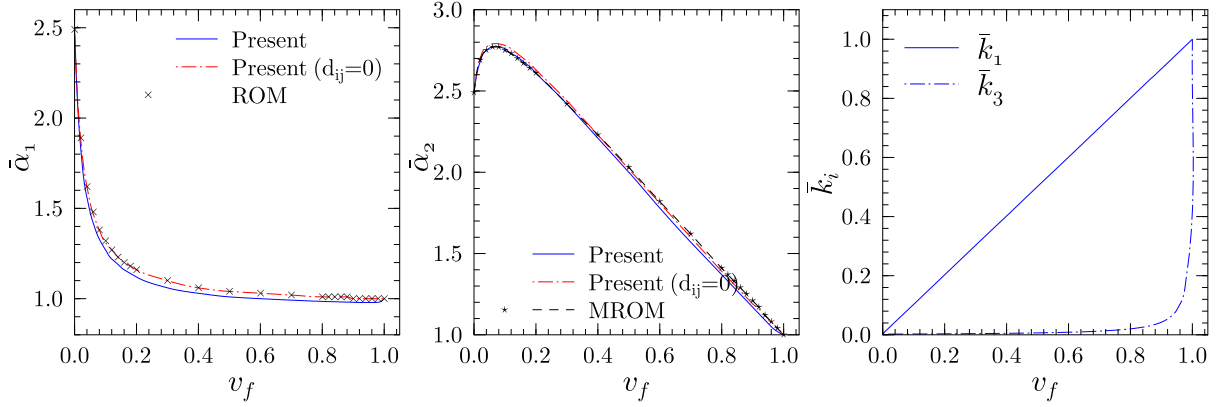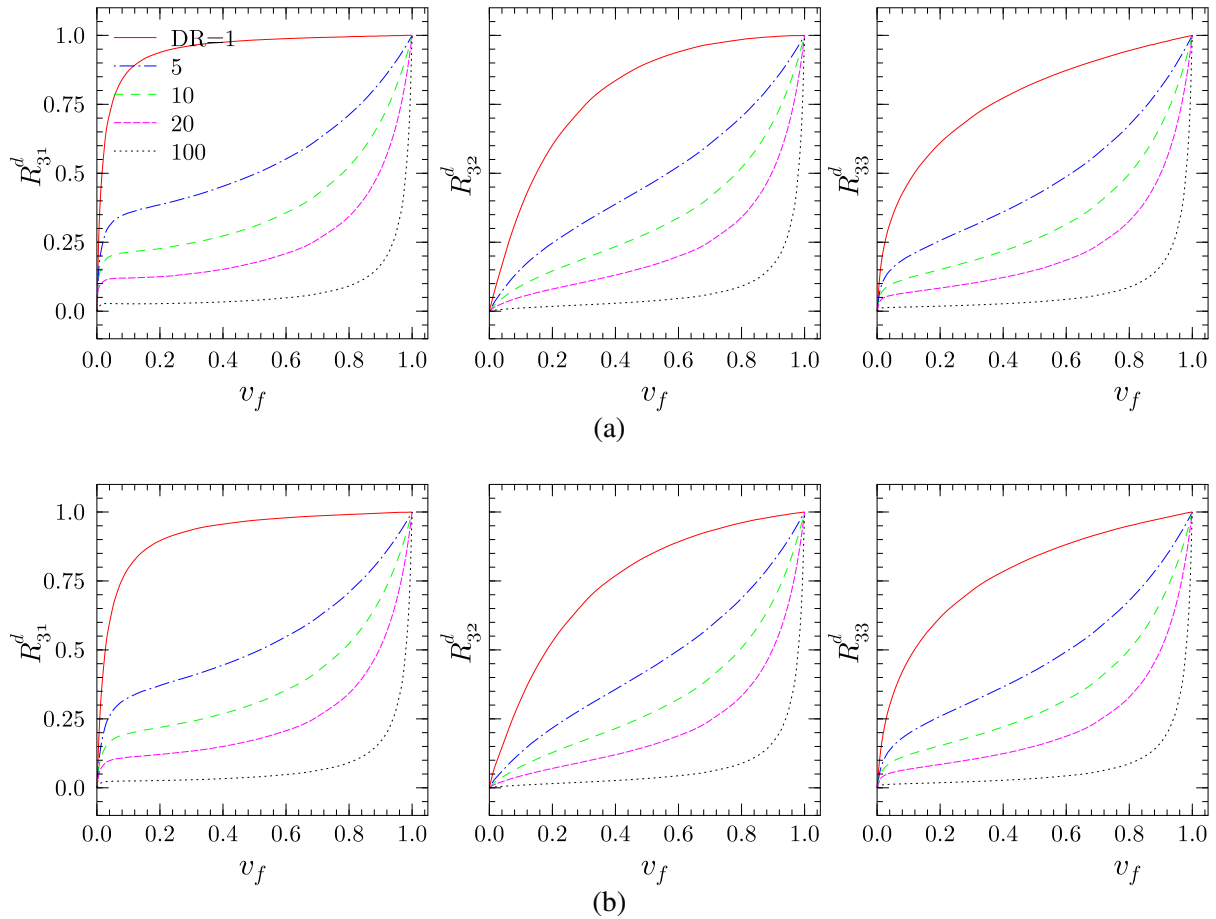
Figure 7. Variation of effective thermal expansion coefficients and conductivities of PZT-5H/epoxy with fiber volume fraction.

Figure 8. Effect of DR on effective piezoelectric strain constants of PFRCs, for (a) PZT-7A/epoxy and (b) PZT-5H/epoxy.

at a low fiber volume fraction of 10%, there is a large difference between the results of d_{3i}^e at DR $= 1$ and DR $= 5$. This is because the electric field E_3^p across the piezoelectric phase in model A is only a fraction of the overall electric field $\bar{E}_3$ and is dependent on both v_3^p and DR. For example, ignoring the electromechanical coupling, E_3^p in model A can be easily obtained, using (2)–(8), as

$$\frac{E_3^p}{\bar{E}_3} = \frac{1}{v_3^p + v_3^m(\text{DR})}.$$

Thus, E_3^p is more sensitive to DR at a lower value of the fiber volume fraction. The performance of these PFRCs can thus be improved either by using suitable matrix materials with dielectric constant ϵ_{33}^m of the order of ϵ_{33}^p or by directly applying electric fields across the piezoelectric fibers. It is also revealed from Figure 8 that above 90% of the value of d_{31} of the bulk PZT can be achieved in the PFRC with a fiber volume fraction of only 25% for the uniform field case. While the values of effective piezoelectric strain constants d_{3i}^e vary from zero to those of the piezoelectric fibers, the effective values of piezoelectric stress constants e_{3i} can exceed those of the bulk piezoelectric material at an intermediate value of v_f (Figure 4). The latter leads to an impression that a higher electromechanical coupling can be achieved in PFRC than the bulk piezoelectric material [Ray 2006], which is clearly not true. Thus, the fundamental constants d_{ij} (and not e_{ij}) should be used for evaluating the effective electromechanical coupling property of PFRC.

The variations of nondimensionalized effective dielectric and pyroelectric constants with v_f are plotted in Figure 9. The effective pyroelectric constant almost follows the ROM for DR of unity, but
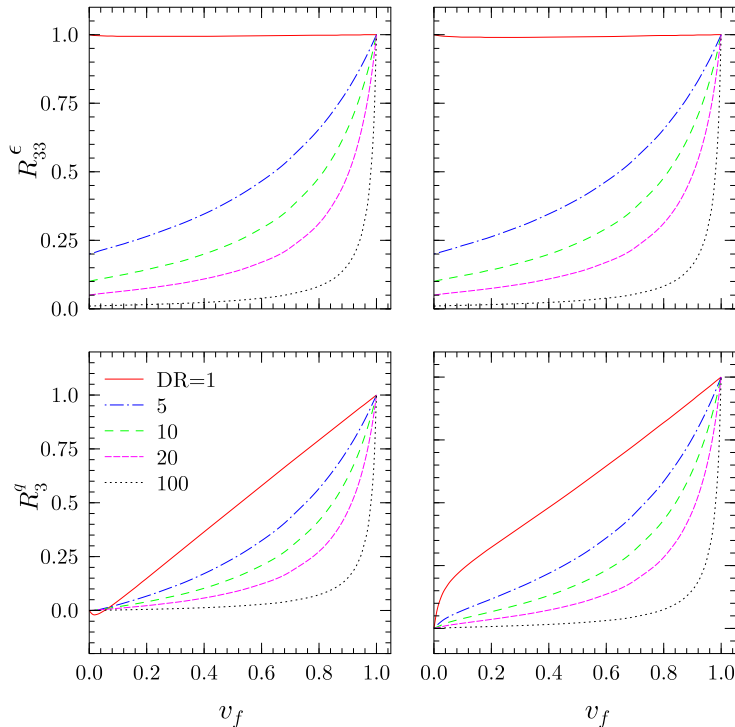

Figure 9. Variation of effective dielectric and pyroelectric constants of PFRCs with fiber volume fraction: PZT-7A/epoxy (left column) and PZT-5H/epoxy (right column).

Entity	PZT-5H/epoxy				PZT-7A/epoxy			
	0.2	0.4	0.6	0.8	0.2	0.4	0.6	0.8
s_{11}^e	65.12	34.76	23.91	18.56	43.25	22.59	15.39	11.83
s_{22}^e	233.7	141.3	85.06	45.60	222.8	132.8	77.99	39.37
s_{33}^e	225.1	135.1	80.58	43.16	212.6	125.3	72.20	35.28
s_{12}^e	−30.39	−16.63	−11.27	−8.15	−21.75	−11.66	−7.79	−5.61
s_{13}^e	−20.84	−10.33	−7.08	−6.07	−10.73	−4.39	−2.60	−2.11
s_{23}^e	−89.88	−39.37	−17.06	−8.07	−86.29	−35.26	−13.16	−4.28
s_{44}^e	1250	950	650	345	1250	950	640	340
s_{55}^e	1010	680	420	215	1010	680	420	210
s_{66}^e	1010	680	420	214	1000	670	410	200
d_{31}^e	−33.15	−41.25	−57.01	−94.32	−5.24	−6.39	−8.88	−15.14
d_{32}^e	−19.33	−33.07	−51.94	−91.40	−3.36	−5.48	−8.45	−14.99
d_{33}^e	50.38	73.52	111.2	195.3	8.72	12.70	19.56	35.81
ϵ_{33}^e	2.27	3.30	5.16	9.49	0.12	0.18	0.28	0.53
q_3^e	−0.99	−2.00	−3.74	−7.78	0.44	1.15	2.47	5.71
α_1^e	11.07	10.24	9.94	9.77	3.09	2.19	1.84	1.59
α_2^e	25.33	21.53	17.46	13.48	25.12	19.06	12.93	7.01
α_3^e	22.02	17.03	12.32	7.96	24.12	17.55	11.23	5.40
k_1^e	11.20	20.90	30.60	40.30	1.56	1.62	1.68	1.74
k_2^e	0.78	1.18	1.95	4.14	0.63	0.80	1.01	1.33
k_3^e	0.79	1.19	1.97	4.26	0.63	0.80	1.01	1.33

Table 2. Effective electrothermoelastic properties for PFRCs: s_{ij}^e in $10^{-12}\,\mathrm{Pa}^{-1}$, d_{3j}^e in pm V^{-1}, ϵ_{33}^e in $\mathrm{nF\,m}^{-1}$, q_3^e in $\mu\mathrm{C\,m}^{-2}\,\mathrm{K}^{-1}$, α_i^e in $10^{-6}\,\mathrm{K}^{-1}$, k_i^e in $\mathrm{W\,m}^{-1}\,\mathrm{K}^{-1}$.

drops sharply as the DR increases. Effective material properties for PZT-5H/epoxy with DR $= 20$ and PZT-7A/epoxy with DR $= 26.2$ are presented in Table 2 for four values of the fiber volume fraction ($v_f = 0.2, 0.4, 0.6, 0.8$). These properties are directly useful for 1D/2D/3D thermoelectromechanical analysis of smart laminated structures integrated with these PFRCs.

4. Conclusions

A coupled 3D isofield-based micromechanical model is presented for calculating effective electrother-moelastic properties of piezoelectric fiber-reinforced composite (PFRC) materials with poling and elec-tric field applied along the normal to the fiber direction. The model employs the isofield method on representative volume elements (RVEs) of two possible connectivities, A and B, for the fiber and matrix phases. The two RVEs are combined in sequences AB and BA, so as to achieve transverse isotropy in the effective properties when both the constituents are transversely isotropic. The model considers differen-tial electric fields in fiber and matrix phases due to their different dielectric constants. The assumption of uniform electric field across the two phases made by some existing theories can be achieved as a special case of the present formulation when the dielectric ratio (DR) is unity.

Results presented for two PFRC systems, PZT-7A/epoxy and PZT-5H/epoxy, reveal that it is possible to achieve effective piezoelectric strain constant d_{31} of magnitude greater than 90% of that of bulk piezoelectric material with even a low fiber volume fraction of 25%, if the DR is close to unity. The effective piezoelectric constants, however, reduce drastically as the DR increases. While the magnitudes of effective d_{3i} of PFRC vary from zero to those of the piezoelectric fiber, the effective e_{3i} can exceed its value for the fiber at an intermediate value of v_f, giving a false notion of improved performance, for DR close to unity. Therefore, constants d_{ij} should be used to evaluate electromechanical coupling of PFRCs. The effective pyroelectric constant nearly follows the rule of mixtures for DR of unity, but reduces as DR increases.

Appendix A

The isofield micromechanics model, also known as the uniform field model, is basically a generalization of the strength of materials approach of estimating overall properties of two-phase composite materials. The basic assumption, as the name implies, is that all fields are uniform within each material phase. This assumption makes the model independent of the geometry of its microstructure, and it is possible to refer to the two phases in a RVE as two cuboidal blocks connected at a plane. This forms the basis of the RVE representation in Figure 2. The model actually violates some of the compatibility and equilibrium conditions at the interface. However, it has been successfully used in the past (see, for example, [Jones 1975]) for estimating mechanical and transport (conductivity, thermal expansion coefficient) properties of fiber-reinforced composite materials. The large mismatch in the material properties makes the method particularly well suited for modeling such fibrous composites. Its accuracy largely depends on the spatial scale of the phase distributions. For finely distributed phases, it is expected to yield good estimates of the averaged response. A major advantage of this method is that it provides simple analytical solutions for the effective properties, which allows physical insight into the problem.

For uniaxial fields, the uniform field assumption leads to the well-known rules of mixtures, comprising of parallel and series (Voigt and Reuss) additions for estimating the effective properties along and perpendicular to the fiber direction, respectively. For the 3D fields, however, it is not as straightforward. Typically, the conventional model with parallel connectivity yields very good estimates of longitudinal properties, but the series model at best yields a lower bound of the transverse properties. This is improved by using a combination model, wherein the two phases in the RVE are connected at two orthogonal planes (x_1-x_2 and x_1-x_3) parallel to the fiber direction (x_1), instead of only one plane as in the conventional model, as shown in Figure 2 (combined model). The combination model is treated as follows: blocks of piezoelectric (p) and matrix (m) phases are connected at the x_1-x_2 plane and they together are connected to a matrix phase at the x_1-x_3 plane, and vice versa. Thus, models A and B with the two phases connected at the x_1-x_2 and x_1-x_3 planes, respectively, form the building blocks of the combination model.

The other assumptions in the present model are that

- the two material phases are perfectly bonded,

- deformation and electric fields are small enough that linear constitutive equations can be applied, and

- the piezoelectric materials are uniformly polarized along the x_3 direction.

Appendix B

The expressions for the effective Young's moduli Y_i^e, shear modulus G_{12}^e, and thermal expansion coefficients α_i^e according to the rule of mixtures (ROM)/inverse rule of mixtures (IROM) and the modified rule of mixtures (MROM) based on the strength of materials approach can be found in [Gibson 2007]. These are given below:

- ROM/IROM

$$Y_1^e = Y_1^p v_f + Y_1^m v_m, \qquad \alpha_1^e = \frac{Y_1^p \alpha_1^p v_f + Y_1^m \alpha_1^m v_m}{Y_1^p v_f + Y_1^m v_m},$$

$$\frac{1}{Y_2^e} = \frac{v_f}{Y_2^p} + \frac{v_m}{Y_2^m}, \qquad \frac{1}{G_{12}^e} = \frac{v_f}{G_{12}^p} + \frac{v_m}{G_{12}^m}.$$

- MROM

$$Y_2^e = Y_2^m \left[(1 - \sqrt{v_f}) + \frac{\sqrt{v_f}}{1 - \sqrt{v_f}(1 - Y_2^m/Y_2^p)} \right],$$

$$G_{12}^e = G_{12}^m \left[(1 - \sqrt{v_f}) + \frac{\sqrt{v_f}}{1 - \sqrt{v_f}(1 - G_{12}^m/G_{12}^p)} \right],$$

$$\alpha_2^e = (1 + v_m)\alpha_2^m v_m + (1 + v_p)\alpha_2^p v_f - \alpha_1 v_{12}^e, \quad \text{where} \quad v_{12}^e = v_{12}^p v_f + v^m v_m.$$

Acknowledgements

References

[Aboudi 1998] J. Aboudi, "Micromechanical prediction of the effective coefficients of thermo-piezoelectric multiphase composites", *J. Intell. Mater. Syst. Struct.* **9**:9 (1998), 713–722.

[Auld 1973] B. A. Auld, *Acoustic fields and waves in solids*, vol. 1, Wiley, New York, 1973.

[Bent 1994] A. A. Bent, "Piezoelectric fiber composites for structural actuation", Master's thesis, Massachusetts Institute of Technology, Cambridge, MA, 1994, Available at http://hdl.handle.net/1721.1/12322.

[Chan and Unsworth 1989] H. L. W. Chan and J. Unsworth, "Simple model for piezoelectric ceramic/polymer 1-3 composites used in ultrasonic transducer applications", *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **36**:4 (1989), 434–441.

[Chen 2006] C.-D. Chen, "On the singularities of the thermo-electro-elastic fields near the apex of a piezoelectric bonded wedge", *Int. J. Solids Struct.* **43**:5 (2006), 957–981.

[Chopra 2002] I. Chopra, "Review of state of art of smart structures and integrated systems", *AIAA J.* **40**:11 (2002), 2145–2187.

[Dong et al. 2006] X.-J. Dong, G. Meng, and J.-C. Peng, "Vibration control of piezoelectric smart structures based on system identification technique: numerical simulation and experimental study", *J. Sound Vib.* **297**:3–5 (2006), 680–693.

[Dunn and Taya 1993] M. L. Dunn and M. Taya, "Micromechanics predictions of the effective electroelastic moduli of piezoelectric composites", *Int. J. Solids Struct.* **30**:2 (1993), 161–175.

[Gibson 2007] R. F. Gibson, *Principles of composite material mechanics*, 2nd ed., CRC Press, Boca Raton, FL, 2007.

[Jones 1975] R. M. Jones, *Mechanics of composite materials*, Hemisphere, New York, 1975.

[Kapuria and Hagedorn 2007] S. Kapuria and P. Hagedorn, "Unified efficient layerwise theory for smart beams with segmented extension/shear mode, piezoelectric actuators and sensors", *J. Mech. Mater. Struct.* **2**:7 (2007), 1267–1298.

[Kapuria and Kumari 2010] S. Kapuria and P. Kumari, "Three-dimensional piezoelasticity solution for dynamics of cross-ply cylindrical shells integrated with piezoelectric fiber reinforced composite actuators and sensors", *Compos. Struct.* **92**:10 (2010), 2431–2444.

[Kar-Gupta and Venkatesh 2005] R. Kar-Gupta and T. A. Venkatesh, "Electromechanical response of 1-3 piezoelectric composites: effect of poling characteristics", *J. Appl. Phys.* **98**:5 (2005), 054102.

[Kumar and Chakraborty 2009] A. Kumar and D. Chakraborty, "Effective properties of thermo-electro-mechanically coupled piezoelectric fiber reinforced composites", *Mater. Des.* **30**:4 (2009), 1216–1222.

[Levin et al. 2000] V. M. Levin, T. Michelitsch, and I. Sevostianov, "Spheroidal inhomogeneity in a transversely isotropic piezoelectric medium", *Arch. Appl. Mech.* **70**:10 (2000), 673–693.

[Levin et al. 2008] V. M. Levin, F. J. Sabina, J. Bravo-Castillero, R. Guinovart-Díaz, R. Rodríguez-Ramos, and O. C. Valdiviezo-Mijangos, "Analysis of effective properties of electroelastic composites using the self-consistent and asymptotic homogenization methods", *Int. J. Eng. Sci.* **46**:8 (2008), 818–834.

[Mallik and Ray 2003] N. Mallik and M. C. Ray, "Effective coefficients of piezoelectric fiber-reinforced composites", *AIAA J.* **41**:4 (2003), 704–710.

[Park et al. 2010] S. Park, C. Lee, and H. Sohn, "Reference-free crack detection using transfer impedances", *J. Sound Vib.* **329**:12 (2010), 2337–2348.

[Ray 2006] M. C. Ray, "Micromechanics of piezoelectric composites with improved effective piezoelectric constant", *Int. J. Mech. Mater. Des.* **3**:4 (2006), 361–371.

[Sabina et al. 2001] F. J. Sabina, R. Rodríguez-Ramos, J. Bravo-Castillero, and R. Guinovart-Díaz, "Closed-form expressions for the effective coefficients of a fibre-reinforced composite with transversely isotropic constituents, II: Piezoelectric and hexagonal symmetry", *J. Mech. Phys. Solids* **49**:7 (2001), 1463–1479.

[Smith and Auld 1991] W. A. Smith and B. A. Auld, "Modeling 1-3 composite piezoelectrics: thickness-mode oscillations", *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **38**:1 (1991), 40–47.

[Uchino 2000] K. Uchino, *Ferroelectric devices*, Marcel Dekker, New York, 2000.

SANTOSH KAPURIA: kapuria@am.iitd.ac.in
Department of Applied Mechanics, Indian Institute of Technology Delhi, Hauz Khas, New Delhi 110016, India
http://web.iitd.ac.in/~am/

POONAM KUMARI: kpmech.iitd@gmail.com
Department of Applied Mechanics, Indian Institute of Technology Delhi, Hauz Khas, New Delhi 110016, India

∎msp

ACCURATE SIMULATION OF MIXED-MODE COHESIVE CRACK PROPAGATION IN QUASI-BRITTLE STRUCTURES USING EXACT ASYMPTOTIC FIELDS IN XFEM: AN OVERVIEW

BHUSHAN LAL KARIHALOO AND QI-ZHI XIAO

The extended finite element (XFEM) enriches the standard local FE approximations with known information about the problem, with the use of the partition of unity. This allows the use of meshes that do not conform to a discontinuity and avoids adaptive re-meshing as the discontinuity grows as required with the conventional FEM. When the crack tip asymptotic field is available and used as the enrichment function, XFEM is more accurate than FEM allowing the use of a much coarser mesh around the crack tip. Such asymptotic fields have been known for a long time for traction-free cracks (the Williams expansions) but have only recently been derived for cohesive cracks (Karihaloo–Xiao expansions). In this paper an overview of latter expansions is given for a range of cohesive laws and their usefulness in the simulation of cohesive crack propagation is demonstrated on two examples of concrete and fibre-reinforced concrete flexural members.

1. Introduction

The cohesive zone (or crack) model of Hillerborg et al. [1976] has been extensively used in the study of localization and failure in quasi-brittle materials (such as concrete and fibre-reinforced concrete) and structures. Borst et al. [2004] have given a concise overview of the various ways for the numerical implementation of the cohesive zone methodology.

The knowledge of the asymptotic crack tip displacement fields is especially useful in the recently developed extended finite element methodology (XFEM) (see [Moës et al. 1999; Strouboulis et al. 2001; Babuška et al. 2003; Karihaloo and Xiao 2003], for example). XFEM enriches the standard local FE approximations with known information about the problem, with the use of the partition of unity (PU). It avoids meshes conforming with the discontinuity and adaptive re-meshing as the discontinuity grows as is the case with the FEM. In [Karihaloo and Xiao 2003] we demonstrated that for a crack with traction-free faces, when the crack tip asymptotic field is available and used as enrichment function, XFEM not only avoids using a mesh conforming with the crack but is also more accurate than FEM. Hence XFEM can use a much coarser mesh around the crack tip. However, when the enrichment function differs from the true asymptotic crack tip field, the mesh needs to be refined in the same manner as in the FEM. Thus it is necessary to know the true asymptotic displacement field around a cohesive crack tip in order to exploit fully the advantages of XFEM.

The application of XFEM to the simulation of the growth of cohesive cracks in quasi-brittle materials has received considerable attention. However, adjacent to the cohesive crack tip, the enrichment function is chosen as the jump function [Wells and Sluys 2001; Hansbo and Hansbo 2004; Zi and Belytschko 2003]

or a (branch) function which does not represent the true asymptotic nature of the displacement/stress field there [Moës and Belytschko 2002]. Since there is no singularity at the tip of a cohesive crack, a stress criterion is often used to judge the initiation and propagation of the crack. Therefore a reliable analysis of cohesive crack propagation requires an accurate knowledge of the crack tip field. However, although no singularity exists at the tip of a cohesive crack, the stresses obtained by direct differentiation of the displacements are not accurate, and cannot be used to predict accurately the growth of the tip, exactly as in the traction-free cracks.

In [Xiao and Karihaloo 2006a; Karihaloo and Xiao 2008; 2010] we have obtained universal asymptotic expansions at a cohesive crack tip, analogous to the Williams expansions at a traction-free crack tip for any normal cohesion-separation law (softening law) that can be expressed in a special polynomial with integer or fractional powers. This special form ensures that the radial and angular variations of the asymptotic fields are separable as in the Williams expansions. The coefficients of the expansions of course depend nonlinearly on the softening law and the boundary conditions. They demonstrated that many commonly used cohesion-separation laws, e.g., rectangular, linear, bilinear and exponential, can indeed be expressed very accurately in this special form. They also obtained universal asymptotic expansions when the cohesive crack faces are subjected to Coulomb friction. In [Liu et al. 2004; Xiao and Karihaloo 2006b; Xiao et al. 2007] we used this true crack tip asymptotic displacement field as a crack tip enrichment function in XFEM for the simulation of mode I cohesive crack propagation in quasi-brittle materials to improve the prediction of the stress field ahead of the cohesive crack even with a coarse mesh.

In cohesive cracks, the friction is considered for a finite opening. In this sense frictional cohesive cracks are different from the frictional contact of crack faces, where the crack faces are in contact and not open. However, in cohesive cracks, although the crack faces are not in contact because of the applied cohesive stresses, frictional forces can come into play between the faces when there is relative sliding. Many studies on the mixed mode cohesive cracks can also be found in the literature, but there is doubt about the accuracy of the cohesion-sliding relation because it is difficult to isolate it from frictional forces between the rough cohesive crack faces in quasi-brittle materials such as concrete.

In this paper, the Karihaloo–Xiao asymptotic expansions are reviewed for a range of cohesive laws and their usefulness in the simulation of cohesive crack propagation by XFEM is demonstrated on two examples of concrete and fibre-reinforced concrete flexural members.

This paper is organised as follows: Section 2 gives cohesive laws of concrete and fibre-reinforced concrete suitable for the asymptotic analysis of cohesive cracks; Section 3 gives a brief overview of the mathematical formulation and boundary conditions and of the asymptotic fields for several cases; Section 4 discusses the implementation of the asymptotic fields in XFEM. Two illustrative examples of concrete and fibre-reinforced concrete flexural members are given and discussed in Section 5.

2. Cohesive laws for concrete and fibre-reinforced concrete

Cornelissen et al. [1986] introduced the following exponential relation to fit their results from uniaxial tests on double edge notched normal and lightweight concrete panels:

$$\frac{\sigma}{f_t} = f\left(\frac{w}{w_c}\right) - \frac{w}{w_c} f(1), \qquad f\left(\frac{w}{w_c}\right) = \left[1 + \left(C_1 \frac{w}{w_c}\right)^3\right] e^{-C_2 w/w_c} \qquad (1)$$

It fits their experimental results with a high degree of accuracy. In (1), σ and f_t are the stress normal to the cohesive crack face and the uniaxial tensile strength, respectively; w and w_c are the opening displacement of the cohesive crack faces, and the critical opening displacement of the pre-existing macrocrack tip at which the cohesive crack tip begins to grow; and C_1 and C_2 are fitting parameters. Details of the test set up as well as the cohesive relation (1) can be found in [Karihaloo 1995]. Other widely used softening laws for concrete are the linear relation

$$\hat{\sigma} = 1 - \hat{w}, \tag{2}$$

the bilinear relation

$$\hat{\sigma}_y = \begin{cases} 1 - (1 - \hat{f}_1)\dfrac{\hat{w}}{\hat{w}_1} & \text{if } 0 \leq \hat{\sigma}_y \leq \hat{f}_1, \\[2mm] \dfrac{\hat{f}_1}{1 - \hat{w}_1}(1 - \hat{w}) & \text{if } \hat{f}_1 < \hat{\sigma}_y \leq 1, \end{cases} \tag{3}$$

and the power-law relationship

$$\hat{\sigma}_y^m + \hat{w}^{2m} = 1. \tag{4}$$

In (2)–(4) we have $\hat{\sigma}_y = \sigma_y/f_t$, $\hat{w} = w/w_c$, $\hat{f}_1 = f_1/f_t$, $\hat{w}_1 = w_1/w_c$, and m ranges from 0.20 to 0.27 for different concrete grades. Moreover, the two linear parts of (3) can be rewritten into two linear laws. The first part can be written into (2) using a new definition of w_c as

$$w_c = \frac{w_1}{1 - \hat{f}_1}. \tag{5}$$

The second part can be written into (2) using a new definition of f_t as

$$f_t = \frac{f_1}{1 - \hat{w}_1}. \tag{6}$$

We shall also consider two further cohesion-separation laws not commonly associated with concrete, namely

$$\hat{\sigma}_y = 1 \tag{7}$$

and

$$\hat{\sigma}_y = 1 - \hat{w}^{2(L+1)}, \quad L = 1, 2, 3, \ldots. \tag{8}$$

The cohesion-separation law (7) is akin to the Dugdale [1960] model of plasticity in thin metallic sheets, which is also frequently used in fibre-reinforced polymeric materials, whereas the law (8) is commonly used in studying the fracture process zone embedded within the large plastic zone in metals.

For the high-performance fibre-reinforced concrete CARDIFRC [Benson and Karihaloo 2005], tensile tests showed that the post-peak cohesion-separation curve can be accurately represented by the polynomial

$$\hat{\sigma}_y = 116.76\hat{w}^7 - 468.08\hat{w}^6 + 738.12\hat{w}^5 - 612.10\hat{w}^4 + 272.71\hat{w}^3 - 59.33\hat{w}^2 + 2.89\hat{w} + 1, \tag{9}$$

in which $f_t = 16\,\text{MPa}$ and $w_c = 6.5\,\text{mm}$.

In [Xiao and Karihaloo 2006a] we showed that the polynomial

$$\hat{\sigma}_y = 1 + \sum_{i=1}^{L} \alpha_i \hat{w}^{(2/3)i} - \left(1 + \sum_{i=1}^{L} \alpha_i\right)\hat{w}^{(2/3)(L+1)} \tag{10}$$

not only fits the experimental results of Cornelissen et al. (Equation (1) above), but also many commonly used cohesion-separation laws, including the rectangular (7), linear (2), and bilinear (3).

The choice of the special form of the cohesion-separation relation (10) involving as it does fractional powers of $\hat{w}$ may seem strange but it ensures that the asymptotic fields at the cohesive crack tip are separable into radial and angular variations exactly as Williams expansions at a traction-free crack tip.

In [Karihaloo and Xiao 2008] we showed that separable asymptotic crack tip fields are also obtainable for the following cohesion-separation relation not involving fractional powers of $\hat{w}$:

$$\hat{\sigma}_y = 1 + \sum_{i=1}^{L} \alpha_i \hat{w}^{2i} - \left(1 + \sum_{i=1}^{L} \alpha_i\right) \hat{w}^{2(L+1)}, \tag{11}$$

where L is an integer.

3. Mathematical formulation

The mathematical formulation is fully described in [Karihaloo and Xiao 2008], so that only a very brief overview will suffice here. For plane problems, the stresses and displacements in the Cartesian coordinate system centred at the cohesive crack tip can be expressed in terms of two analytic functions $\phi(z)$ and $\chi(z)$ of the complex variable $z = re^{i\theta}$:

$$\sigma_x + \sigma_y = 2\left[\phi'(z) + \overline{\phi'(z)}\right],$$
$$\sigma_x - \sigma_y + 2i\tau_{xy} = 2\left[\bar{z}\phi''(z) + \chi''(z)\right], \tag{12}$$
$$2\mu(u + iv) = \kappa\phi(z) - z\overline{\phi'(z)} - \overline{\chi'(z)},$$

where a prime denotes differentiation with respect to z and an overbar the complex conjugate. In (12), $\mu = E/[2(1+\nu)]$ is the shear modulus; the Kolosov constant is $\kappa = 3 - 4\nu$ for plane strain or $\kappa = (3-\nu)/(1+\nu)$ for plane stress; E and ν are Young's modulus and Poisson's ratio, respectively.

For a general plane mixed mode I + II problem, the complex functions $\phi(z)$ and $\chi(z)$ can be chosen as series of complex eigenvalue Goursat functions

$$\phi(z) = \sum_{n=0} A_n z^{\lambda_n} = \sum_{n=0} A_n r^{\lambda_n} e^{i\lambda_n\theta}, \quad \chi(z) = \sum_{n=0} B_n z^{\lambda_n+1} = \sum_{n=0} B_n r^{\lambda_n+1} e^{i(\lambda_n+1)\theta}, \tag{13}$$

where the complex coefficients are $A_n = a_{1n} + ia_{2n}$ and $B_n = b_{1n} + ib_{2n}$. The eigenvalues λ_n and coefficients $a_{1n}, a_{2n}, b_{1n}, b_{2n}$ are real.

Substitution of the complex functions (13) into (12) gives the complete series expansions of the displacements and stresses near the tip of the crack. For details, see [Karihaloo and Xiao 2008].

These solutions need to satisfy the proper symmetry conditions along the line of extension ahead of the cohesive crack, and boundary conditions on the cohesive crack faces. Considering frictional cohesive crack faces, the boundary conditions on the cohesive crack ($0 = r = l_p$) are

$$\sigma_y|_{\theta=\pi} = \sigma_y|_{\theta=-\pi} \neq 0, \quad \tau_{xy}|_{\theta=\pi} = \tau_{xy}|_{\theta=-\pi} = -\mu\sigma_y|_{\theta=\pm\pi}, \tag{14}$$

where μ_f equals the positive or negative value of the coefficient of kinetic friction, which is assumed to be constant, depending on the relative sliding direction of the two crack faces. Specifically, $\mu_f > 0$ when the sliding displacement δ is positive and $\mu_f < 0$ when $\delta < 0$. The length of the process (cohesive) zone l_p is

either prescribed (i.e., an initial cohesive zone exists before the loading is applied, and does not propagate under the present loading) or is determined by the condition $\hat{w} = 1$ in the normal cohesion-separation relation (10) or (11) at the instant of growth of the pre-existing traction-free crack.

The separated crack tip asymptotic solutions can be obtained after satisfying the boundary conditions (14). The complete asymptotic solutions are composed of two parts, corresponding to integer ($\lambda_n = n + 1$) and noninteger ($\lambda_n = (2n + 3)/2$) eigenvalues, respectively. These solutions are given in [Karihaloo and Xiao 2008].

The displacements corresponding to the first integer eigenvalue are

$$
\begin{aligned}
2\mu u &= r\{[a_{10}^I(\kappa - 1)\cos\theta - 2\mu_f\sin\theta] - a_{20}^I(\kappa + 1)\sin\theta - 2b_{10}^I(\cos\theta + \mu_f\sin\theta)\}, \\
2\mu v &= r\{[a_{10}^I(\kappa - 1)\sin\theta + 2\mu_f\cos\theta] - a_{20}^I(\kappa + 1)\cos\theta + 2b_{10}^I(\sin\theta - \mu_f\cos\theta)\}.
\end{aligned}
\tag{15}
$$

The displacements corresponding to the second integer eigenvalue are

$$
\begin{aligned}
2\mu u &= r^2\{a_{11}^I(\kappa\cos 2\theta - 2 - 3\mu_f\sin 2\theta) + a_{21}^I[-(\kappa+1)\sin 2\theta + 2] - 3b_{11}^I(\cos 2\theta + \mu_f\sin 2\theta)\}, \\
2\mu v &= r^2\{a_{11}^I(\kappa\sin 2\theta - 2 - 3\mu_f\cos 2\theta) + a_{21}^I[(\kappa + 1)\cos 2\theta + 2] + 3b_{11}^I(\sin 2\theta - \mu_f\cos 2\theta)\}.
\end{aligned}
\tag{16}
$$

The displacements corresponding to the first noninteger (fractional) eigenvalue are

$$
\begin{aligned}
2\mu u &= r^{3/2}\{a_{10}^f[(\kappa + \tfrac{1}{2})\cos\tfrac{3}{2}\theta - \tfrac{3}{2}\cos\tfrac{1}{2}\theta] + a_{20}^f[-(\kappa + \tfrac{5}{2})\sin\tfrac{3}{2}\theta - \tfrac{3}{2}\sin\tfrac{1}{2}\theta]\}, \\
2\mu v &= r^{3/2}\{a_{10}^f[(\kappa - \tfrac{1}{2})\sin\tfrac{3}{2}\theta - \tfrac{3}{2}\sin\tfrac{1}{2}\theta] + a_{20}^f[(\kappa - \tfrac{5}{2})\cos\tfrac{3}{2}\theta + \tfrac{3}{2}\cos\tfrac{1}{2}\theta]\}.
\end{aligned}
\tag{17}
$$

The displacements corresponding to the second noninteger eigenvalue are

$$
\begin{aligned}
2\mu u &= r^{5/2}\{a_{11}^f[(\kappa + \tfrac{3}{2})\cos\tfrac{5}{2}\theta - \tfrac{5}{2}\cos\tfrac{1}{2}\theta] + a_{21}^f[-(\kappa + \tfrac{7}{2})\sin\tfrac{5}{2}\theta + \tfrac{5}{2}\sin\tfrac{1}{2}\theta]\}, \\
2\mu v &= r^{5/2}\{a_{11}^f[(\kappa - \tfrac{3}{2})\sin\tfrac{5}{2}\theta + \tfrac{5}{2}\sin\tfrac{1}{2}\theta] + a_{21}^f[(\kappa - \tfrac{7}{2})\cos\tfrac{5}{2}\theta + \tfrac{5}{2}\cos\tfrac{1}{2}\theta]\}.
\end{aligned}
\tag{18}
$$

Here superscript I distinguishes coefficients associated with integer eigenvalues, whereas superscript f distinguishes coefficients associated with noninteger (fractional) eigenvalues.

For integer eigenvalues the opening displacement (COD) behind the cohesive zone tip and the sliding displacement of the cohesive crack faces vanish

$$
w = v|_{\theta=\pi} - v|_{\theta=-\pi} = 0, \quad \delta = u|_{\theta=\pi} - u|_{\theta=-\pi} = 0,
\tag{19}
$$

while σ_y and τ_{xy} are nonzero along the cohesive crack faces with

$$
\hat{\sigma}_y = \frac{\sigma_y}{f_t}\bigg|_{\substack{\theta=\pm\pi \\ 0\le r\le l_p}} = \sum_{n=0} c_n r^n = 1 + \sum_{n=1} c_n r^n,
\tag{20}
$$

where

$$
c_n = \frac{(n+1)(n+2)(a_{1n}^I + b_{1n}^I)\cos n\pi}{f_t}
\tag{21}
$$

with

$$
c_0 = \frac{2(a_{10}^I + b_{10}^I)}{f_t},
\tag{22}
$$

since $\sigma_y|_{\theta=\pm\pi} = f_t$ at $r = 0$.

For noninteger eigenvalues,

$$\sigma_y|_{\theta=\pm\pi} = 0, \tag{23}$$

$$\hat{w} = \frac{w}{w_c}\bigg|_{0 \le r \le l_p} = \sum_{n=0} \bar{d}_n r^{(2n+3)/2} = \bar{d}_0 r^{3/2} + \sum_{n=1} \bar{d}_n r^{(2n+3)/2}, \tag{24}$$

$$\delta|_{0 \le r \le l_p} = -\frac{\kappa+1}{\mu} \sum_{n=0} a_{2n}^f r^{(2n+3)/2} \sin \frac{2n+3}{2}\pi, \tag{25}$$

$$\bar{d}_n = \frac{\kappa+1}{\mu w_c} a_{1n}^f \sin \frac{2n+3}{2}\pi. \tag{26}$$

Set $d_0 = \bar{d}_0$ and $d_n = \bar{d}_n/d_0$ for $n \ge 1$. Then

$$d_0 = -\frac{\kappa+1}{\mu w_c} a_{10}^f, \quad d_n = -\frac{a_{1n}^f}{a_{10}^f} \sin \frac{2n+3}{2}\pi \tag{27}$$

By enforcing the cohesion-separation law (10) or (11), nonlinear relationships between d_n and c_n are obtained. The general relationships can be found in [12]. In [13], these general relationships have been explicitly derived for the simpler cohesion-separations laws (7) and (8).

4. Implementation in XFEM

To model the cohesive cracks in XFEM [Moës et al. 1999; Strouboulis et al. 2001; Babuška et al. 2003; Karihaloo and Xiao 2003], a standard local FE displacement approximation around the crack is enriched with discontinuous Heaviside functions along the crack faces behind the crack tip including the open traction-free part, and the crack tip asymptotic displacement fields at nodes surrounding the cohesive crack tip using the PU. The approximation of displacements for an element can be expressed in the form

$$\begin{Bmatrix} u^h(x) \\ v^h(x) \end{Bmatrix} = \sum_{i \in I} \phi_i(x) \begin{Bmatrix} u_{0i} \\ v_{0i} \end{Bmatrix} + \sum_{j \in J \cap I} \phi_j(x) H(x) \begin{Bmatrix} b_{1j} \\ b_{2j} \end{Bmatrix} + \sum_{m \in M_k \cap I} \phi_m(x) \begin{Bmatrix} u_m \\ v_m \end{Bmatrix}^{(tip,\kappa)}, \tag{28}$$

where I is the set of all nodes in the element, (u_{0i}, v_{0i}) are the regular degrees of freedom at node i, ϕ_i is the FE shape function associated with node i, J is the subset of nodes whose support is intersected by the crack but do not cover any cohesive crack tips, the function $H(x)$ is the Heaviside function centred on the crack discontinuity, and (b_{1j}, b_{2j}) are the corresponding additional degrees of freedom. M_k is the subset of nodes that are enriched around the cohesive crack tip k with the asymptotic displacements $u^{(tip,\kappa)}$ and $v^{(tip,\kappa)}$.

The general asymptotic field for each integer eigenvalue has three independent terms, e.g., (15)–(16), and each noninteger (fractional) eigenvalue has two independent terms, e.g., (17)–(18). The fields for integer and noninteger eigenvalues need to be used together to produce the crack opening and non-vanishing tractions.

In order to improve the accuracy of stresses, we may use eight terms corresponding to one noninteger eigenvalue and two integer eigenvalues as crack tip enrichment functions in (28)

$$\left\{ \begin{matrix} u \\ v \end{matrix} \right\}^{(\text{tip})} = \mathbf{\Phi}(r, \theta) \boldsymbol{q}^{(\text{tip})} \tag{29}$$

where $\boldsymbol{q}^{(\text{tip})} = [a_{10}^I, a_{20}^I, b_{10}^I, a_{10}^f, a_{20}^f, a_{11}^I, a_{21}^I, b_{11}^I]^T$ are additional nodal degrees of freedom at the enriched nodes to be solved together with nodal displacements of conventional FEM.

In (29), $\mathbf{\Phi}$ is the matrix formed by (r, θ) terms dependent on the additional unknown coefficients in (15)–(18).

For a mode I frictionless cohesive crack, the leading asymptotic displacement terms which correspond to a first noninteger eigenvalue adopted in [Xiao and Karihaloo 2006a] that gives a normal displacement discontinuity over the cohesive-crack faces can be used as crack tip enrichment functions:

$$u = \frac{r^{3/2}}{2\mu} a_1 [(\kappa + \tfrac{1}{2}) \cos \tfrac{3}{2}\theta - \tfrac{3}{2} \cos \tfrac{1}{2}\theta], \tag{30}$$

$$v = \frac{r^{3/2}}{2\mu} a_1 [(\kappa - \tfrac{1}{2}) \sin \tfrac{3}{2}\theta - \tfrac{3}{2} \sin \tfrac{1}{2}\theta]. \tag{31}$$

There are two possible ways to implement the above expansions in the XFEM. The first is to consider directly the nonlinear relationships between d_n and c_n and to obtain the cohesive stresses from the expansions. This avoids iterations on the cohesion-separation law, but requires solution of a system of algebraic equations with nonlinear constraints. The second way is to treat c_n and d_n as independent variables, and satisfy the cohesion-separation law (10) or (11) iteratively. The cohesive stresses in this case are obtained from the cohesion-separation law but not the expansions. However, these expansions can be used to smooth the numerically computed results. This way is generally more convenient to implement [11-13], and will also be used in the two examples below.

5. Examples, results and discussion

In this section, we will analyze typical mode I cohesive cracking problems of plain and fibre-reinforced concrete using the above asymptotic fields and XFEM. We consider three-point bend beams without any initial crack (Figure 1, left) made of plain concrete with the constant softening law (7), and of CARDIFRC with the softening law (9). We actually assume a very small initial crack of length 0.1 mm at the bottom midpoint of the beam. A state of plane strain is considered for all specimens. The geometrical parameters are $l = 4b$, and $t = b$, where t is the specimen thickness in the out-of-plane direction. The x-direction of nodes with coordinates $(0, 0)$ and $(0, l)$ and y-direction of the node with coordinates $(b, l/2)$ are constrained; the central "point" load is distributed over two elements.

We consider first the three-point bend plain concrete beams with $b = 150$ mm. This problem was solved in [11] with XFEM using the enrichment functions (30) and (31) for the linear and bilinear cohesion separation relations (2) and (3). It was shown that very accurate results could be obtained in this way with a coarse mesh (Figure 1, right) that would have required a much finer mesh with conventional FEM.

We have re-analysed the three-point bend beam using the constant traction law (7) with

$$f_t = 3.14 \, \text{MPa}, \quad w_c = 0.03886 \, \text{mm}, \quad G_F = 122 \, \text{N/m} \tag{32}$$
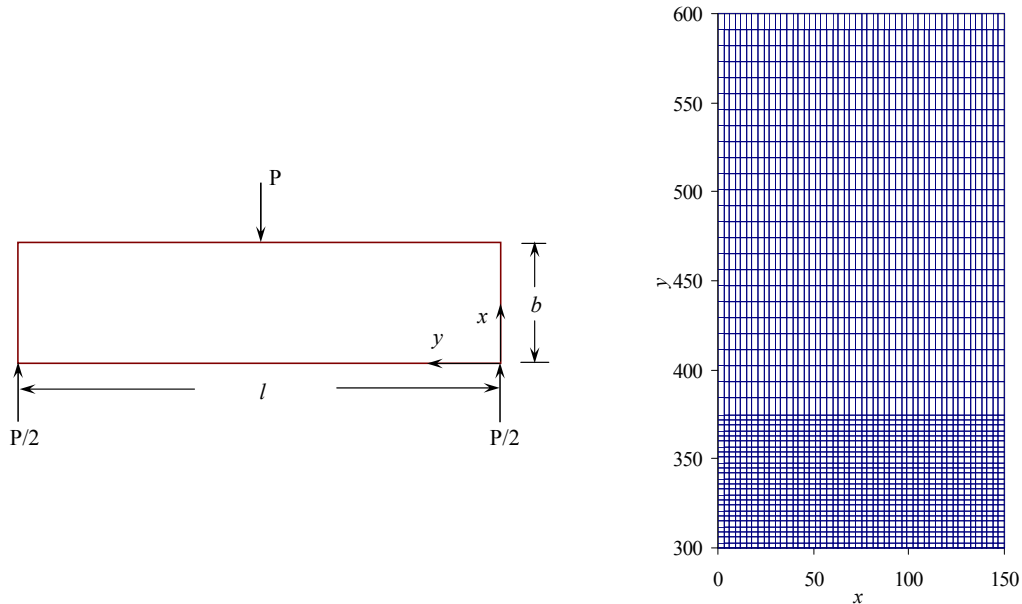
Figure 1. Left: three-point bend beam. Right: coarse mesh for one half of the specimen to the left of midspan (note rotation).

in order to study the influence of the shape of the cohesion-separation law on the global response. We have used the same coarse mesh as before (Figure 1, right). However to improve the accuracy further, we have used (15)–(18) as crack tip enrichment functions instead of the leading asymptotic displacement terms of a mode I frictionless cohesive crack (30) and (31) adopted in [Xiao and Karihaloo 2006a]. The nondimensional load-midspan deflection curve is shown in Figure 2 and compared with that of the bilinear law (*loc. cit.*). It is clear that although the tensile strength f_t and specific fracture energy G_F are the same, the initial branch of the cohesion-separation law produces very different responses: the constant traction law produces a higher peak load but a more brittle overall response with a significant snapback.
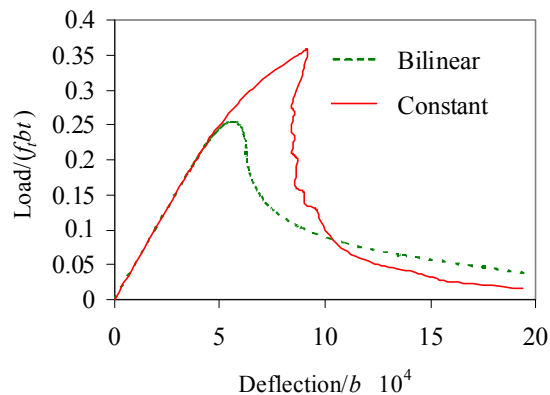
Figure 2. The nondimensional load-midspan deflection curves of the three-point bend beam with constant traction and bilinear softening laws.

Three-point bend fibre-reinforced concrete beam. We now consider a smaller three-point bend beam $b = 100$ mm made of CARDIFRC. We found it convenient to approximate the polynomial form of its cohesion separation relationship (9) by the bilinear relationship (3) with the material parameters $E = 48$ GPa, $f_t = 16$ MPa, $w_1 = 2.29$ mm, $f_1 = 3.9088$ MPa and the Poisson's ratio ν equals 0.2. The constraint and loading conditions are identical to the plain concrete beam. The discretised mesh is also identical to that shown in Figure 1, right, after scaling the coordinates by a factor of 2/3. The asymptotic displacements (30) and (31) are used as crack tip enrichment functions.

In the simulation, the first increment of the cohesive crack is 2.9 mm; thereafter the cohesive crack propagates by a segment of length 2 mm after each step. When the cohesive crack has developed over the entire depth of the specimen at a midspan deflection of about 3 mm, jump functions are used to enrich all nodes on the crack faces. The simulation is continued by increasing the deflection by 0.5 mm in each step. The load-midspan deflection curves are shown in Figure 3. The overall shape of the curve, peak load and the corresponding displacement from the XFEM simulation using the bilinear law (figure inset) agrees very well with the experiment (thin curve). The effect of large deformation has been ignored in the present simulation, which may be significant at the final stage. If this effect is taken into account, the agreement may be further improved.
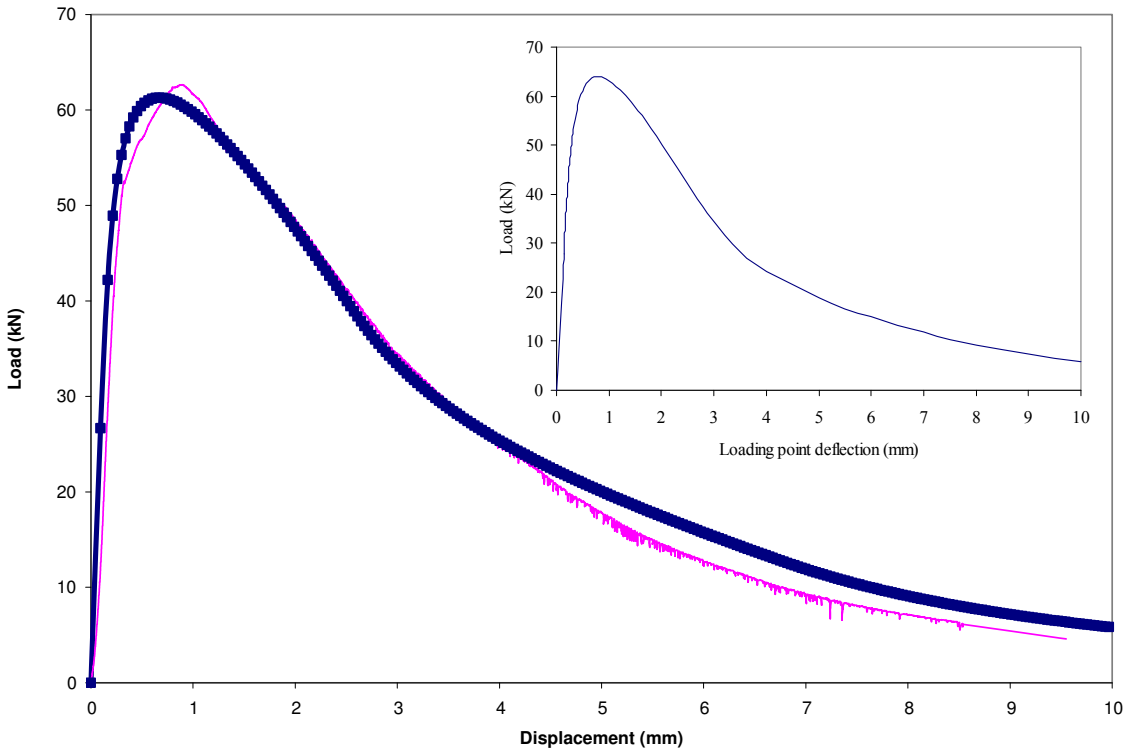

Figure 3. Large graph: load-deflection curves from experiments (thin curve) compared with an approximate method (thick curve/squares) in which the cohesive zone is regarded as a "hinge" in the central plane. Inset: load-deflection curve from the XFEM simulation. The agreement with the experimental curve is excellent both in terms of the shape and the maximum load carrying capacity.

References

[Babuška et al. 2003] I. Babuška, U. Banerjee, and J. E. Osborn, "Survey of meshless and generalized finite element methods: a unified approach", *Acta Numer.* **12** (2003), 1–125.

[Benson and Karihaloo 2005] S. D. P. Benson and B. L. Karihaloo, "CARDIFRC^{TM} - Development and mechanical properties, III: uniaxial tensile response and other mechanical properties", *Mag. Concrete Res.* **57**:8 (2005), 433–443.

[Borst et al. 2004] R. Borst, M. G. Gutierrez, G. N. Wells, J. J. C. Remmers, and H. Askes, "Cohesive-zone models, higher-order continuum theories and reliability methods for computational failure analysis", *Int. J. Numer. Methods Eng.* **60**:1 (2004), 289–315.

[Cornelissen et al. 1986] H. A. W. Cornelissen, D. A. Hordijk, and H. W. Reinhardt, "Experimental determination of crack softening characteristics of normalweight and lightweight concrete", *Heron* **31**:2 (1986), 45–56.

[Dugdale 1960] D. S. Dugdale, "Yielding of steel sheets containing slits", *J. Mech. Phys. Solids* **8**:2 (1960), 100–104.

[Hansbo and Hansbo 2004] A. Hansbo and P. Hansbo, "A finite element method for the simulation of strong and weak discontinuities in solid mechanics", *Comput. Methods Appl. Mech. Eng.* **193**:33-35 (2004), 3523–3540.

[Hillerborg et al. 1976] A. Hillerborg, M. Modeer, and P. E. Petersson, "Analysis of crack formation and crack growth in concrete by means of fracture mechanics and finite elements", *Cement Concrete Res.* **6**:6 (1976), 773–781.

[Karihaloo 1995] B. L. Karihaloo, *Fracture mechanics and structural concrete*, Longman, Harlow, Essex, 1995.

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

[Karihaloo and Xiao 2008] B. L. Karihaloo and Q. Z. Xiao, "Asymptotic fields at the tip of a cohesive crack", *Int. J. Fract.* **150** (2008), 55–74.

[Karihaloo and Xiao 2010] B. L. Karihaloo and Q. Z. Xiao, "Asymptotic fields ahead of mixed mode frictional cohesive cracks", *ZAMM Z. Angew. Math. Mech.* **90**:9 (2010), 710–720.

[Liu et al. 2004] X. Y. Liu, Q. Z. Xiao, and B. L. Karihaloo, "XFEM for direct evaluation of mixed mode SIFs in homogeneous and bimaterials", *Int. J. Numer. Methods Eng.* **59**:8 (2004), 1103–1118.

[Moës and Belytschko 2002] N. Moës and T. Belytschko, "Extended finite element method for cohesive crack growth", *Eng. Fract. Mech.* **69**:7 (2002), 813–833.

[Moës et al. 1999] N. Moës, J. Dolbow, and T. Belytschko, "A finite element method for crack growth without remeshing", *Int. J. Numer. Methods Eng.* **46**:1 (1999), 131–150.

[Strouboulis et al. 2001] T. Strouboulis, K. Copps, and I. Babuška, "The generalized finite element method", *Comput. Methods Appl. Mech. Eng.* **190**:32-33 (2001), 4081–4193.

[Wells and Sluys 2001] G. N. Wells and L. J. Sluys, "A new method for modelling cohesive cracks using finite elements", *Int. J. Numer. Methods Eng.* **50**:12 (2001), 2667–2682.

[Xiao and Karihaloo 2006a] Q. Z. Xiao and B. Karihaloo, "Asymptotic fields at frictionless and frictional cohesive crack tips in quasibrittle materials", *J. Mech. Mater. Struct.* **1**:5 (2006), 881–910.

[Xiao and Karihaloo 2006b] Q. Z. Xiao and B. L. Karihaloo, "Improving the accuracy of XFEM crack tip fields using higher order quadrature and statically admissible stress recovery", *Int. J. Numer. Methods Eng.* **66**:9 (2006), 1378–1410.

[Xiao et al. 2007] Q. Z. Xiao, B. L. Karihaloo, and X. Y. Liu, "Incremental-secant modulus iteration scheme and stress recovery for simulating cracking process in quasi-brittle materials using XFEM", *Int. J. Numer. Methods Eng.* **69**:12 (2007), 2606–2635.

[Zi and Belytschko 2003] G. Zi and T. Belytschko, "New crack-tip elements for XFEM and applications to cohesive cracks", *Int. J. Numer. Methods Eng.* **57**:15 (2003), 2221–2240.

BHUSHAN LAL KARIHALOO: karihaloob@cardiff.ac.uk
School of Engineering, Cardiff University, Cardiff, CF24 3AA, United Kingdom

QI-ZHI XIAO: qizhi.xiao@lusas.com
LUSAS FEA Ltd, Forge House, 66 High Street, Kingston-upon-Thames, KT1 1HN, United Kingdom

msp

MECHANICS OF MATERIALS AND STRUCTURES:
A SIMULATION-DRIVEN DESIGN APPROACH

LENNART KARLSSON, ANDREAS PAHKAMAA, MAGNUS KARLBERG,
MAGNUS LÖFSTRAND, JOHN GOLDAK AND JONAS PAVASSON

Engineering product development has developed considerably over the past decade. In order for industry to keep up with continuously changing requirements, it is necessary to develop new and innovative simulation methods. However, few tools and methods for simulation-driven design have been applied in industrial settings and proven to actually drive the development and selection of the ideal solution. Such tools, based on fundamental equations, are the focus of this paper.

In this paper the work is based on two cases of mechanics of materials and structures: welding and rotor dynamical simulations. These two examples of simulation-driven design indicate that a larger design space can be explored and that more possible solutions can be evaluated. Therefore, the approach improves the probability of innovations and finding optimal solutions.

A calibrated block dumping approach can be used to increase the efficiency of welding simulations when many simulations are required.

1. Introduction

Engineering product development has evolved considerably over the past decade. The vision of the sustainable society and a global context for many industries is now forcing manufacturing companies to develop innovative ways of developing their products. In addition, the product life-cycle will likely be shorter in the future, partly because customers will demand shorter delivery times. Suppliers will have to be more efficient than today and be able to assemble products from a large number of parts from subsuppliers. Increasingly, the value of the product originates outside the physical artifact (the hardware), in the form of services and add-ons. Instead of hardware ownership, other needs such as availability, productivity and risk minimization increasingly need to be catered for. In such a challenging environment and in order for industry to keep up with demand, it is necessary to develop new and innovative simulation methods, knowledge management methods and tools and methods for distributed work, to name a few.

In this paper, the focus is mainly on simulation-driven design applied on welding simulations (mechanics of materials) and rotor dynamical design (mechanics of structures). Below, the basic equations for modeling and simulation of welding and the basic equations for modeling and simulation of nonlinear dynamics are solved. For these two important problems, the way in which this fundamental modeling can effectively be used in simulation-driven product development is demonstrated.
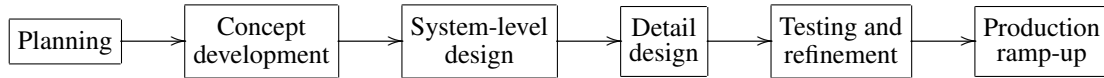
277

Figure 1. An example of a product development process.

Most industrial companies of any size follow some development process when developing their products. Figure 1 shows an example of such a development process from planning to production ramp-up [Ulrich and Eppinger 1995].

Product development literature provides a broad view of how to understand customer needs, develop and sell products, and includes discussions concerning best practices [Ulrich and Eppinger 1995; Wheelwright and Clark 1992; Cross 2000]. For example, Smith and Reinertsen [1998] offer a general view and aim to describe methods for generating a product to meet customer needs. Recently, product development processes have been extended to cover the whole process from needs to recycling.

Within the hardware product development domain, numerous tools have been developed to support the creation of excellent goods; e.g., computer aided engineering for geometric representation [LaCourse 1995] and the finite element method for stress calculation, exemplified in this paper. Typically, this work has been about making knowledge explicit and expressible and support tools have over time been developed to aid the creation of the hardware.

Simulations are ideally supposed to be used in the earliest stages of the development process. Often, however, there are few analyses involved until the detailed design stage is reached. At this stage simulations are often used in a verifying sense, meaning that the designer submits a model to the analysts who perform the analysis and then deliver the results to the designer. The results are used by the designer to verify whether or not the proposed design will meet the criteria. If not, the designer needs to upgrade the initial design, and then propose a new design for the analysts who perform a new simulation and so on, see Figure 2.

This sequential simulation usage workflow in combination with little usage of simulations in the earliest stages of the development process gives an ineffective product development process. To overcome this problem strategies for simulation-driven development have been proposed by some researchers [Courter 2009; Sellgren 1995; Wall 2007]. Naturally, such work has been ongoing for a relatively long time, as exemplified by [Hansen 1974] and [Gero 1981]. The development processes discussed in literature are often analyzed from an abstract point of view, while few reports exist on how simulation tools shall be designed and used to enable such methodologies. That is, few tools and methods have been applied in industrial settings and proven to actually drive the development and selection of the ideal solution.

Given these challenges, this paper is focused on briefly introducing the area of simulation driven design (SDD) [Courter 2009], and based on that approach describing a block dumping technique applied in welding simulations. Another application of the SDD approach is presented in the application area of rotor dynamical design.
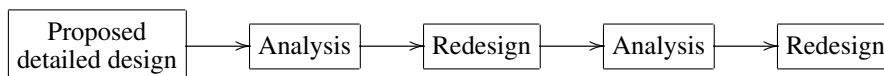

Figure 2. Sequential verifying simulation usage workflow.

Thus, the objective of this paper is to show how tools for modeling and simulation of mechanics of materials and structures shall be designed and used to enable simulation-driven product development. The objective is exemplified through two cases of mechanics of materials and structures: welding and rotor dynamical simulations.

2. Simulation-driven design methodology

In this paper a simulation-driven methodology is used [Bylund 2004; Goldak et al. 2007]. Bylund states that one way of reaching simulation-driven design is by providing simulation tools that can be used by the designers themselves. Designers thereby avoid having to send models and results back and forth, often among several different people. For designers not used to analysis work to be able to perform the necessary analyses, tools that are *intuitive and include expert knowledge* through equations of mechanics of materials and structures are required. The results from the analyses must further be *derived quickly* and must be *easy to interpret.* Hence, the demands on the simulation tools affect the preprocessing, the solver and the postprocessing.

3. Simulation-driven design of welded structures

Welding is one of the most commonly used methods of joining metal pieces. In fusion welding, the metal pieces are heated until they melt together, resulting in strong coupling between thermal, mechanical and metallurgical (microstructural) properties. Due to the complexity of the welding applications, the governing equations as well as the thermal, mechanical and metallurgical couplings, computational support is necessary for prediction of distortions and residual stresses. The development of the field of computational welding mechanics has been described in, for example, [Karlsson 1986; Goldak and Akhlaghi 2005; Lindgren 2001a; 2001b; 2001c; 2007]. Much research has focused on predictions of residual stresses and distortions [Chen and Sheng 1992; Lee et al. 2008; Ueda et al. 1988; Ueda and Yuan 1993].

Several reports exist regarding validation of predictions of residual stresses and deformation [Barroso et al. 2010; Karlsson et al. 1989; Mochizuki et al. 2000; Ueda et al. 1986]. The European Network on Neutron Techniques Standardization for Structural Integrity (NeT) has a benchmark problem for a single weld bead-on-plate specimen [Truman and Smith 2009]. By using finite element simulations, NeT members have predicted and measured residual stresses and thermal fields of this benchmark problem by different methods, the results of which have been compiled [Smith and Smith 2009a; 2009b]. In those reports, different sources of errors are discussed, and it was also found that "there is much room for improvement" regarding prediction accuracy.

In product development, residual stresses and distortions often need predicting to verify alignment tolerances, strength demands, fatigue requirements, etc. It is important, for example, to keep track of the residual stress- and deformation history when simulating a sequence of manufacturing processes [Deng et al. 2009; Åström 2004]. Different approaches have been developed for how to use welding simulations to predict suitable sequences of weld paths. Troive et al. [1998] compared different predefined paths, while Voutchkov et al. [2005] used surrogate models to solve a combinatorial weld path planning problem. Efforts have further been made to show effects of deposition sequences regarding, for instance, residual stresses and distortions [Ogawa et al. 2009]. Although a lot of research has been conducted on welding

simulations, there is still a need for development of more efficient welding simulations and methodologies for using such simulations to support design processes.

In this article, a strategy for improved efficiency of welding simulations is used as proposed in [Pahka-maa et al. 2010], aiming at developing a simulation-driven design methodology for welding simulations. In this work, the modeling and simulation software VrWeld from Goldak Technologies Inc. [Goldak Technologies 2010] has been used to demonstrate the proposed strategy.

3.1. *Welding simulation theory.* The field of welding simulations described in this paper (computational welding mechanics) is partially built upon earlier work within the fields of thermal, mechanical and metallurgical (microstructural) properties of materials. The principles and applications of welding simulations have been described in [Goldak and Akhlaghi 2005; Lindgren 2001b]. Welding simulation is a good example of how mechanics of materials and structures can be put to practical use with the support of computers. The essential features of computational weld mechanics (CWM) are [Goldak 2009]:

- It requires solving the nonlinear, coupled 3D transient partial differential equations (PDEs) for heat flow (conservation of energy), microstructure evolution and stress-strain evolution (conservation of momentum).

- The material properties are temperature, history dependent and involve phase changes.

- The welding process usually adds material (filler metal), making the geometric domain a time-dependent free surface problem.

- The boundary conditions applied by fixtures, clamps and tack welds are complex and transient.

- The geometry of welded structures is often complex with many parts.

- Modeling the heat source of the arc is itself complex.

3.1.1. *Welding simulation software (VrWeld).* VrWeld, part of the VrSuite software package, is a finite element program used to simulate welding processes. VrSuite [Goldak Technologies 2010] uses operator splitting to solve the system of equations needed to model manufacturing processes such as welding and heat treating and the in-service behavior of assemblies of such manufactured parts. Each equation in the system is solved with the algorithm, domain, parameters, initial conditions, boundary conditions, length scale and time scale that best approximate the physics of the phenomena that the equation models. Maps are created to map data from this equation to each equation that it is coupled to. Each equation is solved iteratively using solvers such as frontal, multifrontal, ICCG, GMRES, MG and AMG [Saad 1996]. Limits and bifurcation points are computed using the Wriggers–Simo algorithm [Wriggers and Simo 1990]. A version of Crisfield's arc-length method [Ramm 1981] is used to follow the path of the solution in nonlinear analysis. The meshing is largely automated. Domain decomposition and adaptive meshing play an important role. Solvers run in-core using processors with 8 GB of RAM. The CPU-intensive code is written in C++. The high-level code is written in a scripting language such as Tcl/Tk or HTML.

3.1.2. *Block dumping (fast simulations in VrWeld).* Traditional transient welding simulations often use a moving heat source such as the Gaussian double ellipsoid [Goldak and Akhlaghi 2005]; see Figure 3. This approach has proved to accurately predict the temperature history, residual stresses and welding distortions [Goldak and Akhlaghi 2005; Lindgren 2001a; 2001b; 2001c; 2007]. It is suggested that the
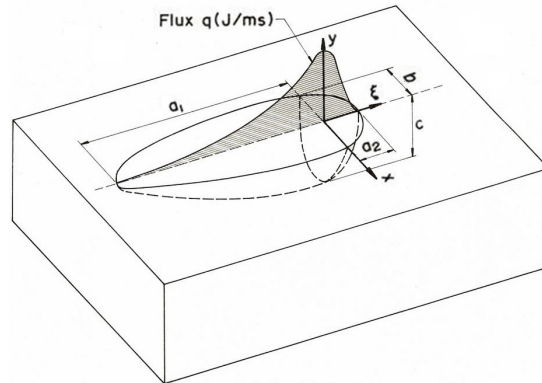
Figure 3. Double ellipsoid heat source with Gaussian heat distribution [Goldak and Akhlaghi 2005].

heat source may move no more than half of the weld pool length to function properly in three dimensional welding simulations [Goldak et al. 1986]. This suggestion causes a transient welding simulation that uses a moving heat source to be quite time consuming, even with present-day computational power. The problem gets even more difficult when a large design space is to be explored, meaning that many simulations are required.

A faster way to simulate the thermal history of a weld is the block dumping method [Goldak et al. 1986], also referred to as the prolonged Gaussian heat source [Cai and Zhao 2003]. The block dumping approach heats the whole weld, or large pieces of it, in a single time step. This approach can reduce the number of time steps in the welding simulation significantly and thereby reduce the total calculation time without any appreciable decrease in accuracy [Pahkamaa et al. 2010]. With the block dumping technique, the same parameters as the Gaussian Ellipsoid heat source are used to calculate the amount of heat distributed to each element, i.e., voltage, current, efficiency, welding speed and parameters a_1, a_2, b and c in Figure 3. The total amount of heat distributed to the model is identical to the heat input of a traditional transient simulation. The difference is the time in which the heat is diffused. In each block dump, the amount of heat added is exactly the same as the heat that would be added in a transient analysis for this length of weld, and it is added in exactly the same way with the same transient time steps. However, no heat diffusion is done until all heat has been added for this length of weld. Then one heating time step is done with these applied nodal thermal loads to heat this length of weld. Following this heating time step, the heat equation is solved for the structure as it cools with exponentially increasing length of time steps. Usually, 4 to 6 cooling time steps are applied. The designer can choose to allow each block dump to cool to ambient temperature or use a shorter cooling time in which the structure does not fully cool, the latter approach is used in this paper. In the heating and cooling steps, the thermal-elastoplastic stress analysis problem is solved. This process is applied for each block dump in the sequence of block dumps specified by the designer.

3.2. *Simulation strategy.* The welding simulation strategy proposed in [Pahkamaa et al. 2010] (Figure 4), which will be verified in this study, reduces the calculation time for welding simulations by replacing the traditional moving heat source with a calibrated block dumping heat source. The block dumping heat
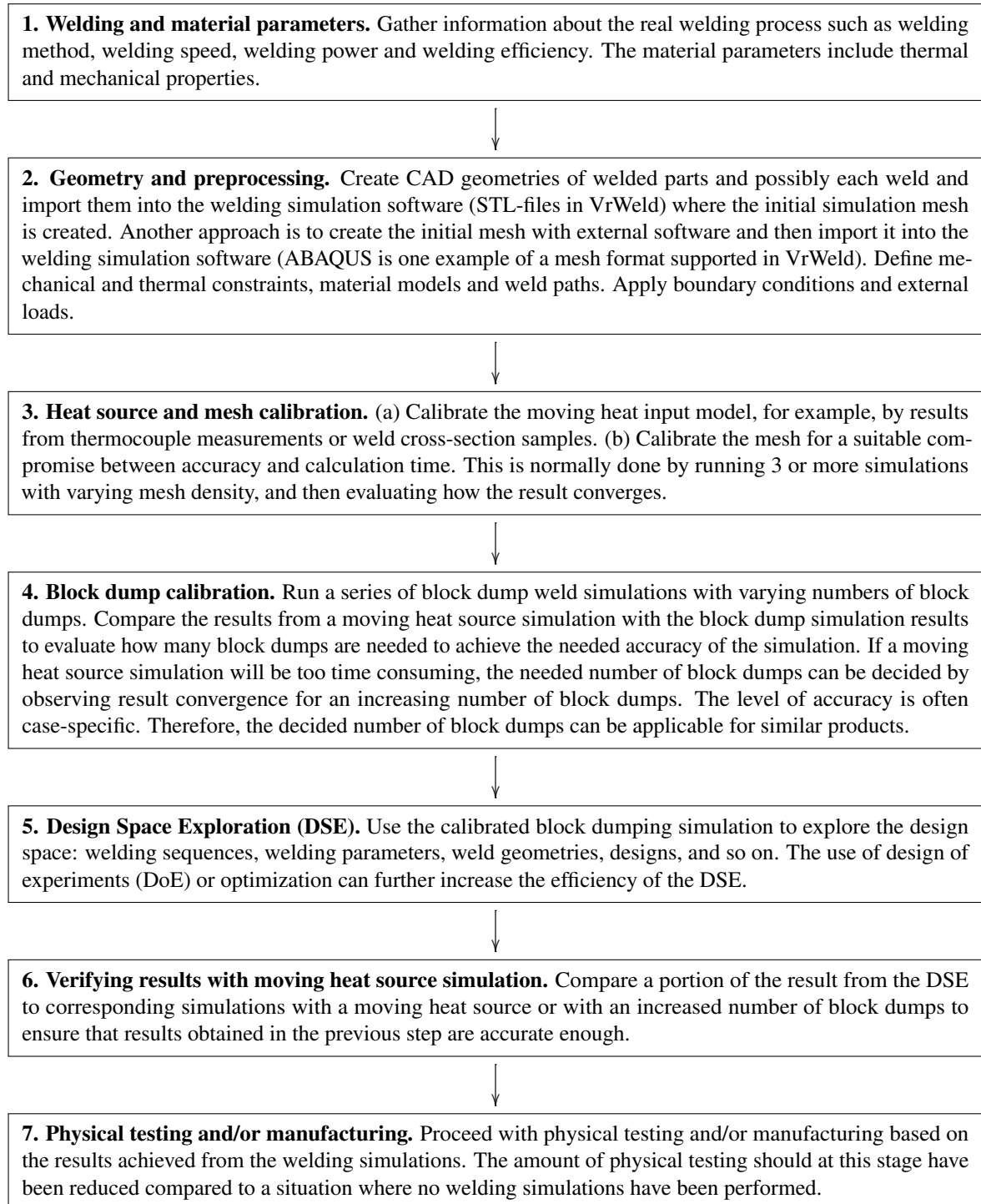
1. Welding and material parameters. Gather information about the real welding process such as welding method, welding speed, welding power and welding efficiency. The material parameters include thermal and mechanical properties.

2. Geometry and preprocessing. Create CAD geometries of welded parts and possibly each weld and import them into the welding simulation software (STL-files in VrWeld) where the initial simulation mesh is created. Another approach is to create the initial mesh with external software and then import it into the welding simulation software (ABAQUS is one example of a mesh format supported in VrWeld). Define mechanical and thermal constraints, material models and weld paths. Apply boundary conditions and external loads.

3. Heat source and mesh calibration. (a) Calibrate the moving heat input model, for example, by results from thermocouple measurements or weld cross-section samples. (b) Calibrate the mesh for a suitable compromise between accuracy and calculation time. This is normally done by running 3 or more simulations with varying mesh density, and then evaluating how the result converges.

4. Block dump calibration. Run a series of block dump weld simulations with varying numbers of block dumps. Compare the results from a moving heat source simulation with the block dump simulation results to evaluate how many block dumps are needed to achieve the needed accuracy of the simulation. If a moving heat source simulation will be too time consuming, the needed number of block dumps can be decided by observing result convergence for an increasing number of block dumps. The level of accuracy is often case-specific. Therefore, the decided number of block dumps can be applicable for similar products.

5. Design Space Exploration (DSE). Use the calibrated block dumping simulation to explore the design space: welding sequences, welding parameters, weld geometries, designs, and so on. The use of design of experiments (DoE) or optimization can further increase the efficiency of the DSE.

6. Verifying results with moving heat source simulation. Compare a portion of the result from the DSE to corresponding simulations with a moving heat source or with an increased number of block dumps to ensure that results obtained in the previous step are accurate enough.

7. Physical testing and/or manufacturing. Proceed with physical testing and/or manufacturing based on the results achieved from the welding simulations. The amount of physical testing should at this stage have been reduced compared to a situation where no welding simulations have been performed.

Figure 4. Suggested welding simulation strategy for simulation-driven design. From [Pahkamaa et al. 2010].

Figure 5. Rear frame and rear axle bridge.

source is calibrated to give a good compromise between simulation accuracy and calculation time, thus allowing a larger design space to be explored during a given period of time. A rear axle bridge from a Volvo Construction Equipment wheel loader is used as a demonstrator case to show that steps 1–6 work on an industrial application. Twenty different welding sequences are compared to find the welding sequence that gives the smallest welding distortion at key positions.

3.3. *Welding case study.* The object of this case study is a rear axle bridge from a Volvo Construction Equipment wheel loader. The axle bridge is positioned in the rear frame as in Figure 5. An axle working as a pivot for the rear axle assembly is mounted in the two holes in the axle bridge. The axle is supported by journal bearings, which are lubricated with oil from the rear axle differential. The concentricity between the two holes and the parallelism and perpendicular alignment between the plates are two important tolerance demands for the axle bridge. Due to these requirements, it is important that the welding process used to manufacture the axle bridge does not introduce excessive deformations. Therefore, the aim of this case study is to derive a suitable welding sequence that minimizes the welding distortion in the axle bridge. Identifying the proper welding sequences is done by use of the proposed simulation strategy; see Figure 4.

3.3.1. *Welding and material properties.* The three plates are joined by welds, marked a–d in Figure 6; the welds have a throat size of 6 mm. The axle bridge is manually tack-welded with 40 mm long tack welds at start, mid and end of the four welds. The tack welding is performed in a separate fixture; the tack welded axle bridge is then positioned in the welding fixture shown in Figure 6. An automated MAG welding process then applies the four welds, with the following parameters:

$$\text{Power}:\ 10880\,\text{W (34 V, 320 A)}$$
$$\text{Efficiency}:\ 85\%$$
$$\text{Welding speed}:\ 37\,\text{cm/min}$$

3.3.2. *Geometry and preprocessing.* A CAD model of the axle bridge assembly was provided by the manufacturer. This CAD model was imported to Siemens PLM NX6, where small features were removed to ease the meshing process. The idealized part shown in Figure 7 was exported to the welding
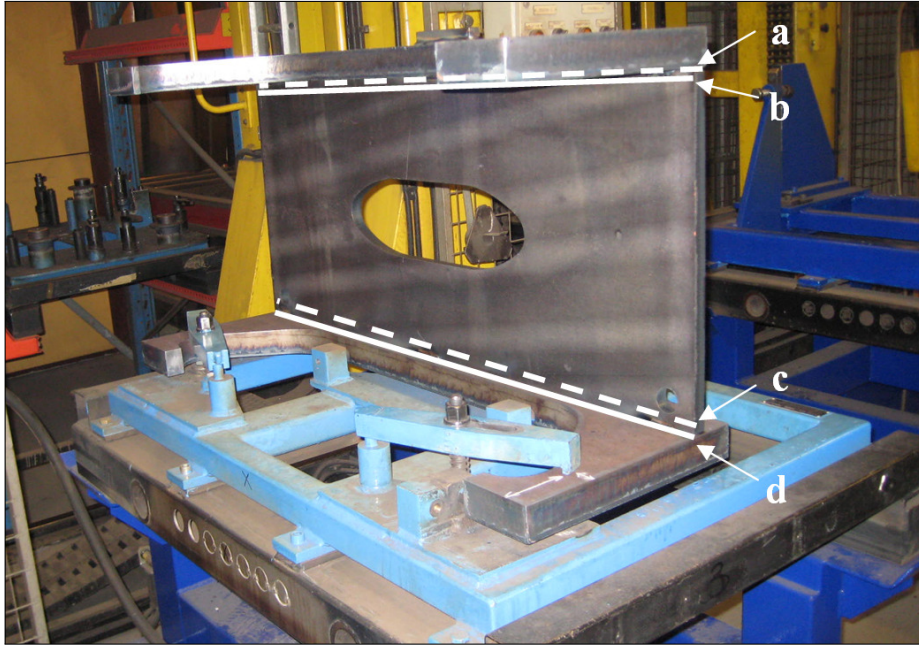
Figure 6. Real axle bridge placed in its welding fixture. The weld joints (a–d) are marked with white lines; welds a and c are placed underneath the center plate.

simulation software (VrWeld) in STL-format. The simulation model was then created in VrWeld by defining weld joints, assigning materials, creating an initial mesh and assigning initial conditions and boundary conditions, etc. The material model described in [Andersson 1978] was used for both plates and filler metal. The used fixture shown in Figure 7 prevents rigid body motions (locks six DOFs) without restraining the growth/shrinkage of the plates. This fixture modeling method was used, since the real fixture only clamps one of the plates, and should therefore not have a large impact on the welding distortions. The welding distortion in the rear axle bridge was measured as the global displacement of points P1 and P2. The global ambient temperature is set to 300° K.
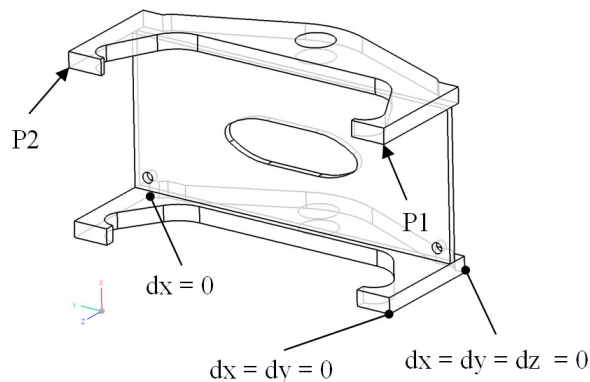
Figure 7. Idealized axle bridge, evaluation points and constraints.
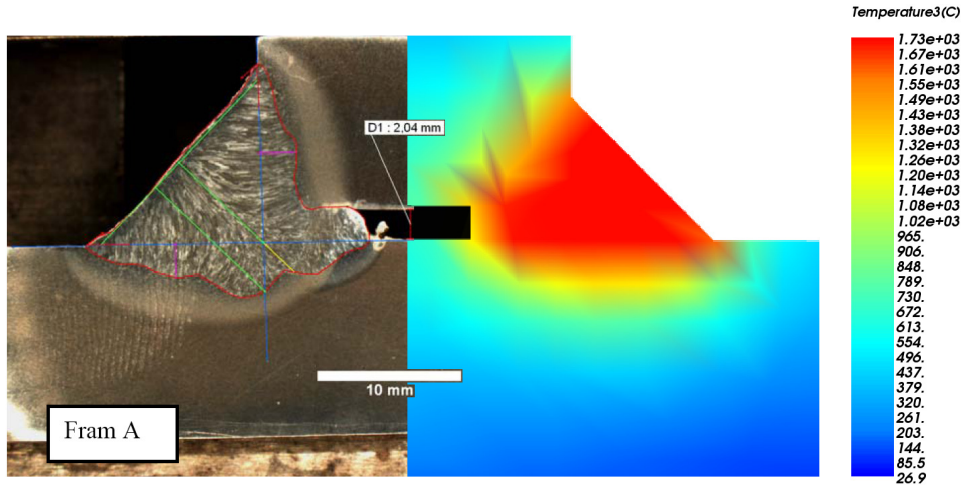
Figure 8. Comparison of weld cross section photograph and simulated heat input.

3.3.3. *Moving heat input calibration.* The moving heat input model was calibrated by use of a weld cross section photograph, see Figure 8. The trial and error method was used to find suitable heat input parameters that describe the shape of the Double Ellipsoid heat source. The calibrated parameters are as follows (see Figure 3):

$$a_1 = 10\,\text{mm} \qquad a_2 = 6\,\text{mm} \qquad b = 6\,\text{mm} \qquad c = 8\,\text{mm}$$

3.3.4. *Mesh and block dump calibration.* Three models with different mesh density were created, as follows:

Mesh #	Volume elements	Elements in weld joint (cross-section)
1	13237	5
2	27142	7–9
3	42665	12–15

For each of them, seven welding simulations were run, with an increasing number of block dumps (4, 8, 15, 20, 30, 40, 50). Ten cooling time steps were used in each block dump simulation. The "Fast Simulation" option in VrWeld uses the same number of block dumps for all welds. Therefore, a full-length weld will be welded with the same number of block dumps as a half-length weld. Thus, a welding sequence with full-length welds and, for example, 20 block dumps will be simulated in 90 time steps (4 welds × 20 block dumps + 10 cooling time steps), while a welding sequence with half-length welds will be simulated in 170 time steps (8 welds × 20 block dumps + 10 cooling time steps). The case used to calibrate the mesh and number of block dumps has four half-length welds and two full-length welds (see Appendix).

The absolute displacement of points P1 and P2 (see Figure 7) was used to evaluate the results. Figure 9 shows results for point P1, similar results were obtained for point P2. From Figure 9 it can be concluded that there is little difference in results between mesh 2 and 3, while the results from mesh 1 deviate significantly. The difference between mesh 2 and 3 is only 0.1 mm when 40 block dumps are used. Based on these results, mesh number 2 and 40 block dumps will be used for the DSE. The chosen simulation
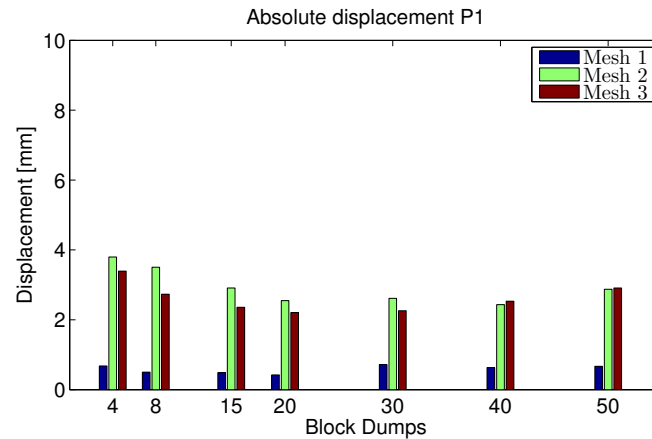
Figure 9. Mesh and block dump calibration. There is a difference of 0.1 mm between Mesh 2 and Mesh 3 when 40 block dumps are used.
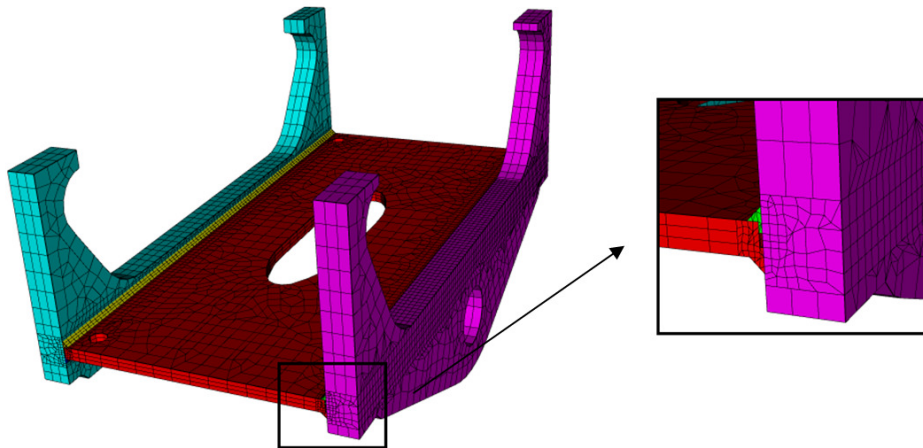
Figure 10. The chosen simulation mesh has 8 elements across the weld joint and 27142 elements in total.

mesh is shown in Figure 10. To verify predictions made by use of the chosen mesh and number of block dumps, the results should be compared with the results from simulations with a moving heat source. If the same best welding sequence is being predicted by the two models, the chosen mesh and number of block dumps are considered to be sufficient.

3.3.5. *Design space exploration.* A structure that has four welds, such as the axle bridge, can be welded in 384 ($2^4 \times 4!$) different ways. The large number of combinations is possible when allowing the welds to be laid in any order (4!) and in both directions (2^4). If only half welds are allowed, the number of possible welding combinations grows to about 10 million possible combinations ($2^8 \times 8!$). Design space exploration (DSE) refers to exploring the set of possible welding combinations and selecting a subset which meets the requirements. In this case, it is not practical to explore the entire design space. To show how welding sequences influence the displacements in P1 and P2, twenty different sequences are
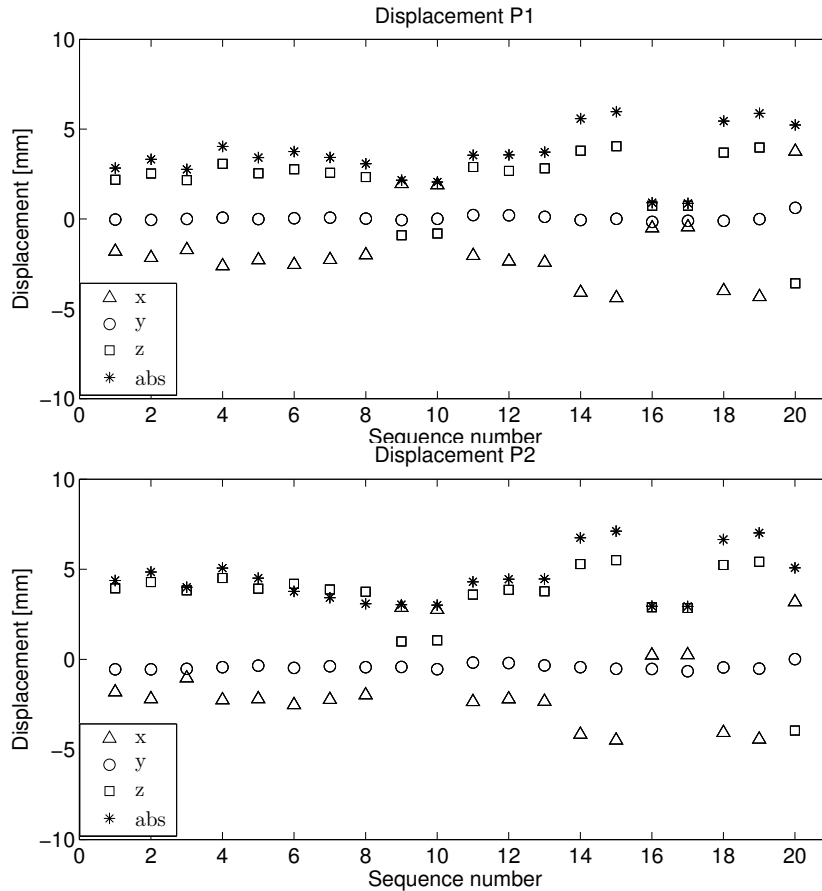
Figure 11. Comparison of displacement of points P1 (top) and P2 (bottom) for welding sequences 1–20. The displacement is plotted as *x*-, *y*- and *z*-components and absolute displacement.

simulated; see the Appendix. The results from the twenty welding sequences are shown in Figure 11. Here, the displacement in the *x*-, *y*- and *z*-directions as well as the absolute value is presented for points P1 and P2 for each welding sequence. These results show that welding sequences 9, 10, 16 and 17 give the smallest displacements for P1 and P2, while welding sequences 14, 15, 18 and 19 give the largest displacements in these points. To further illustrate the difference between the welding sequences, the sum of the absolute displacement of points P1 and P2 is shown in Figure 12. This figure shows that welding sequences 16 and 17 give the smallest total displacement (3.8 mm), while welding sequences 15 and 19 give the largest total displacement (13.0 mm). Figure 13 shows the displacement field amplified 20 times for welding sequences 17 and 15.

3.3.6. *Verification of simulation strategy.* To ensure that the chosen number of block dumps gives sufficient accuracy, it is suggested that a small number of moving heat source simulations are performed. In this case, to show that high accuracy has been achieved with the block dumping method, simulations with a moving heat source for all twenty welding sequences were conducted. It was found that the block
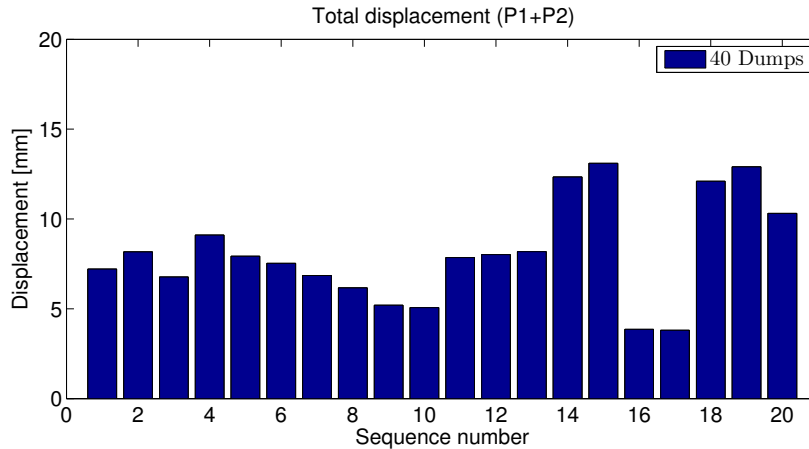
Figure 12. Total displacement calculated as the sum of the absolute displacement of points P1 and P2. Notice that welding sequences 16 and 17 give the smallest total displacement (4 mm), while welding sequences 15 and 19 give the largest total displacement (13 mm).
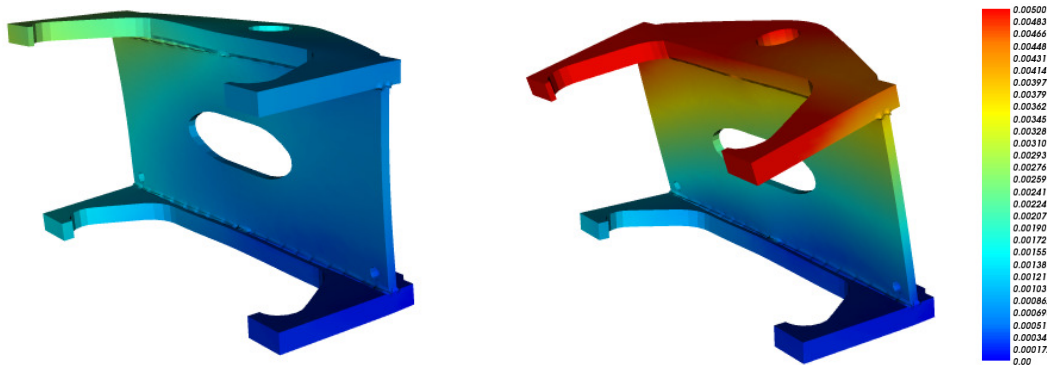
Figure 13. Comparison of best (#17) and worst (#15) welding sequences. Colors indicate displacements, given in mm on the scale on the right and magnified 20 times in the geometric representation.

dumped simulations deviate by 0.7 mm (mean value) compared to the moving heat source simulations. Both simulation approaches predict the same welding sequences to give the smallest and largest welding distortions in P1 and P2. Welding sequences 1–4 are welded with full-length welds only and simulated in 170 time steps. Welding sequence 20 is welded with two full-length welds and four half-length welds and simulated in 250 time steps. Welding sequences 5–19 are welded with eight half-length welds and simulated in 330 time steps. The moving heat source is simulated in 450 time steps. Hence, the number of time steps has then been reduced by 60% (welding sequences 1–4), 45% (welding sequence 20) and 25% (welding sequences 5–19).

3.4. *Welding case discussion.* This case study has shown that calibrated block dumping simulations can reduce the calculation time compared to moving heat source simulations. The results from the block
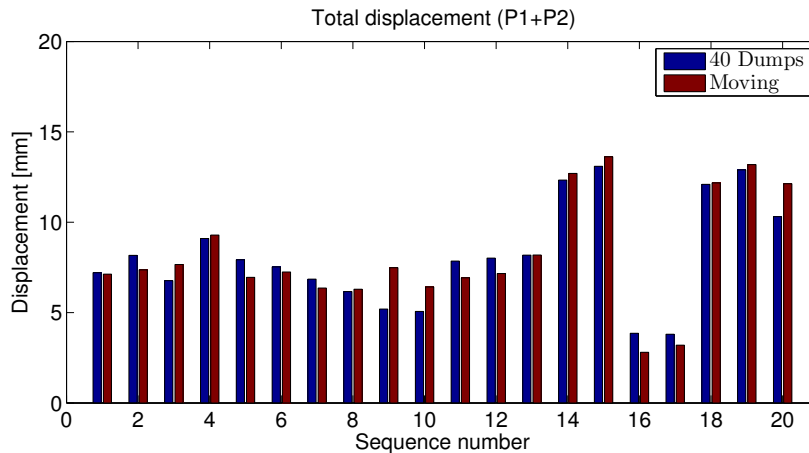
Figure 14. Comparison between block dumped and moving heat source simulations.

dumped simulations were able to point out the best (16 and 17) and worst (15 and 19) welding sequences (Figure 14), which was identical to the results obtained with the moving heat source simulations. The block dump calibration and mesh calibration for this case study was performed by running 21 simulations. Thus, using the calibrated block dump method to simulate as few as twenty welding sequences will not save much time. The strength of this method becomes obvious when a larger design space is explored, when hundreds or thousands of welding simulations are performed.

This work is done on the assumption that the 3D transient welding simulations approach used in this paper is able to predict residual stresses and welding distortions correctly. Future work should involve validation of the simulation results obtained in this paper.

This case study has shown that it should be possible to use a simulation-driven design approach in the design process regarding welded structures. Twenty different welding sequences were compared, showing that welding sequences 16 and 17 gave the smallest welding distortions in P1 and P2 and could therefore be considered as the best suited for this structure.

Future work should explore the possibilities of combining the calibrated block dumping simulations with design of experiments or optimization routines to further rationalize the design process of welded structures. A computer program made to generate suitable welding sequences to be explored would significantly decrease designer manual labor. The welding sequences listed in Table 2 where defined manually, which was a quite time consuming task.

It would also be valuable to perform a similar case study where the block dumps are defined by their length instead of the number of block dumps per weld.

4. Simulation-driven rotor dynamical design

Few CAE systems support rotor dynamical analysis, and if they do they have restricted functionality. Therefore, specialized software is commonly used when designing rotating machinery. These specialized rotor dynamical codes often include a lot of functionality, but for nonexperts, they can be difficult to use. In order to facilitate simulation-driven design, an in-house rigid body rotor dynamical demonstrator code

called RESORS has been developed at Luleå University of Technology, Division of Computer Aided Design. The aim is to enable nonexperts to conduct advanced nonlinear rotor dynamical analyses very early in the product development process. By this strategy, sending analysis requests to experts can be avoided, allowing possibilities for exploring a larger part of the design space.

4.1. *Rotor dynamical theory.* Rotor dynamical theory was basically developed during the last century, but started earlier, with Rankine [1869], who incorrectly predicted that the first critical speed could not be exceeded. In the late nineteenth century, de Laval practically proved (see [Childs 1993]) that rotating machinery can run supercritically, which was later verified by Jeffcott [1919]. Since it was introduced, the Jeffcott rotor model has been widely used for different scientific purposes; see for example [Childs 1982; Karlberg and Aidanpää 2003; Karlberg and Aidanpää 2004]. In linear rotor dynamics, eigenfrequencies depend on the spin speed due to the gyroscopic effect. In addition, the direction of the vibration is important, and forward whirl therefore has to be distinguished from backward whirl [Genta 1999]. Some rotor dynamical systems become strongly nonlinear due to for example clearances [Chu and Zhang 1997; Edwards et al. 1999; Ganesan 1996; Goldman and Muszynska 1995; Muszynska and Goldman 1995], processes [Karlberg and Aidanpää 2007], bearings [Harris 1991] etc. Hence, in rotor dynamical software, it is important to enable predictions of linear as well as nonlinear analysis of rotating machinery

4.2. *Simulation strategy.* To show how RESORS rationalizes a design process at early stages of the product development process, an SDD approach is applied on an industrial development case. The SDD approach can (in this example) be divided into three steps, starting with *gathering of information* — requirements, limitations, parameter ranges, etc. Even if nonlinear systems are to be designed, the designer should start with a *linearized simulation model*. This model is primarily used to decide the mesh density and to set suitable damping but also to get an understanding of critical speeds and mode shapes (which may be of interest even in nonlinear systems). To avoid vibration problems due to nonlinearities, at the final step, a *fully nonlinear simulation model* is studied. Since common postprocessing methods fail in nonlinear systems (Campbell diagrams, critical speeds, mode shapes, etc.), other methods are required.

4.3. *Rotor dynamical case study.* In every product development project there are limits to the design space. Before starting modeling and simulation work, information regarding design space boundaries needs to be gathered.

4.3.1. *Information gathering.* In the industrial case used for this paper an overhung rotor system will be designed. The shaft will be supported by radial bearings in order to enable axial displacement during operation. The machinery will be powered by an electrical motor running at 1500 rpm. Earlier concepts were subcritical, but *in the scenario presented here a supercritical machine is to be developed.*

In previous designs, bearings without clearance were used. Table 1, on the next page, lists the needed properties from the latest design, which will be used as a starting point.

The damping ratios are between 0.6% and 16% when considering all vibration modes. In this scenario it is assumed that all parameters shown in Table 1 are fixed except for the rotor mass, the pedestal stiffness, the clearance and the allowed rotor mass. The rotor mass is allowed to be varied between 1800 kg and 2200 kg, the pedestal stiffness between 10^7–10^9 N/m and the clearance between 0.01–0.1 mm.

Shaft :	Density	$7800\,\mathrm{kg/m^3}$	**Rotor :**	Shaft position	Right end
	Young's modulus	$200\,\mathrm{GPa}$		Mass	$2000\,\mathrm{kg}$
	Length	$1\,\mathrm{m}$		Polar mass I	$200\,\mathrm{kg\,m^2}$
	Radius	$0.1\,\mathrm{m}$		Transversal mass I	$100\,\mathrm{kg\,m^2}$
Pedestals :	Shaft position	$0\,/\,0.8\,\mathrm{m}$	**Load :**	Rotor unbalance	$10^{-4}\,\mathrm{m}$
	Type	Radial, no clearance		Gravity	$9.8\,\mathrm{m/s^2}$
	Isotropic stiffness	$10^8\,\mathrm{N/m}$ (both)			

Table 1. Properties of latest design (I = moment of inertia).

4.3.2. *Analysis by use of linearized simulation model.* To rapidly converge on suitable solutions a linearized simulation model, without clearance, is initially used. The objective of this analysis is to find suitable pedestal stiffness and rotor mass that give a supercritical system.

Preprocessing. Figure 16a (next page) shows the graphical user interface (GUI) for the preprocessing step in RESORS. Here, the designer is requested to enter data for the model (concept) to be analyzed. The GUI is designed so that the in-data fields are directly coupled to physical properties, and thus easy to understand. The designer enters data about the shaft, the rotor (disc), the pedestals (supporting structure), the load, the spin speed and damping. To simplify implementation of pedestal data, different types of supporting structures (stiffness matrices) are predefined for the designer to test (see Figure 16b). Another common issue to deal with is damping. In RESORS (and other commercial software) proportional damping is implemented and hence two parameters, α and β, can be chosen to get suitable damping. The choice of α and β usually requires some experience; therefore, a procedure has been implemented in RESORS showing the minimum and maximum damping ratio for each set of parameters. This gives direct feedback to the designer, who can then choose suitable ranges for the damping ratios (see Figure 16c). Entering all in-data to the preprocessor typically takes a few minutes.

Mesh convergence. In the next step of the design process it is suggested that the mesh density is checked by means of variation of eigenvalues. Figure 15 shows the preprocessed model (by use of in-data from Table 1) with three different mesh densities: two, seven and fifteen elements.

Eigenfrequencies are derived without numerical simulation in time domain. In RESORS, the derivation of Campbell diagrams is therefore implemented directly in the postprocessor and is usually obtained
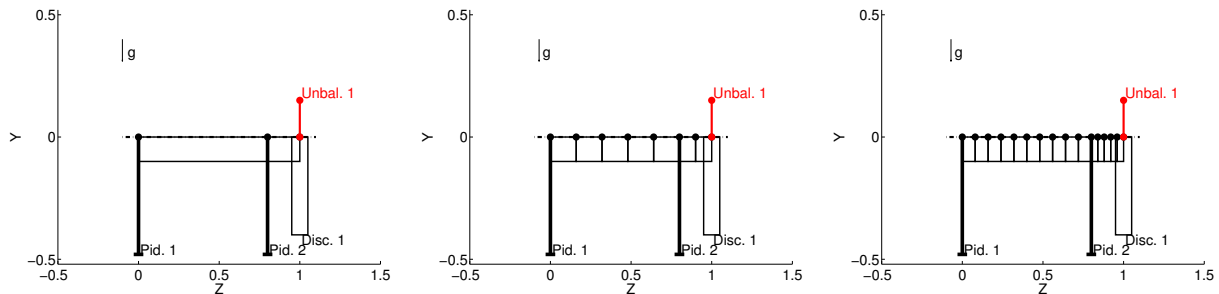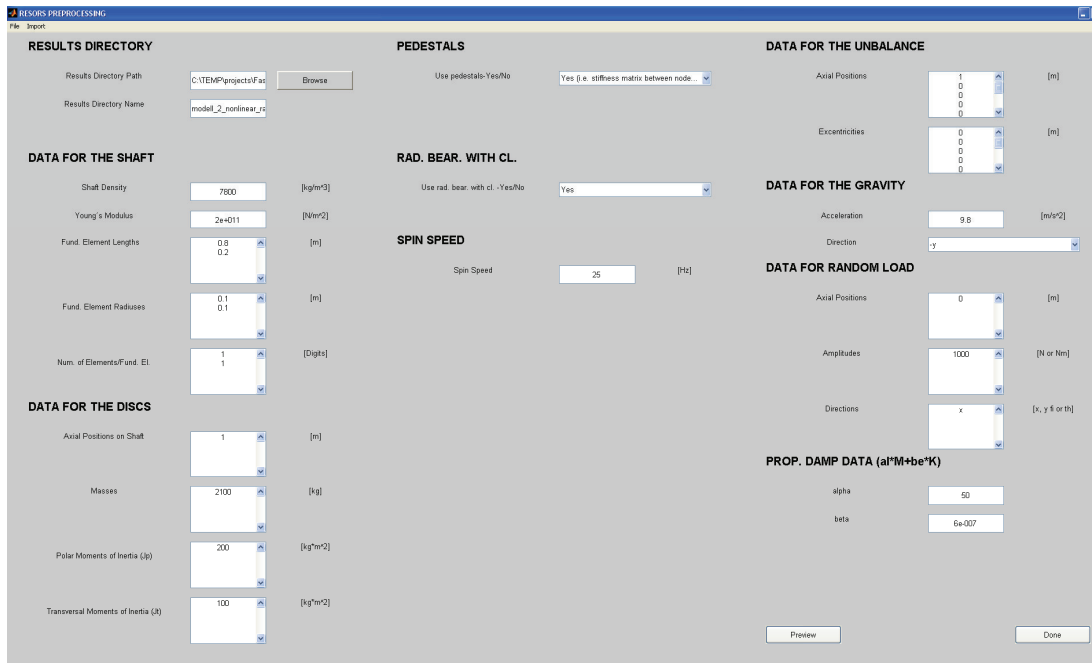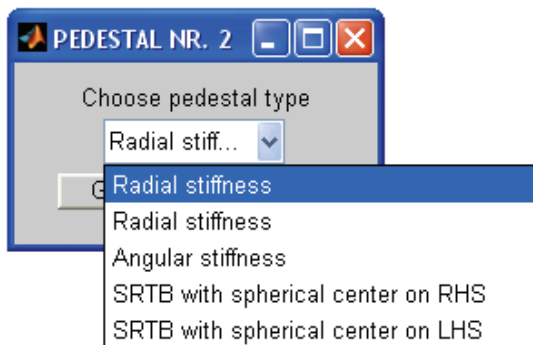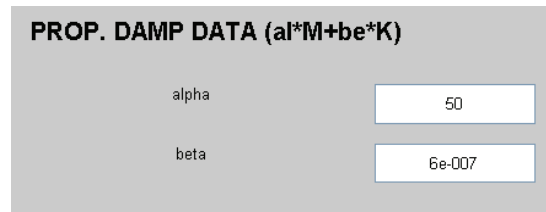

Figure 15. Three different mesh densities.

(a)

(b)

(c)

Figure 16. (a) Preprocessing GUI in RESORS. (b) Pedestal selection. (c) Damping ratio selection.

within a few seconds. Figure 17 (page 293) shows how the eigenfrequencies as a function of spin speed (Campbell diagram) depend on the mesh density of the case studied (the black dots, triangles and plus markers). From this figure it can be concluded that, except for the second forward whirl mode at large spin speeds, the eigenfrequencies show little mesh density dependency.

In this scenario, a machine running just above the first forward whirl frequency is to be designed and, hence, even the coarsest mesh gives enough accuracy and will therefore be used henceforth.

Selection of pedestal stiffness and rotor mass. When a suitable mesh density has been obtained, the designer can test different concepts by changing parameters in the preprocessor, conduct numerical simulation (when needed) and postprocess and analyze the results. The blue ring in Figure 17 shows that the
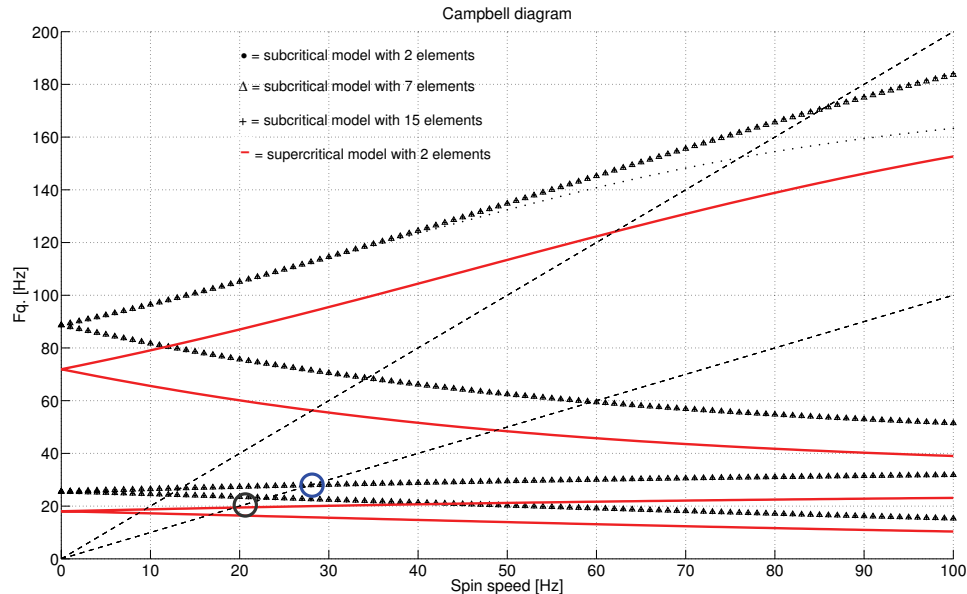
Figure 17. Campbell diagram for different mesh densities.

first critical speed, i.e., the intersection between the unbalance excitation frequency and the first forward whirl frequency (dashed line and black dotted line in Figure 17), is somewhere between 25 Hz and 30 Hz. Since a supercritical machine will be designed, the critical speed must be decreased below 25 Hz. The red lines in Figure 17 show the Campbell diagram when increasing the mass to 2100 kg and decreasing the pedestal stiffness to 5×10^7 N/m. The gray ring in Figure 17 shows that, by these design justifications, the critical speed is around 20 Hz and, hence, the machine will run supercritically. A set of parameters that give a suitable Campbell diagram can typically be found in a few minutes.

4.3.3. *Analyses by use of fully nonlinear simulation model.* Karlberg [2010] showed that the pedestal stiffness may depend on the amount of clearance in bearings. Karlberg et al. [2010] showed that this clearance can give rise to a significant decrease in resonance frequencies. It was also shown that due to clearance, the unbalance may excite both forward and backward whirl motion. Hence, in order to indicate that the machine will run supercritically as predicted by the linearized model, a fully nonlinear analysis including the clearance must be conducted.

Analysis of clearance dependency. In this step of the analysis process, the designer returns to the preprocessor and changes the linear pedestals with pedestals including bearings with clearance, which is done from drop-down lists. In RESORS the possibility of varying parameters has been implemented. This is done by choosing the parameter to be varied from a drop-down list, and then entering the parameter ranges and where on the model to store the data. Since the system is nonlinear, regular analysis measures such as eigenfrequencies cannot be used. Therefore, a transfer function approach has been implemented in RESORS, where random load is applied to the model leading to vibrations, which are measured. Figure 18 shows the transfer function of $\dot{\theta}$ as a function of the bearing clearance at the rotor when a random load was applied in the *x*-direction at the rear bearing. Red indicates high vibration amplitudes,
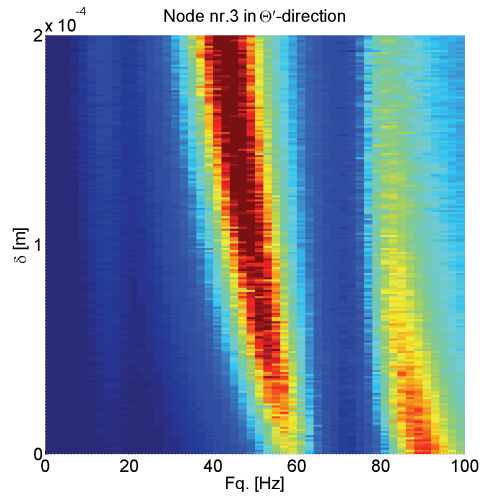
Figure 18. Resonance frequencies as a function of bearing clearance.

while blue indicates low vibration amplitudes. The model has a spin speed of 25 Hz Figure 18 shows that the first four eigenfrequencies at zero clearance are 16 Hz, 18 Hz, 58 Hz and 92 Hz.

Figure 18 shows that when the clearance in the bearings is increased to 0.2 mm the resonance frequencies decrease up to 20 Hz. The resonance frequencies within this clearance range do not coincide with the spin speed, but a clearance below 0.1 mm is still recommended, since the first multiple of the spin speed (50 Hz) may also excite the system and hence cause resonance. In order to verify that a suitable design has been developed, in the final step the designer runs an unbalance response simulation of the fully nonlinear model. The dashed line in Figure 19 is the unbalance response for the largest acceptable clearance (0.1 mm). As a reference, the unbalance response of the linearized model is also shown (solid line in Figure 19). Unlike the linear model, analysis of the nonlinear unbalance response shows that critical speeds can occur at more intersections than between the spin speed and the first forward whirl frequency.

The time needed to conduct the nonlinear analysis described above is usually around one hour.

4.4. *Rotor dynamical case discussion.* To show how simulation-driven rotor dynamical design can be conducted a tool named RESORS has been developed and evaluated. The proposed methodology consists of three steps: *gathering of information*, analysis by use of *linearized models*, and by use of *nonlinear models*.

In RESORS, information based on expert knowledge has been implemented, such as different pedestal types, choice of damping parameters, postprocessing tools as transfer functions for nonlinear systems, etc., meaning that the designer does not have to bother about rotor dynamical details. The GUI of RESORS has further been designed to be intuitive and easy to use, featuring e.g., clear descriptions of in-data fields and units, drop-down lists for specific choices, few steps in pre and postprocessing, etc. The post processor in RESORS is designed so that important engineering measures are directly derived without further data processing, e.g., Campbell diagrams, mode shapes, transfer functions and load responses (maximum vibration amplitudes).
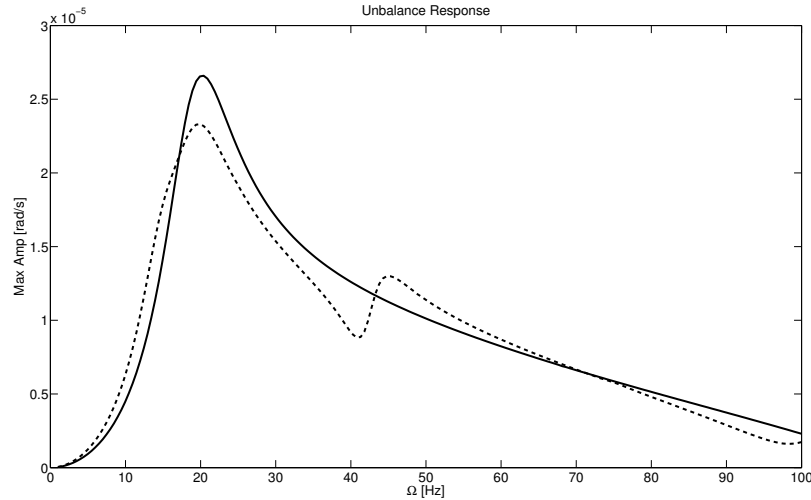
Figure 19. Unbalance response of linearized model (solid line) and nonlinear model (dashed line).

For the specific case used in this paper an overcritical rotating machine including bearings with clearance was to be developed, meaning that the system becomes strongly nonlinear and, hence, traditional analysis fails. By comparing the results from the linearized model (used to derive a suitable mesh density) with the results from the fully nonlinear model it was found that the resonance frequencies have a strong dependency on the clearance. Hence, a nonlinear analysis is necessary when designing rotating machinery supported on bearings with clearance. Although advanced equations and simulation strategies are required for this type of analysis, with RESORS, analysis is still easy for the designer.

The total time to finish one loop of the proposed three-step methodology is typically around one and a half hours, to be compared with traditional "over the wall" strategies, which usually take days or even weeks per iteration. Hence, by use of the proposed simulation-driven design methodology and effective and intuitive simulation software, a lot of time is saved (days or even weeks). This time can instead be used to test other possible solutions to the problem — in other words, a larger design space can be explored, thereby improving the innovation probability.

5. Conclusions

To show how tools for modeling and simulation of mechanics of materials and structures can be designed and used to enable simulation-driven product development two case studies have been used. In both cases, fundamental equations as discussed in Sections 3.1 and 4.1, serve as basis for the software design and development in this paper.

The first case study concerned welding simulations of a Volvo Construction Equipment wheel loader rear axle bridge (mechanics of materials), while the second case study concerned design of a supercritical rotating machine (mechanics of structures).

In the welding case study, a simulation-driven design methodology is applied in order to find suitable welding sequences for a Volvo Construction Equipment wheel loader rear axle bridge. It can be concluded

that by use of a calibrated mesh density and block dumping heat source, the efficiency of the welding simulations was improved with sufficient simulation accuracy.

A program (RESORS) enabling simulation-driven design has been developed and used for the rotor dynamical case study. This software has been designed for improved usability compared to commercial codes used for rotor dynamical analysis through the GUI, being developed purposely for the specific application. The software has been found to be intuitive, efficient and easy to use for designers and improves postprocessing and analysis even for challenging, nonlinear problems.

The methodology used in the two widely different cases both shows that the efficiency in the analysis of challenging problems can be significantly improved through developing advanced tools suitable for use by design engineers. Important features for enabling simulation design by designers are (i) a graphical use interface enabling simulation-driven design; (ii) software built on ambient expert knowledge, made usable for the designer; (iii) significantly decreased time from preprocessing to analysis compared to commercially available codes; and (iv) postprocessing and analysis made easy for designers.

These two examples of simulation-driven design by designers indicate that a larger design space can be explored and that more possible solutions can be evaluated. Therefore, the approach improves the probabilities of innovation and finding optimal solutions.

Furthermore, the calibrated block dumping results presented in Section 3.3 can be used to increase the efficiency of welding simulations when many simulations are required.

References

[Andersson 1978] B. A. B. Andersson, "Thermal stresses in a submerged-arc welded joint considering phase transformations", *J. Eng. Mater. Technol. (ASME)* **100**:4 (1978), 356–362.

[Åström 2004] P. Åström, *Simulation of manufacturing processes in product development*, Ph.D. thesis, Luleå University of Technology, Luleå, 2004, Available at http://epubl.ltu.se/1402-1544/2004/56/index-en.html.

[Barroso et al. 2010] A. Barroso et al., "Prediction of welding residual stresses and displacements by simplified models: experimental validation", *Mater. Des.* **31**:3 (2010), 1338–1349.

[Bylund 2004] N. Bylund, *Simulation driven product development applied to car body design*, Ph.D. thesis, Luleå University of Technology, Luleå, 2004, Available at http://epubl.luth.se/1402-1544/2004/33/index-en.html.

[Cai and Zhao 2003] Z. Cai and H. Zhao, "Efficient finite element approach for modelling of actual welded structures", *Sci. Technol. Weld. Join.* **8**:3 (2003), 195–204.

[Chen and Sheng 1992] Y. Chen and I. C. Sheng, "Residual stress in weldment", *J. Therm. Stresses* **15**:1 (1992), 53–69.

[Childs 1982] D. W. Childs, "Fractional-frequency rotor motion due to nonsymmetric clearance effects", *J. Eng. Power* **104**:3 (1982), 533–541.

[Childs 1993] D. Childs, *Turbomachinery rotordynamics: phenomena, modeling and analysis*, Wiley, New York, 1993.

[Chu and Zhang 1997] F. Chu and Z. Zhang, "Periodic, quasi-periodic and chaotic vibrations of a rub-impact rotor system supported on oil film bearings", *Int. J. Eng. Sci.* **35**:10–11 (1997), 963–973.

[Courter 2009] B. Courter, "Simulation-driven product development: will form finally follow function?", White paper, Space-Claim, Concord, MA, 2009, Available at http://files.spaceclaim.com/whitepapers/SPC.CAE.WhitePaper.web2.pdf.

[Cross 2000] N. Cross, *Engineering design methods: strategies for product design*, Wiley, Chichester, 2000.

[Deng et al. 2009] D. Deng, K. Ogawa, S. Kiyoshima, N. Yanagida, and K. Saito, "Prediction of residual stresses in a dissimilar metal welded pipe with considering cladding, buttering and post weld heat treatment", *Comput. Mater. Sci.* **47**:2 (2009), 398–408.

[Edwards et al. 1999] S. Edwards, A. W. Lees, and M. I. Friswell, "The influence of torsion on rotor/stator contact in rotating machinery", *J. Sound Vib.* **225**:4 (1999), 767–778.

[Ganesan 1996] R. Ganesan, "Dynamic response and stability of a rotor-support system with non-symmetric bearing clearances", *Mech. Mach. Theory* **31**:6 (1996), 781–798.

[Genta 1999] G. Genta, *Vibration of structures and machines: practical aspects*, Springer, New York, 1999.

[Gero 1981] J. S. Gero, "Design optimization", *Comput.-Aided Des.* **13**:5 (1981), 252.

[Goldak 2009] J. A. Goldak, "Distortion and residual stress in welds: the next generation", pp. 45–52 in *Trends in welding research: proceedings of the 8th International Conference* (Pine Mountain, GA, 2008), edited by S. A. David et al., ASM International, Materials Park, OH, 2009.

[Goldak and Akhlaghi 2005] J. A. Goldak and M. Akhlaghi, *Computational welding mechanics*, Springer, New York, 2005.

[Goldak et al. 1986] J. A. Goldak, B. Patel, M. Bibby, and J. Moore, "Computational weld mechanics", *AGARD Conf. Proc.* **398** (1986), 1–32.

[Goldak et al. 2007] J. A. Goldak et al., "Designer driven analysis of welded structures", pp. 1025–1038 in *Mathematical modelling of weld phenomena* 8 (Graz, 2006), edited by H.-H. Cerjak et al., Technische Universität Graz, Graz, 2007.

[Goldak Technologies 2010] Goldak Technologies, "VrWeld", 2010, Available at http://www.goldaktec.com/vrweld.html.

[Goldman and Muszynska 1995] P. Goldman and A. Muszynska, "Rotor-to-stator, rub-related, thermal/mechanical effects in rotating machinery", *Chaos Solitons Fractals* **5**:9 (1995), 1579–1601.

[Hansen 1974] H. R. Hansen, "Application of optimization methods within structural design: practical design example", *Comput. Struct.* **4**:1 (1974), 213–220.

[Harris 1991] T. A. Harris, *Rolling bearing analysis*, Wiley, New York, 1991.

[Jeffcott 1919] H. H. Jeffcott, "XXVII. The lateral vibration of loaded shafts in the neighbourhood of a whirling speed.–The effect of want of balance", *Philos. Mag.* (6) **37**:219 (1919), 304–314.

[Karlberg 2010] M. Karlberg, "Approximated stiffness coefficients in rotor systems supported by bearings with clearance", *Int. J. Rotat. Mach.* **2010** (2010). Article ID 540101.

[Karlberg and Aidanpää 2003] M. Karlberg and J.-O. Aidanpää, "Numerical investigation of an unbalanced rotor system with bearing clearance", *Chaos Solitons Fractals* **18**:4 (2003), 653–664.

[Karlberg and Aidanpää 2004] M. Karlberg and J.-O. Aidanpää, "Investigation of an unbalanced rotor system with bearing clearance and stabilising rods", *Chaos Solitons Fractals* **20**:2 (2004), 363–374.

[Karlberg and Aidanpää 2007] M. Karlberg and J.-O. Aidanpää, "Rotordynamical modelling of a fibre refiner during production", *J. Sound Vib.* **303**:3–5 (2007), 440–454.

[Karlberg et al. 2010] M. Karlberg, M. Karlsson, L. Karlsson, and M. Näsström, "Dynamics of rotor systems with clearance and weak pedestals in full contact", in *13th International Symposium on Transport Phenomena and Dynamics of Rotating Machinery: ISROMAC 13* (Honolulu, HI, 2010), International Symposium on Transport Phenomena and Dynamics of Rotating Machinery **13**, Curran Associates, Red Hook, NY, 2010.

[Karlsson 1986] L. Karlsson, "Thermal stresses in welding", pp. 299–389 in *Thermal stresses I*, edited by R. B. Hetnarski, Elsevier, Amsterdam, 1986.

[Karlsson et al. 1989] L. Karlsson, M. Jonsson, L.-E. Lindgren, M. Näsström, and L. Troive, "Residual stresses and deformations in a welded thin-walled pipe", pp. 7–11 in *Weld residual stresses and plastic deformation* (Honolulu, HI, 1989), edited by E. Rybicki et al., Pressure Vessels and Piping **173**, ASME, New York, 1989.

[LaCourse 1995] D. E. LaCourse, *Handbook of solid modeling*, McGraw-Hill, New York, 1995.

[Lee et al. 2008] C.-H. Lee, K.-H. Chang, and C.-Y. Lee, "Comparative study of welding residual stresses in carbon and stainless steel butt welds", *Proc. Inst. Mech. Eng. B, J. Eng. Manuf.* **222**:12 (2008), 1685–1694.

[Lindgren 2001a] L.-E. Lindgren, "Finite element modeling and simulation of welding, 2: Improved material modeling", *J. Therm. Stresses* **24**:3 (2001), 195–231.

[Lindgren 2001b] L.-E. Lindgren, "Finite element modeling and simulation of welding, 1: Increased complexity", *J. Therm. Stresses* **24**:2 (2001), 141–192.

[Lindgren 2001c] L.-E. Lindgren, "Finite element modeling and simulation of welding, 3: Efficiency and integration", *J. Therm. Stresses* **24**:4 (2001), 305–334.

[Lindgren 2007] L.-E. Lindgren, *Computational welding mechanics: thermomechanical and microstructural simulations*, CRC Press, Boca Raton, FL, 2007.

[Mochizuki et al. 2000] M. Mochizuki, M. Hayashi, and T. Hattori, "Numerical analysis of welding residual stress and its verification using neutron diffraction measurement", *J. Eng. Mater. Technol.* (*ASME*) **122**:1 (2000), 98–103.

[Muszynska and Goldman 1995] A. Muszynska and P. Goldman, "Chaotic responses of unbalanced rotor/bearing/stator systems with looseness or rubs", *Chaos Solitons Fractals* **5**:9 (1995), 1683–1704.

[Ogawa et al. 2009] K. Ogawa, D. Deng, S. Kiyoshima, N. Yanagida, and K. Saito, "Investigations on welding residual stresses in penetration nozzles by means of 3D thermal elastic plastic FEM and experiment", *Comput. Mater. Sci.* **45**:4 (2009), 1031–1042.

[Pahkamaa et al. 2010] A. Pahkamaa, L. Karlsson, and J. Pavasson, "A method to improve efficiency in welding simulations for simulation driven design", pp. 81–90 in *Proceedings of the ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference* (*IDETC/CIE2010*)*, 3: 30th Computers and Information in Engineering Conference* (Montreal, 2010), ASME, New York, 2010.

[Ramm 1981] E. Ramm, *Nonlinear finite element analysis in structural mechanics: strategies for tracing the nonlinear response of near limit points*, Springer, Berlin, 1981.

[Rankine 1869] W. A. Rankine, "On the centrifugal force of rotating shafts", *Eng.* **27** (1869), 249.

[Saad 1996] Y. Saad, *Iterative methods for sparse linear systems*, PWS, Boston, 1996. 2nd ed. published by SIAM, Philadelphia, 2003.

[Sellgren 1995] U. Sellgren, *Simulation driven design: a functional view of the design process*, Licentiate thesis, Royal Institute of Technology, Stockholm, 1995, Available at http://www.md.kth.se/~ulfs/Research/LIC/LIC_main.pdf.

[Smith and Reinertsen 1998] P. G. Smith and D. G. Reinertsen, *Developing products in half the time: new rules, new tools*, Wiley, New York, 1998.

[Smith and Smith 2009a] M. C. Smith and A. C. Smith, "NeT bead-on-plate round robin: comparison of residual stress predictions and measurements", *Int. J. Press. Vessels Pip.* **86**:1 (2009), 79–95.

[Smith and Smith 2009b] M. C. Smith and A. C. Smith, "NeT bead-on-plate round robin: comparison of transient thermal predictions and measurements", *Int. J. Press. Vessels Pip.* **86**:1 (2009), 96–109.

[Troive et al. 1998] L. Troive, M. Näsström, and M. Jonsson, "Experimental and numerical study of multi-pass welding process of pipe-flange joints", *J. Press. Vessel Technol.* (*ASME*) **120**:3 (1998), 244–251.

[Truman and Smith 2009] C. E. Truman and M. C. Smith, "The NeT residual stress measurement and modelling round robin on a single weld bead-on-plate specimen", *Int. J. Press. Vessels Pip.* **86**:1 (2009), 1–2.

[Ueda and Yuan 1993] Y. Ueda and M. G. Yuan, "Prediction of residual stresses in butt welded plates using inherent strains", *J. Eng. Mater. Technol.* (*ASME*) **115**:4 (1993), 417–423.

[Ueda et al. 1986] Y. Ueda, K. Fukuda, and Y. C. Kim, "New measuring method of axisymmetric three-dimensional residual stresses using inherent strains as parameters", *J. Eng. Mater. Technol.* (*ASME*) **108**:4 (1986), 328–334.

[Ueda et al. 1988] Y. Ueda, Y. C. Kim, and M. G. Yuan, "A predicting method of welding residual stress using source of residual stress", *Q. J. Jpn. Weld. Soc.* **6**:1 (1988), 59–64.

[Ulrich and Eppinger 1995] K. T. Ulrich and S. D. Eppinger, *Product design and development*, McGraw-Hill, New York, 1995.

[Voutchkov et al. 2005] I. Voutchkov, A. J. Keane, A. Bhaskar, and T. M. Olsen, "Weld sequence optimization: the use of surrogate models for solving sequential combinatorial problems", *Comput. Methods Appl. Mech. Eng.* **194**:30–33 (2005), 3535–3551.

[Wall 2007] J. Wall, *Simulation-driven design of complex mechanical and mechatronic systems*, Ph.D. thesis, Blekinge Institute of Technology, Karlskrona, 2007, Available at http://tinyurl.com/Wall-2007-diss.

[Wheelwright and Clark 1992] S. C. Wheelwright and K. B. Clark, *Revolutionizing product development: quantum leaps in speed, efficiency and quality*, Free Press, New York, 1992.

[Wriggers and Simo 1990] P. Wriggers and J. C. Simo, "A general procedure for the direct computation of turning and bifurcation points", *Int. J. Numer. Methods Eng.* **30**:1 (1990), 155–176.

Appendix: Design space exploration: simulated welding sequences

Sequence number	Welding sequence	WeldPath	nbSubPasses	StartWF 1st | 2nd	EndWF 1st | 2nd	StartTime (s) 1st | 2nd
Dump and Mesh Calibration		a (1) b (2) c (3,4) d (5,6)	1 1 2 2	0.0 1.0 0.0 | 1.0 0.0 | 1.0	1.0 0.0 0.5 | 0.5 0.5 | 0.5	0 158 317 | 396 475 | 554
1		a (1) b (2) c (3) d (4)	1 1 1 1	0.0 0.0 0.0 0.0	1.0 1.0 1.0 1.0	0 158 317 475
2		a (1) b (3) c (2) d (4)	1 1 1 1	0.0 0.0 0.0 0.0	1.0 1.0 1.0 1.0	0 317 158 475
3		a (1) b (2) c (3) d (4)	1 1 1 1	0.0 1.0 0.0 1.0	1.0 0.0 1.0 0.0	0 158 317 475
4		a (1) b (3) c (2) d (4)	1 1 1 1	0.0 1.0 0.0 1.0	1.0 0.0 1.0 0.0	0 317 158 475
5		a (1,2) b (3,4) c (5,6) d (7,8)	2 2 2 2	0.0 | 1.0 0.0 | 1.0 0.0 | 1.0 0.0 | 1.0	0.5 | 0.5 0.5 | 0.5 0.5 | 0.5 0.5 | 0.5	0 | 79 158 | 237 317 | 396 475 | 554
6		a (1,2) b (3,4) c (5,6) d (7,8)	2 2 2 2	0.5 | 0.5 0.5 | 0.5 0.5 | 0.5 0.5 | 0.5	0.0 | 1.0 0.0 | 1.0 0.0 | 1.0 0.0 | 1.0	0 | 79 158 | 237 317 | 396 475 | 554
7		a (1,2) b (3,4) c (5,6) d (7,8)	2 2 2 2	0.5 | 0.5 0.5 | 0.5 0.0 | 1.0 0.0 | 1.0	0.0 | 1.0 0.0 | 1.0 0.5 | 0.5 0.5 | 0.5	0 | 79 158 | 237 317 | 396 475 | 554
8		a (1,2) b (3,4) c (5,6) d (7,8)	2 2 2 2	0.5 | 0.5 0.0 | 1.0 0.5 | 0.5 0.0 | 1.0	0.0 | 1.0 0.5 | 0.5 0.0 | 1.0 0.5 | 0.5	0 | 79 158 | 237 317 | 396 475 | 554
9		a (1,3) b (2,4) c (5,7) d (6,8)	2 2 2 2	0.0 | 1.0 1.0 | 0.0 0.0 | 1.0 1.0 | 0.0	0.5 | 0.5 0.5 | 0.5 0.5 | 0.5 0.5 | 0.5	0 | 158 79 | 237 317 | 475 396 | 554

10		a (1,3)	2	0.5 | 0.5	0.0 | 1.0	0 | 158
		b (2,4)	2	0.5 | 0.5	1.0 | 0.0	79 | 237
		c (5,7)	2	0.5 | 0.5	0.0 | 1.0	317 | 475
		d (6,8)	2	0.5 | 0.5	1.0 | 0.0	396 | 554
11		a (1,2)	2	0.5 | 0.0	1.0 | 0.5	0 | 79
		b (3,4)	2	0.5 | 0.0	1.0 | 0.5	158 | 237
		c (5,6)	2	0.5 | 0.0	1.0 | 0.5	317 | 396
		d (7,8)	2	0.5 | 0.0	1.0 | 0.5	475 | 554
12		a (1,2)	2	0.5 | 1.0	0.0 | 0.5	0 | 79
		b (3,4)	2	0.5 | 1.0	0.0 | 0.5	158 | 237
		c (5,6)	2	0.5 | 0.0	1.0 | 0.5	317 | 396
		d (7,8)	2	0.5 | 0.0	1.0 | 0.5	475 | 554
13		a (1,2)	2	0.5 | 1.0	0.0 | 0.5	0 | 79
		b (3,4)	2	0.5 | 0.0	1.0 | 0.5	158 | 237
		c (5,6)	2	0.5 | 1.0	0.0 | 0.5	317 | 396
		d (7,8)	2	0.5 | 0.0	1.0 | 0.5	475 | 554
14		a (1,3)	2	0.0 | 1.0	0.5 | 0.5	0 | 158
		b (2,4)	2	0.0 | 1.0	0.5 | 0.5	79 | 237
		c (5,7)	2	0.0 | 1.0	0.5 | 0.5	317 | 475
		d (6,8)	2	0.0 | 1.0	0.5 | 0.5	396 | 554
15		a (1,3)	2	0.5 | 0.5	0.0 | 1.0	0 | 158
		b (2,4)	2	0.5 | 0.5	0.0 | 1.0	79 | 237
		c (5,7)	2	0.5 | 0.5	0.0 | 1.0	317 | 475
		d (6,8)	2	0.5 | 0.5	0.0 | 1.0	396 | 554
16		a (1,4)	2	0.0 | 1.0	0.5 | 0.5	0 | 237
		b (2,3)	2	0.0 | 1.0	0.5 | 0.5	79 | 158
		c (5,8)	2	0.0 | 1.0	0.5 | 0.5	317 | 554
		d (6,7)	2	0.0 | 1.0	0.5 | 0.5	396 | 475
17		a (1,4)	2	0.5 | 0.5	0.0 | 1.0	0 | 237
		b (2,3)	2	0.5 | 0.5	0.0 | 1.0	79 | 158
		c (5,8)	2	0.5 | 0.5	0.0 | 1.0	317 | 554
		d (6,7)	2	0.5 | 0.5	0.0 | 1.0	396 | 475
18		a (1,5)	2	0.0 | 1.0	0.5 | 0.5	0 | 317
		b (2,6)	2	0.0 | 1.0	0.5 | 0.5	79 | 396
		c (3,7)	2	0.0 | 1.0	0.5 | 0.5	158 | 475
		d (4,8)	2	0.0 | 1.0	0.5 | 0.5	237 | 554
19		a (1,5)	2	0.5 | 0.5	0.0 | 1.0	0 | 317
		b (2,6)	2	0.5 | 0.5	0.0 | 1.0	79 | 396
		c (3,7)	2	0.5 | 0.5	0.0 | 1.0	158 | 475
		d (4,8)	2	0.5 | 0.5	0.0 | 1.0	237 | 554
20		a (1)	1	1.0	0.0	0
		b (2)	1	0.0	1.0	158
		c (4,6)	2	1.0 | 0.0	0.5 | 0.5	396 | 554
		d (3,5)	2	1.0 | 0.0	0.5 | 0.5	317 | 475

LENNART KARLSSON: Lennart.Karlsson@ltu.se
The Faste Laboratory, Division of Computer Aided Design, Luleå University of Technology, SE-97187 Luleå, Sweden
http://www.ltu.se/tfm/cad/home/d743/d18287/leka?l=en

ANDREAS PAHKAMAA: Andreas.Pahkamaa@ltu.se
Division of Computer Aided Design, LuleåUniversity of Technology, SE-97187 Luleå, Sweden
http://www.ltu.se/tfm/cad/home/d743/d18287/1.52354?l=en

MAGNUS KARLBERG: Magnus.Karlberg@ltu.se
The Faste Laboratory, Division of Computer Aided Design, Luleå University of Technology, SE-97187 Luleå, Sweden
http://www.ltu.se/tfm/cad/home/d743/d18287/magkar?l=en

MAGNUS LÖFSTRAND: Magnus.Lofstrand@ltu.se
The Faste Laboratory, Division of Computer Aided Design, Luleå University of Technology, SE-97187 Luleå, Sweden
http://www.ltu.se/tfm/cad/home/d743/d18287/maglof?l=en

JOHN GOLDAK: jgoldak@mrco2.carleton.ca
Department of Mechanical and Aerospace Engineering, Carleton University, Ottawa, ON K1S 5B6, Canada

JONAS PAVASSON: Jonas.Pavasson@ltu.se
The Faste Laboratory, Division of Computer Aided Design, Luleå University of Technology, SE-97187 Luleå, Sweden
http://www.ltu.se/tfm/cad/home/d743/d18287/jonpav?l=en

DISSIPATION ENERGY AS A STIMULUS FOR CORTICAL BONE ADAPTATION

Natarajan Chennimalai Kumar, Iwona Jasiuk and Jonathan Dantzig

We present a finite element study of a poroelastic rectangular beam subjected to oscillatory bending loads. This geometric model is chosen for simplicity, as an idealized representation of cortical bone. We then propose the use of the dissipation energy of the poroelastic flow as a mechanical stimulus for bone adaptation, and show that it can predict the effect of frequency of the applied load. Surface adaptation in the model depends on the weighted average of the mechanical stimulus in a "zone of influence" near each surface point, in order to incorporate the non-locality in the mechanotransduction of osteocytes present in the lacunae. We show that the dissipation energy stimulus and the resulting increase in second moment of inertia of the cross section increase linearly with frequency in the low frequency range (less than 10 Hz) and saturate at the higher frequency range (greater than 10 Hz). Similar non-linear adaptation frequency response also has been observed in numerous experiments. Our framework is readily extended to the modeling of cortical bone using actual bone geometries.

1. Introduction

Bone is a mechanically sensitive biological tissue, which adapts its size, shape, mass and density based on its mechanical environment [Cowin 2001]. It has been long recognized that a dynamic stimulus is required for bone adaptation [Rubin and Lanyon 1984; Turner 1998; Lanyon and Rubin 1984]. Researchers have shown that the adaptation depends on a combination of different mechanical stimuli such as the magnitude [Rubin and Lanyon 1984; Burr et al. 2002; Lanyon et al. 1982] and frequency of applied load [Burr et al. 2002; Lanyon et al. 1982; Hsieh and Turner 2001; Warden and Turner 2004], number of cycles [Rubin and Lanyon 1984; Turner 1998], and bouts of the applied loading [Robling et al. 2000; Robling et al. 2001; 2002]. The adaptation response is initiated when the applied strain exceeds a threshold value, and increases with the magnitude of applied strain [Rubin and Lanyon 1984; Lanyon et al. 1982]. Turner et al. [1994] observed significant cortical bone adaptation when the loading frequency exceeded 0.5 Hz. At low frequencies (between 0.5 and 10 Hz), it has been observed that adaptation in rat ulnae follows an approximately linear dose-response relationship with frequency [Hsieh and Turner 2001]. Warden and Turner [2004] found no significant increase in the adaptation response when the frequency of loading was increased beyond 10 Hz. Rubin and coworkers investigated the effect of very low magnitude high frequency (greater than 30 Hz) loading on sheep standing on a vibrating plate 20 minutes per day for a year. Trabecular bone volume increased more than 30%, but no significant changes were found in the cortical bone [Rubin et al. 2001; 2002]. Qin et al. [2001; 2000; 1999] showed that cortical bone responds to applied pressure gradients in the intermedullary fluid through periosteal bone formation, without any applied mechanical loading.

Several groups have examined the relationship between mechanical loading on bone and the associated fluid flow inside the various internal porosities. Knothe Tate and Knothe observed the fluid flow in and out of cortical bone in sheep forearms under applied load [Knothe Tate and Knothe 2000]. Knothe Tate et al. showed similar load-induced fluid flow in rat tibia subjected to bending, and also indicated the role of the fluid flow in mechanotransduction [Knothe Tate et al. 2000]. Numerous mechanisms have been proposed to explain the effect of fluid flow on the osteocytes e.g., via shear stress on the cell surface [Reich et al. 1990], drag force on the transverse fibrils that tether the osteocyte to the canalicular walls [Weinbaum et al. 1994], strain generated electric potentials [Pollack et al. 1984; Salzstein and Pollack 1987], or biochemical diffusive gradients [Robling et al. 2008].

Weinbaum and co-workers developed an analytical model that describes the hierarchical nature of the lacunocanalicular porosity in the cortical bone, using the results from Zhang and Cowin [Zhang and Cowin 1994] at the macroscale, coupled to mechanical model at the cellular length-scale [Zhang and Cowin 1994; You et al. 2001; Han et al. 2004]. These works demonstrated that the fluid-induced shear stress and drag force on the osteocyte process tethering fibers can amplify local strains by more than 50 times compared to the tissue level strains [Han et al. 2004]. They also showed that the strain amplification factor varies non-linearly with respect to the loading frequency, in a similar way as the adaptation response observed in the experiments mentioned earlier. Fritton and Weinbaum's extensive review [2009] provides a more complete description of fluid flow induced mechanotransduction in cortical bone.

In [Chennimalai Kumar et al. 2010] we developed a framework for modeling cortical bone adaptation which included: elastic finite element (FE) analysis of loading applied to geometrically accurate models of bone, extraction of a mechanical stimulus from the FE results to be used in an adaptation law, and simulation of adaptation by direct modification of the FE model. The adaptation procedure is implemented in a generic way so that it can be used to simulate the effect of the different mechanical loading parameters such as magnitude, frequency, number of bouts of loading, time between bouts, and other factors. We showed that an elastic material model and a simple growth law using strain energy density as the mechanical stimulus could predict the effect of load magnitude and the number of bouts of loading on adaptation of rat ulnae [Chennimalai Kumar et al. 2010]. In this paper, we extend that work to incorporate poroelastic analysis. Previous works have included poroelasticity in the adaptation equation, using the fluid shear stress and octahedral strain to compute the tissue phenotype for fracture [Lacroix and Prendergast 2002; Prendergast et al. 1997]. We propose a new mechanical stimulus based on the viscous dissipation energy due to the fluid flow as the stimulus for adaptation. We show that this stimulus can simulate the dependence of adaptation on loading frequency similar to the adaptation response observed experimentally. We consider the idealized bone geometry in the shape of a rectangular beam (following [Zhang and Cowin 1994]). This allows us to focus on the frequency trends by conducting a parametric study with relatively small computational resources.

The paper is organized as follows: We begin with a brief description of the theory of poroelasticity as it applies to bone, and identify a problem with an analytical solution to test our formulation. We then develop a measure of the mechanical stimulus to be used in an adaptation law. The results of simulations performed for a range of frequencies are presented, and these results are then interpreted in the context of mechanotransduction.

2. Theory of poroelasticity

Cortical bone is a porous structure with different porosities at different length scales [Cowin 1999], the most important being:

Vascular porosity: which is comprised of the cylindrical pathways of blood flow in the Haversian and Volkmann canals that form the osteonal structure of the cortical bone. The fluid permeability at this scale is of the order of $10^{-12}\,\mathrm{m}^2$ [Zhang et al. 1998a]. This porosity plays an important role in adaptation by delivering nutrients to the cells that form or resorb bone. Since the blood flow is from the vascular system, the pressure of the fluid in the vascular porosity is of the order of the physiological blood pressure.

Lacunocanalicular porosity: consisting of the porous regions of the lacunae surrounding the center of the osteon and the interconnecting canaliculi. The stimuli-sensing osteocyte cells reside within the pores of the lacunae and their processes are housed in the canaliculi. Zhang et al. [1998b] estimated permeability at this scale of porosity to be $10^{-18} - 10^{-21}\,\mathrm{m}^2$.

Collagen-apatite porosity: is seen at the nanostructural scale. This porosity contains water which is considered to be part of bone's collagen-apatite nanostructure. The flow of water through the pores between the organic collagen and the apatite mineral can be neglected [Cowin 2001].

It has been shown in a number of experimental and analytical works that the mechanical loading acting on bone induces the flow of fluid in and out of these different porosities [Knothe Tate 2001]. Weinbaum and co-workers showed the importance of the lacunocanalicular porosity for mechanotransduction in bone, where they hypothesize that osteocytes are deformed through the viscous shear stress due to the fluid flow inside the canaliculi [Weinbaum et al. 1994; You et al. 2001; Han et al. 2004]. These studies motivate the use of a poroelastic material model for the cortical bone with lacunocanalicular porosity.

We provide here a short overview of the theory of poroelasticity to establish the context for the poroelastic analyses presented in this paper, and refer the interested reader to more comprehensive treatments in the literature, such as [Biot 1941; Biot and Willis 1957; Coussy 1995; Detournay and Cheng 1993]. A poroelastic medium is made up of a solid matrix and pores. We consider saturated media, in which all of the pores are filled with fluid. There are four field variables: stress tensor $\boldsymbol{\sigma}$, strain tensor $\boldsymbol{\epsilon}$, pore pressure p, and the variation in fluid content ζ. The medium is characterized by its porosity n_p, the bulk modulus of the solid K_s, shear modulus of the solid G, Poisson's ratio for the solid ν, and the bulk modulus of the fluid K_f. The constitutive equations for an isotropic linear poroelastic material are

$$2G\epsilon_{ij} = \sigma_{ij} - \left(\frac{\nu}{\nu+1}\right)\sigma_{kk}\delta_{ij} + \alpha\left(\frac{1-2\nu}{1+\nu}\right)p\delta_{ij}, \quad 2G\zeta = \alpha\left(\frac{1-2\nu}{1+\nu}\right)\left(\sigma_{kk} + \frac{3p}{B}\right),$$

where we have introduced two additional parameters: the Willis coefficient α, and Skempton's coefficient B. The Willis coefficient can be thought of as the ratio of fluid volume gained (or lost) in a poroelastic element due to volume change when loaded under drained condition ($p = 0$). It can be shown that $\alpha = 1 - K/K_s$, where K is the drained bulk modulus. Similarly, Skempton's coefficient B is obtained from the undrained condition ($\zeta = 0$) as $p = -B\sigma_{kk}/3$.

The fluid flow rate is computed from the pressure using Darcy's law,

$$q_i = -\kappa\frac{\partial p}{\partial x_i} \tag{1}$$
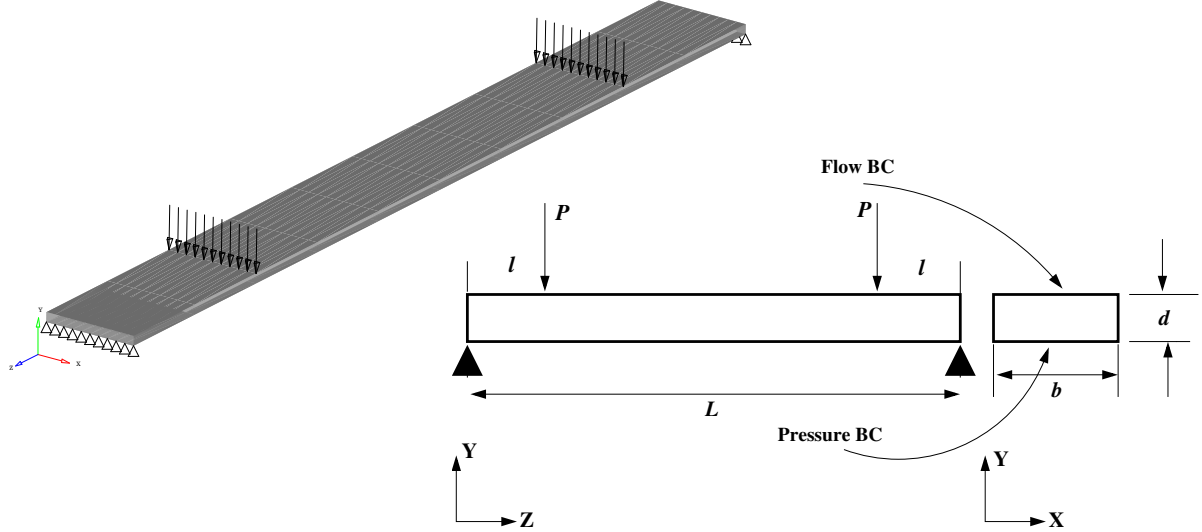
Figure 1. Schematic of the four-point bending setup of the rectangular poroelastic beam: isometric view (top left) and orthographic view with boundary conditions.

where the subscript i represents the coordinate direction, and q_i is the fluid mass flow rate and κ is the hydraulic permeability ($\kappa = k/\mu$, where k is the isotropic intrinsic permeability with units of m^2 and μ is the dynamic viscosity of the fluid). The fluid mass flow rate q_i is related to the fluid flow velocity v_i^{fl} as $q_i = n_p v_i^{\text{fl}}$, where ρ_{fl} is the density of the fluid.

Inserting the constitutive equations and Darcy's law into the mass and momentum balance equations leads to the following equation for the pore pressure:

$$c\nabla^2\left(\sigma_{kk} + \frac{3}{B}p\right) = \frac{\partial}{\partial t}\left(\sigma_{kk} + \frac{3}{B}p\right) \tag{2}$$

where c is the hydraulic diffusivity, defined as

$$c = \frac{2GB^2\kappa(1-\nu)(1+\nu_u)^2}{9(1-\nu_u)(\nu_u - \nu)} \tag{3}$$

where ν_u is the undrained Poisson's ratio. Scaling the length in all directions by a characteristic length d (that is, $X^* = x/d$), and scaling time using the frequency ω ($T^* = \omega t$), Equation (2) becomes

$$c\nabla^{*2}\left(\sigma_{kk} + \frac{3}{B}p\right) = \text{Fo}\frac{\partial}{\partial T^*}\left(\sigma_{kk} + \frac{3}{B}p\right), \tag{4}$$

where $\text{Fo} = \omega d^2/c$ is the Fourier number, which represents the ratio of the timescale for hydraulic diffusion (d^2/c) to the timescale of the applied load ($1/\omega$). For small Fourier number (Fo $\ll$ 1), the transient term on the right hand side can be neglected in comparison to the left hand side, and the pressure solution will be essentially quasi-static. For large values of Fourier number (Fo $\gg$ 1), the Laplacian on the left hand side of the equation can be neglected and the pore pressure follows the stress solution. When Fo is of order one, the two sides balance. This observation will help us explain the results of our simulations.

The objective of this paper is to establish a frequency-dependent mechanical stimulus. To this end, we use as a test problem the bending of a poroelastic rectangular beam, similar to that considered in [Zhang and Cowin 1994]. Since our eventual goal is to analyze geometrically accurate bone models, we solve this problem using the finite element (FE) method, as implemented in the commercial software ABAQUS [2008]. We consider a beam of length L, width b, and thickness d, subjected to a cyclic 4-point bending load of the form $P = P_0/2(1 - \cos \omega t)$. The geometry and boundary conditions (BC) are illustrated in Figure 1. In order to simulate the BCs of a 4-point beam bending, we set $u_y = 0$ at both ends of the beam, and constrain one point, such that $u_x = 0$ and $u_z = 0$ as well to avoid rigid body motion. The bending loads are at a distance l from each end of the beam, as shown. The periosteal (outer) surface of a long bone is highly impermeable and the endosteal (inner) surface is highly permeable [Steck et al. 2003]. To simulate these properties in our model, we apply a zero pressure BC on the bottom surface, and zero flow BC on the top surface. We note that the latter BC is actually implemented via Darcy's law (1) as $\nabla p \cdot \boldsymbol{n} = 0$, where $\boldsymbol{n}$ is the normal vector.

3. Development of the stimulus for the growth law

The general form of the growth law used to model the bone adaptation is

$$\frac{\mathrm{d}b}{\mathrm{d}T} = A(\phi - \phi_{\mathrm{ref}}) \tag{5}$$

where b is a material characteristic (such as density, mass or shape), T is the growth timescale, A is a proportionality constant which we refer to as "gain", ϕ is a mechanical stimulus, and ϕ_{ref} is the reference stimulus that must be exceeded to trigger bone growth. Different types of growth stimuli have been proposed in the literature, such as strain energy density [Weinans et al. 1992; Huiskes et al. 2000], strain [Cowin and Hegedus 1976], daily stress [Carter et al. 1989; Carter et al. 1996], and others. In [Chenni-malai Kumar et al. 2010] we used the strain energy density as the mechanical stimulus to numerically model the bone growth response in a rat ulna using the growth algorithm shown in Figure 2. We describe the approach here briefly.
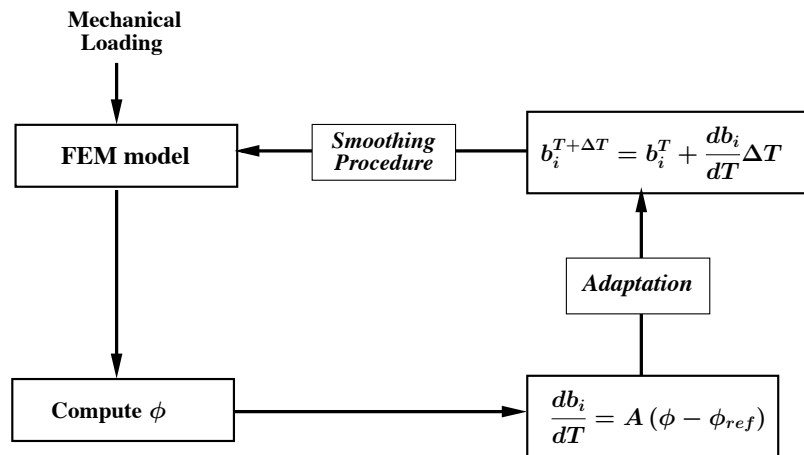
Figure 2. Flowchart describing the implementation of the adaptation procedure.
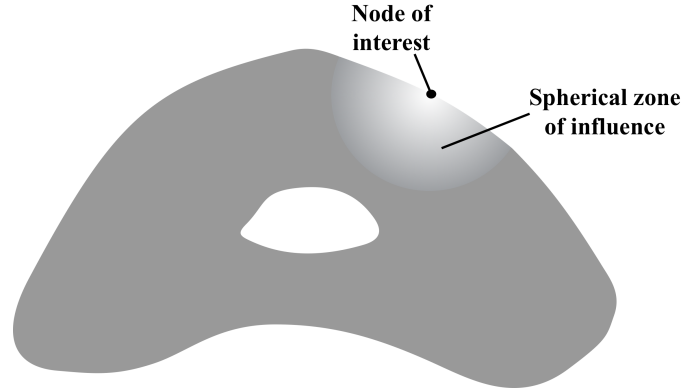
Figure 3. Schematic of the spherical zone of influence on a typical bone cross section. The shading is intended to show the decreasing influence of stimulus with distance from the surface node.

The FE model of the bone is created from micro-CT images, and a set of nodes on the outermost surface are identified as movable. An elastic analysis is carried out to compute the stresses and strains in the bone under an applied load. The strain energy densities in each finite element, including all surface nodes, are determined, and the displacement of a surface node is calculated based on the growth law

$$\frac{\mathrm{d}b_i}{\mathrm{d}T} = \begin{cases} A(\phi_i - \phi_{\mathrm{ref}}) & \text{if } \phi_i \geq \phi_{\mathrm{ref}}, \\ 0 & \text{if } \phi_i < \phi_{\mathrm{ref}}. \end{cases} \tag{6}$$

where b_i is the displacement of the surface node i and ϕ_i is the local strain energy density at node i. Resorption was not observed in the experiments that were simulated [Robling et al. 2002], so we precluded resorption in our model. Note that the growth-timescale T is much longer than the timescale of loading (t in the poroelastic field equations). Once the rates of change of the position of the surface nodes are computed, the new positions of the nodes are computed using a forward Euler scheme for (6), followed by application of a smoothing filter. See [Chennimalai Kumar et al. 2010] for details. The new FE model is then constructed and the procedure is repeated. We were able to find significant agreement between the model and experiments using the strain energy density stimulus, for an elastic material model and quasi-static loading. However, the strain energy density cannot simulate the effect of loading frequency. Our goal in this work is to develop a stimulus similar to the strain energy density that is dependent on frequency.

We would like to have a stimulus that captures the fluid flow in the bone, and reduces it to a convenient scalar quantity. To that end, we choose the dissipation due to the viscous fluid flow [Coussy 1995], defined as

$$\phi = -n_p \boldsymbol{v}^{\mathrm{fl}} \cdot \nabla p = \tfrac{1}{2}\big(n_p \boldsymbol{v}^{\mathrm{fl}}\big) \cdot \boldsymbol{\kappa}^{-1} \cdot \big(n_p \boldsymbol{v}^{\mathrm{fl}}\big) \tag{7}$$

where $\boldsymbol{v}^{\mathrm{fl}}$ is the fluid velocity vector, ∇p is the pressure gradient, and $\boldsymbol{\kappa}$ is the hydraulic permeability tensor. We choose this form because it has been shown in a number of experiments, and hypothesized in analytical models, that the shear stress exerted by the fluid flow on the osteocyte cells is a possible candidate for mechanotransduction. Note that in the work that follows, we will take κ to be isotropic.

Osteocytes are interconnected through the processes inside the canalicular space. We hypothesize that they can communicate with each other through these processes and exchange information on the current state of the stimulus at each location. To simulate this non-local behavior, we propose the use of a spatially averaged stimulus over a "zone of influence" (ZOI), shown schematically in Figure 3. For the sake of simplicity, we choose a spherical zone of influence of radius r. The stimulus that triggers the growth response at a surface node i is then defined as

$$\phi_i = \frac{\int_V \left(\int_0^T \frac{1}{2} n_p \boldsymbol{v}^{\mathrm{fl}} \cdot \kappa^{-1} \cdot n_p \boldsymbol{v}^{\mathrm{fl}} \, \mathrm{d}t \right) f(|\boldsymbol{x}|) \, \mathrm{d}V}{\int_V f(|\boldsymbol{x}|) \, \mathrm{d}V} \tag{8}$$

where $f(|\boldsymbol{x}|)$ is a function that weights the dissipation potential at an inner node by its distance $|\boldsymbol{x}|$ from the surface node i, and V is the volume of the zone of influence. Mullender et al. [1994] introduced a similar zone of influence in a study of internal remodeling, in which they used the strain energy in the neighborhood of each point, weighted by an exponential function, as the stimulus for remodeling. In the results presented below, we chose the form, $f(|\boldsymbol{x}|) = \exp(-5|\boldsymbol{x}|/d)$, where d is the thickness of the beam. The factor five was selected to ensure that any spurious high velocities at the inner surface, resulting from the boundary condition, do not affect the response. Numerical experiments showed that the results were not very sensitive to this factor, as long as it was larger than five.

4. Simulation details

The various material parameters used in the rectangular beam simulations, tabulated in Table 1, were chosen to be similar to properties of rat bone. The material properties given in the table are similar to those used in [Zhang and Cowin 1994]. The permeability of the lacunocanalicular porosity is reported to be in the range of 10^{-18} to 10^{-21} m^2 [Weinbaum et al. 1994; Salzstein and Pollack 1987]. Permeability will be shown to be a very important parameter in this problem, which directly affects the timescale of fluid diffusion in the poroelastic medium. The value of permeability we chose results in a significant range of Fourier numbers (from 12 to 380) for the physiological range of frequencies (1 Hz to 30 Hz). The hydraulic diffusivity corresponding to the permeability given in Table 1 is $c = 0.5606$ mm^2/s. One can then estimate the characteristic time for hydraulic diffusion in a 1 mm thick beam to be of the order of one second.

Property	Value	Units
Young's modulus of bone E	12.0	GPa
Fluid bulk modulus K_f	2.3	GPa
Solid bulk modulus K_s	17.0	GPa
Porosity n_p	0.05	–
Drained Poisson's ratio ν	0.3	–
Intrinsic permeability k	3×10^{-20}	m^2

Table 1. Values of different poroelastic parameters used in the analysis.
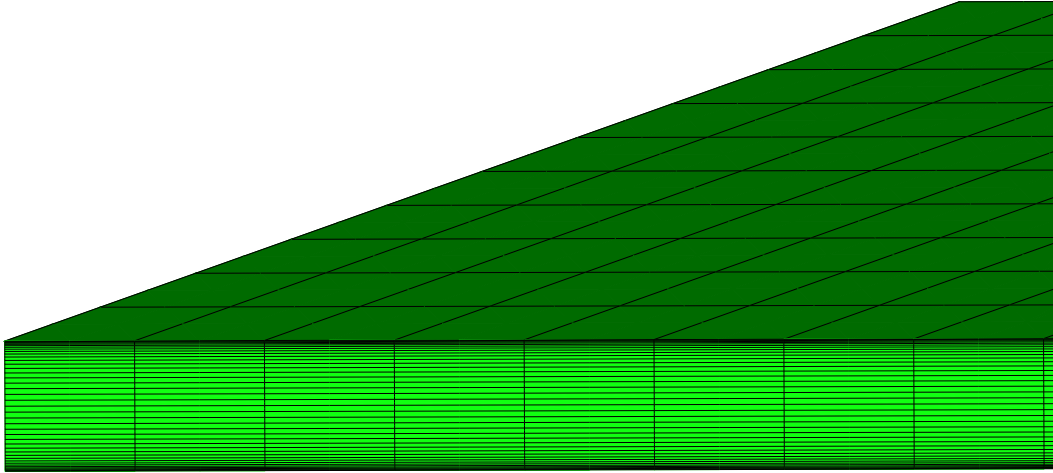
Figure 4. Finite element mesh of the rectangular beam. Note the biased mesh across the cross-section.

The FE mesh for our analyses, shown in Figure 4, consists of 18,821 nodes and 4000 triquadratic hexahedral elements. The elements are graded along the y-axis in order to accurately enforce the zero-flow boundary condition. A grid convergence study showed that this grid resulted in velocities of the order 10^{-4} mm/s on the top surface, which when compared to velocities in the bulk (of the order of 1 mm/s) were judged to be sufficiently small. The poroelastic solver in ABAQUS [2008] uses an unconditionally stable backward Euler integration scheme to solve (2). The accuracy of the solution still depends on the size of the timestep. We performed a convergence study to establish the timestep size as well, and we found that accurate results were obtained for $\Delta t < 0.5 h^2/c$, where h is the characteristic length of the smallest element. For the grid and properties used in the simulations reported below, this criterion gives $\Delta t = 5 \times 10^{-4}$ s. The applied load P_0 was chosen as 100 N.

5. Results

Figure 5 shows the time evolution of pore pressure and the fluid velocity at different points in the rectangular cross section for several different frequencies. In each case, there is an initial transient in both the pore pressure and fluid velocity that dies out after about 0.5 sec, which is of the order of the timescale of hydraulic diffusion noted earlier. After this time, pore pressure and velocity at each point follow a sinusoidal profile with the same frequency as the applied load. This is important, because it allows us to extrapolate the pore pressure and velocity solutions at the end of the initial transient to the entire duration of the experiments, which may be several minutes [Warden and Turner 2004], thus reducing the computational cost significantly. The velocities are maximal at the bottom surface, falling to zero at the top. Note that both the pore pressure at the upper surface, and the velocity at the lower surface, are out of phase with their respective values at the neutral axis.

Figure 6 shows the pore pressure and velocity profiles through the thickness for frequencies at 1 Hz, 5 Hz and 30 Hz, at different times within a single load cycle after the initial transient. We denote the
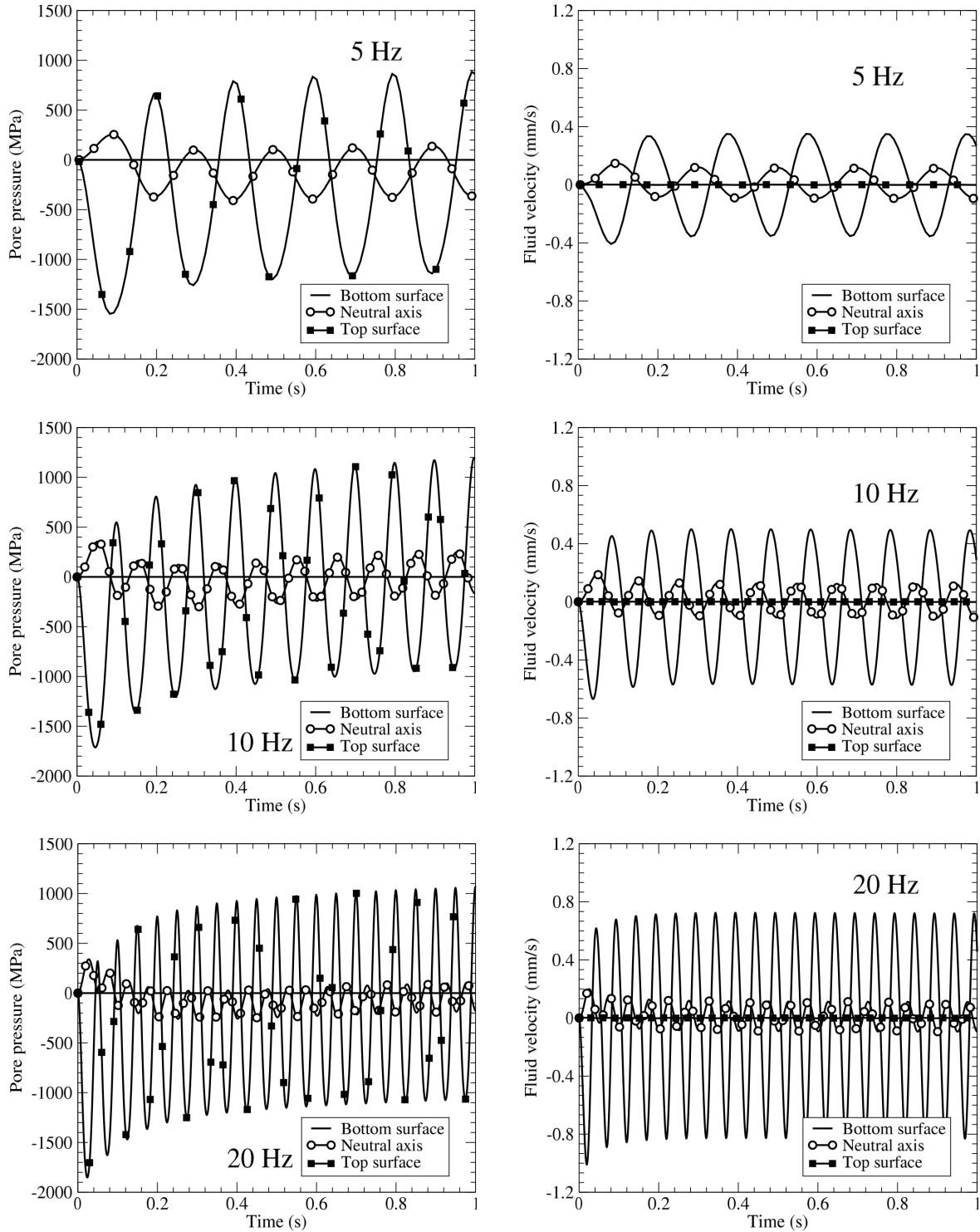
Figure 5. Plots of pore pressure (left column) and fluid flow velocity (right column) solutions as a function of time for three frequencies: 5 Hz, 10 Hz, and 20 Hz.
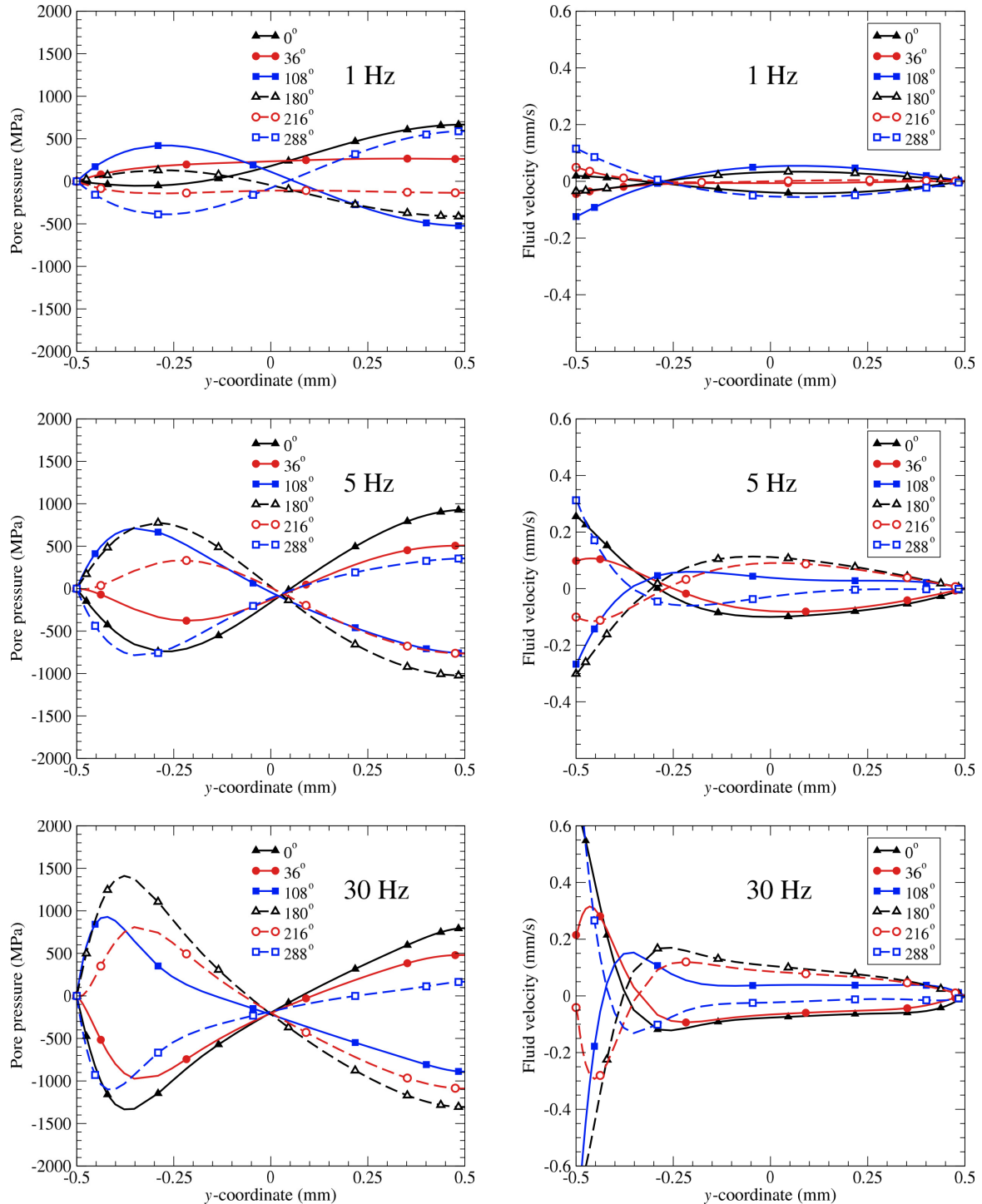
Figure 6. Plots of pore pressure (left column) and fluid flow velocity (right column) solutions across the cross-section at different points of time in a loading cycle (shown here as the phase angle) for three frequencies: 1 Hz, 5 Hz and 30 Hz.
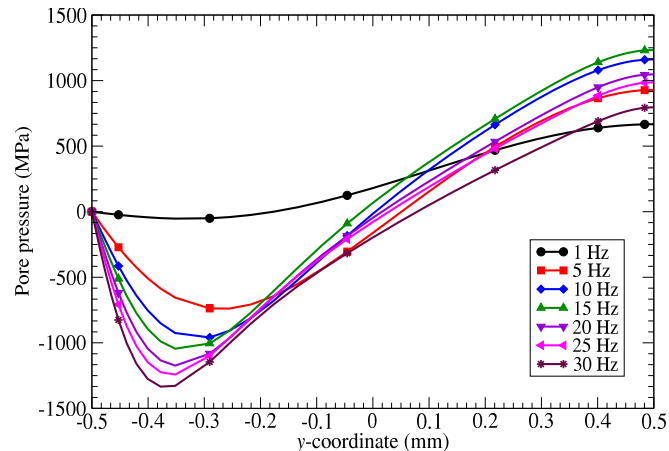
Figure 7. Plot of pore pressure solution across the cross-section for different frequencies at the end of a loading cycle.

various times within the cycle in degrees, where zero corresponds to the maximum in the load cycle. At 1 Hz (Figure 6, top left), the slope of the pore pressure distribution reaches zero midway between the bottom surface and the neutral axis, indicating zero flow velocity there. Figure 6, top right, which shows the corresponding fluid velocity, indicates that the velocity on the bottom surface is of a higher magnitude and has a smooth variation throughout the profile. The zero velocity solution is observed at the same location throughout the loading period.

At 5 Hz (Figure 6, middle), the pore pressure remains very close to zero at the neutral axis throughout the load cycle. The location of the zero-slope in the pressure solution (zero velocity by Darcy's law) varies with time of loading, and the velocity is nearly uniform over much of the beam cross-section above the neutral axis for parts of the load cycle. These effects are more pronounced at 30 Hz (Figure 6, bottom). There is very large variation in pore pressure near the bottom surface, which then falls to almost zero at the neutral axis. The velocity is also very large near the bottom surface, and then nearly uniform (and much smaller) over most of the remainder of the cross-section. Note that the computed velocities and pore pressures are considerably larger than those reported in [Rémond et al. 2008; Nguyen et al. 2010]. This is due to the application of a much larger load in our analysis. These quantities should be proportional to the magnitude of the applied load, because the problem is linear. We will address this issue in a future paper, in which we apply the current analysis to the deformation of a rat ulna.

The results of our calculations are summarized in Figures 7 and 8, which compare the pore pressure and fluid velocity through the thickness for different frequencies. As the frequency increases, the variation in pore pressure increases near the bottom surface, and becomes close to linear between $y = -0.3$ mm to $y = 0.3$ mm of the thickness. It should also be noted that the magnitude of pore pressure near the top surface (where the pressure gradient is zero) increases for frequencies from 1 to 10 Hz, but starts decreasing for frequencies greater than 10 Hz. Figure 8 shows that for frequencies less than 10 Hz, the flow is significant to about 50% of the thickness, i.e., up to the neutral axis, whereas at higher frequencies, even though the velocity at the bottom surface increases, the flow velocity becomes very small before reaching 25% of the thickness.
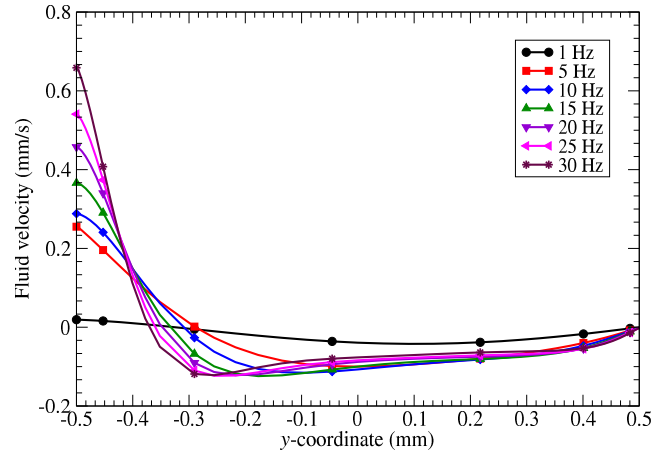
Figure 8. Plot of the fluid velocity pressure solution across the cross-section for different frequencies at the end of a loading cycle.

The dissipation energy, given by the time integral of the dissipation potential (7), was then computed from the solutions extrapolated to 100 s using the results at the end of the initial transient. Figure 9 shows that the dissipation energy in the beam follows a similar trend as the fluid velocity. The dissipation energy at the bottom surface increases with frequency. Since the bottom surface is farthest from the top surface, where significant adaptation occurs, we hypothesize that the effect of the dissipation energy at the bottom surface will be very small on the adaptation at the top. To implement this hypothesis in our analysis, we choose the exponential weighting function described earlier, and also shown in Figure 9.
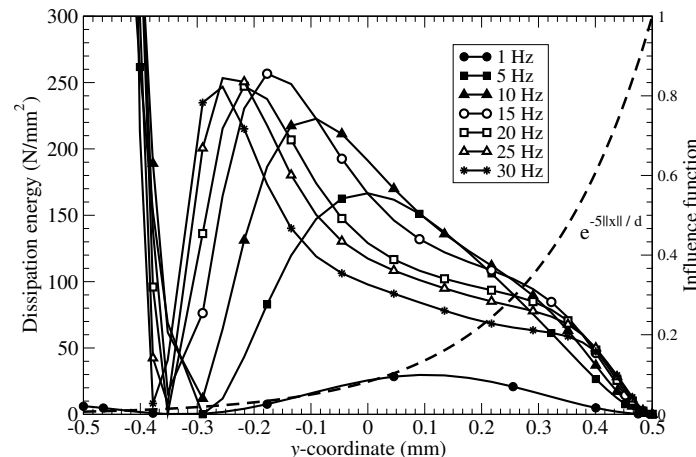
Figure 9. Plot of dissipation energy (time integral of dissipation potential) as a function of the y-coordinate for different frequencies. The influence function is also superimposed on the plot. We chose $\exp(-5\|x\|/d)$ as the influence function for our simulations. We have zoomed in on the plot to show the variation in the dissipation energy inside the beam, and so the very high values near $y = -0.5$ mm are cut off.
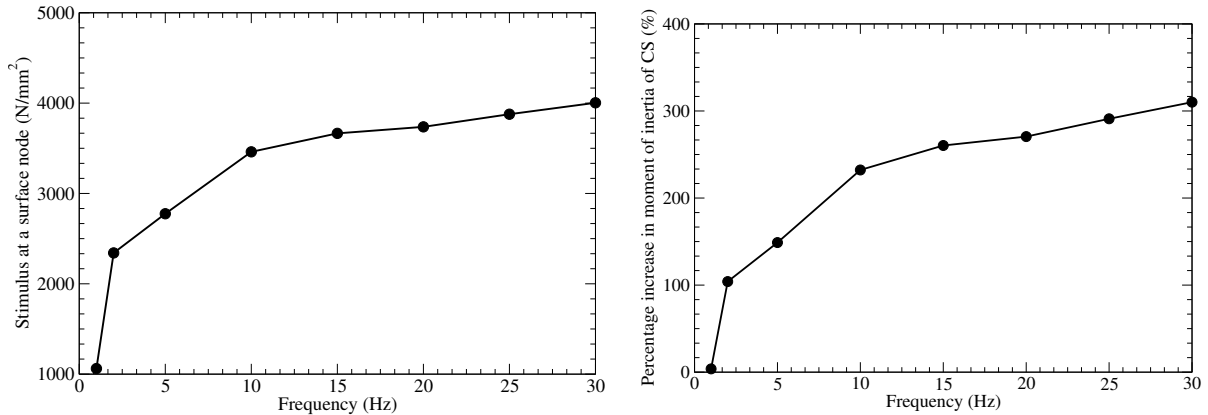
Figure 10. Left: plot of dissipation energy stimulus at a surface node as a function of loading frequency. Right: plot of percentage increase in the moment of inertia of the rectangular cross section due to adaptation as a function of frequency.

Figure 10, left, shows the adaptation stimulus (8) at a top surface node at the midspan of the beam as a function of the frequency of loading. We can see clearly that the adaptation stimulus increases rapidly at low frequencies, then saturates at about 10 Hz. We then used the growth law (6) to simulate growth of the rectangular beam for two growth timesteps, using a sensitivity $\phi_{\mathrm{ref}} = 1000\,\mathrm{N/mm^2}$ and a gain $A = 0.01\,\mathrm{mm^3/N/timestep}$. We express the results as the percentage change in the moment of inertia of the rectangular cross section due to the adaptation. Figure 10, right, shows that the growth depends on frequency in a manner similar to the stimulus. This behavior is typical of that observed in numerous experiments; see [Burr et al. 2002; Lanyon et al. 1982; Hsieh and Turner 2001], and especially [Warden and Turner 2004].

6. Discussion

We have proposed the dissipation energy of the poroelastic flow induced by mechanical loading as the stimulus to trigger adaptation in cortical bone. We implemented this model using a rectangular beam because: (a) the geometry is simple and so the pressure and velocity solutions can be understood in greater detail; (b) and the problem has an analytical solution, which makes it easier to understand the role of the various physical parameters.

We found that at low frequencies, the flow first increases with loading frequency, but that as the frequency increases beyond 10 Hz, the amount of fluid that penetrates through the thickness of the beam decreases. This means that at the higher frequencies, there is less flow seen by the osteocytes in the lacunae, and hence they experience smaller shear and drag forces. The dissipation energy represents the work done by these forces, and thus at higher frequencies, the dissipation energy stimulus saturates, and hence one can expect the adaptation to saturate as well. Weinbaum and co-workers [Zhang and Cowin 1994; You et al. 2001; Han et al. 2004] analyzed the shear and drag forces on the osteocyte cell membrane due to the poroelastic flow at the cellular level, and showed that the amplification of the strains on the osteocyte cell membrane has a similar frequency response to our results for dissipation energy stimulus and increase in moment of inertia (Figure 10). The proposed dissipation energy stimulus that

we have put forth in this paper is thus qualitatively consistent with the strain amplification hypothesis and experimental observations.

We have also proposed the use of a "zone of influence" in order to simulate the effect of the distributed network of osteocytes and their communication. Osteocytes have been identified as the mechanosensory cells in the cortical bone. It has not yet been proven experimentally whether the osteocytes communicate with each other, and if so, what their mode of communication is. A series of experiments by Turner and coworkers on the expression of the sclerostin protein in response to *in vivo* mechanical loading suggest that there is such osteocyte communication, and also suggest the existence of a zone of influence. Robling et al. [2008] found that the expression of sclerostin protein, which is found almost exclusively in the osteocyte cells, is reduced considerably by *in vivo* mechanical loading, and further that the reduction in the sclerostin expression is much greater in the regions experiencing higher strains. The most pertinent observation for our work is that the expression of sclerostin is reduced on a group of osteocyte cells near the regions where higher bone growth is observed. This is a possible indication of the existence of a zone of influence within a real bone. Further experimental investigation of the expression of sclerostin at the scale of the osteocyte network could be used to inform the model as to the shape of the zone of influence.

7. Conclusions

We have performed poroelastic analyses on a simplified geometric model using the commercial FE software ABAQUS. We investigated the variations in pore pressure and fluid velocity with time and location in the beam, and their dependence on frequency. Based on these results, we propose the use of the dissipation energy as a mechanical stimulus for adaptation that can accommodate the effect of frequency. We also included the effect of non-locality of mechanotransduction of osteocytes present in the lacunae in the cortical bone through the use of a zone of influence. The dissipation energy stimulus evaluated in this manner is shown to increase linearly with frequency in the low frequency range, and saturate at the higher frequency range. The implementation of the poroelastic material model and the dissipation energy stimulus can be seamlessly integrated into our framework to simulate adaptation response on cortical bone. We are in the process of extending the poroelastic material model and the dissipation energy stimulus to the actual rat ulna FE model, and to quantitatively validate the numerical model with experimental observations.

Acknowledgments

References

[ABAQUS 2008] *ABAQUS Documentation*, Version 6.7, SIMULIA, Providence, RI, 2008, available at http://tinyurl.com/ ABAQUS-docs-v6-7.

[Biot 1941] M. A. Biot, "General theory of three-dimensional consolidation", *J. Appl. Phys.* **12**:2 (1941), 155–164.

[Biot and Willis 1957] M. A. Biot and D. G. Willis, "The elastic coefficients of the theory of consolidation", *J. Appl. Mech. (ASME)* **24** (1957), 594–601.

[Burr et al. 2002] D. B. Burr, A. G. Robling, and C. H. Turner, "Effects of biomechanical stress on bones in animals", *Bone* **30**:5 (2002), 781–786.

[Carter et al. 1989] D. R. Carter, T. E. Orr, and D. P. Fyhrie, "Relationships between loading history and femoral cancellous bone architecture", *J. Biomech.* **22**:3 (1989), 231–244.

[Carter et al. 1996] D. R. Carter, M. C. H. van der Meulen, and G. S. Beaupré, "Mechanical factors in bone growth and development", *Bone* **18**:1, Supplement 1 (1996), S5–S10.

[Chennimalai Kumar et al. 2010] N. Chennimalai Kumar, J. A. Dantzig, I. M. Jasiuk, A. G. Robling, and C. H. Turner, "Numerical modeling of long bone adaptation due to mechanical loading: correlation with experiments", *Ann. Biomed. Eng.* **38**:3 (2010), 594–604.

[Coussy 1995] O. Coussy, *Mechanics of porous continua*, Wiley, Chichester, 1995.

[Cowin 1999] S. C. Cowin, "Bone poroelasticity", *J. Biomech.* **32**:3 (1999), 217–238.

[Cowin 2001] S. C. Cowin (editor), *Bone mechanics handbook*, 2nd ed., CRC Press, Boca Raton, FL, 2001.

[Cowin and Hegedus 1976] S. C. Cowin and D. H. Hegedus, "Bone remodeling, I: theory of adaptive elasticity", *J. Elasticity* **6**:3 (1976), 313–325.

[Detournay and Cheng 1993] E. Detournay and A. H.-D. Cheng, *Fundamentals of poroelasticity*, edited by C. Fairhurst, Pergamon, 1993.

[Fritton and Weinbaum 2009] S. P. Fritton and S. Weinbaum, "Fluid and solute transport in bone: flow-induced mechanotransduction", *Annu. Rev. Fluid Mech.* **41**:1 (2009), 347–374.

[Han et al. 2004] Y. Han, S. C. Cowin, M. B. Schaffler, and S. Weinbaum, "Mechanotransduction and strain amplification in osteocyte cell processes", *Proc. Nat. Acad. Sci. USA* **101**:47 (2004), 16689–16694.

[Hsieh and Turner 2001] Y.-F. Hsieh and C. H. Turner, "Effects of loading frequency on mechanically induced bone formation", *J. Bone Miner. Res.* **16**:5 (2001), 918–924.

[Huiskes et al. 2000] R. Huiskes, R. Ruimerman, G. H. van Lenthe, and J. D. Janssen, "Effects of mechanical forces on maintenance and adaptation of form in trabecular bone", *Nature* **405**:6787 (2000), 704–706.

[Knothe Tate 2001] M. L. Knothe Tate, "Interstitial fluid flow", pp. 22:1–22:29 in *Bone mechanics handbook*, 2nd ed., edited by S. C. Cowin, CRC Press, Boca Raton, FL, 2001.

[Knothe Tate and Knothe 2000] M. L. Knothe Tate and U. Knothe, "An ex vivo model to study transport processes and fluid flow in loaded bone", *J. Biomech.* **33**:2 (2000), 247–254.

[Knothe Tate et al. 2000] M. L. Knothe Tate, R. Steck, M. R. Forwood, and P. Niederer, "In vivo demonstration of load-induced fluid flow in the rat tibia and its potential implications for processes associated with functional adaptation", *J. Exp. Biol.* **203**:18 (2000), 2737–2745.

[Lacroix and Prendergast 2002] D. Lacroix and P. J. Prendergast, "A mechano-regulation model for tissue differentiation during fracture healing: analysis of gap size and loading", *J. Biomech.* **35**:9 (2002), 1163–1171.

[Lanyon and Rubin 1984] L. E. Lanyon and C. T. Rubin, "Static vs. dynamic loads as an influence on bone remodelling", *J. Biomech.* **17**:12 (1984), 897–905.

[Lanyon et al. 1982] L. E. Lanyon, A. E. Goodship, C. J. Pye, and J. H. MacFie, "Mechanically adaptive bone remodelling", *J. Biomech.* **15**:3 (1982), 141–154.

[Mullender et al. 1994] M. G. Mullender, R. Huiskes, and H. Weinans, "A physiological approach to the simulation of bone remodeling as a self-organizational control process", *J. Biomech.* **27**:11 (1994), 1389–1394.

[Nguyen et al. 2010] V.-H. Nguyen, T. Lemaire, and S. Naili, "Poroelastic behaviour of cortical bone under harmonic axial loading: a finite element study at the osteonal scale", *Med. Eng. Phys.* **32**:4 (2010), 384–390.

[Pollack et al. 1984] S. R. Pollack, N. Petrov, R. Salzstein, G. Brankov, and R. Blagoeva, "An anatomical model for streaming potentials in osteons", *J. Biomech.* **17**:8 (1984), 627–636.

[Prendergast et al. 1997] P. J. Prendergast, R. Huiskes, and K. Søballe, "Biophysical stimuli on cells during tissue differentiation at implant interfaces", *J. Biomech.* **30**:6 (1997), 539–548.

[Qin et al. 1999] Y.-X. Qin, K. J. McLeod, and C. T. Rubin, "Intramedullary pressure induced fluid flow in bone", pp. 490 in *Proceedings of the First Joint BMES/EMBS Conference: serving humanity, advancing technology* (Atlanta, GA, 1999), vol. 1, edited by S. M. Blanchard et al., IEEE, Piscataway, NJ, 1999.

[Qin et al. 2000] Y.-X. Qin, M. Cute, and C. T. Rubin, "Bone morphological adaptation induced by dynamic fluid flow in the absence of matrix strain", *Ann. Biomed. Eng.* **28**:SUPPL 1 (2000), 1–8.

[Qin et al. 2001] Y.-X. Qin, A. Saldanha, and T. Kaplan, "Oscillatory bone fluid flow and its role in initiating remodeling in the absence of matrix strain", pp. 51–52 in *2001 advances in bioengineering: presented at the 2001 ASME International Mechanical Engineering Congress and Exposition* (New York, 2001), edited by B. B. Lieber, BED **51**, ASME, New York, 2001. Paper # BED-23025.

[Reich et al. 1990] K. M. Reich, C. V. Gay, and J. A. Frangos, "Fluid shear stress as a mediator of osteoblast cyclic adenosine monophospate production", *J. Cell. Physiol.* **143**:1 (1990), 100–104.

[Rémond et al. 2008] A. Rémond, S. Naili, and T. Lemaire, "Interstitial fluid flow in the osteon with spatial gradients of mechanical properties: a finite element study", *Biomech. Model. Mechanobiol.* **7**:6 (2008), 487–495.

[Robling et al. 2000] A. G. Robling, D. B. Burr, and C. H. Turner, "Partitioning a daily mechanical stimulus into discrete loading bouts improves the osteogenic response to loading", *J. Bone Miner. Res.* **15**:8 (2000), 1596–1602.

[Robling et al. 2001] A. G. Robling, D. B. Burr, and C. H. Turner, "Recovery periods restore mechanosensitivity to dynamically loaded bone", *J. Exp. Biol.* **204**:19 (2001), 3389–3399.

[Robling et al. 2002] A. G. Robling, F. M. Hinant, D. B. Burr, and C. H. Turner, "Improved bone structure and strength after long-term mechanical loading is greatest if loading is separated into short bouts", *J. Bone Miner. Res.* **17**:8 (2002), 1545–1554.

[Robling et al. 2008] A. G. Robling, P. J. Niziolek, L. A. Baldridge, K. W. Condon, M. R. Allen, I. Alam, S. M. Mantila, J. Gluhak-Heinrich, T. M. Bellido, S. E. Harris, and C. H. Turner, "Mechanical stimulation of bone *in vivo* reduces osteocyte expression of Sost/sclerostin", *J. Biol. Chem.* **283**:9 (2008), 5866–5875.

[Rubin and Lanyon 1984] C. T. Rubin and L. E. Lanyon, "Regulation of bone formation by applied dynamic loads", *J. Bone Joint Surg. Am.* **66**:3 (1984), 397–402.

[Rubin et al. 2001] C. Rubin, A. S. Turner, S. Bain, C. Mallinckrodt, and K. McLeod, "Low mechanical signals strengthen long bones", *Nature* **412**:6847 (2001), 603–604.

[Rubin et al. 2002] C. Rubin, S. Judex, and M. Hadjiargyrou, "Skeletal adaptation to mechanical stimuli in the absence of formation or resorption of bone", *J. Musculoskelet. Neuronal Interact.* **2**:3 (2002), 264–267.

[Salzstein and Pollack 1987] R. A. Salzstein and S. R. Pollack, "Electromechanical potentials in cortical bone, II: experimental analysis", *J. Biomech.* **20**:3 (1987), 271–280.

[Steck et al. 2003] R. Steck, P. Niederer, and M. L. Knothe Tate, "A finite element analysis for the prediction of load-induced fluid flow and mechanochemical transduction in bone", *J. Theor. Biol.* **220**:2 (2003), 249–259.

[Turner 1998] C. H. Turner, "Three rules for bone adaptation to mechanical stimuli", *Bone* **23**:5 (1998), 399–407.

[Turner et al. 1994] C. H. Turner, M. R. Forwood, and M. W. Otter, "Mechanotransduction in bone: do bone cells act as sensors of fluid flow?", *FASEB J.* **8**:11 (1994), 875–878.

[Warden and Turner 2004] S. J. Warden and C. H. Turner, "Mechanotransduction in the cortical bone is most efficient at loading frequencies of 5–10 Hz", *Bone* **34**:2 (2004), 261–270.

[Weinans et al. 1992] H. Weinans, R. Huiskes, and H. J. Grootenboer, "The behavior of adaptive bone-remodeling simulation models", *J. Biomech.* **25**:12 (1992), 1425–1441.

[Weinbaum et al. 1994] S. Weinbaum, S. C. Cowin, and Y. Zeng, "A model for the excitation of osteocytes by mechanical loading-induced bone fluid shear stresses", *J. Biomech.* **27**:3 (1994), 339–360.

[You et al. 2001] L. You, S. C. Cowin, M. B. Schaffler, and S. Weinbaum, "A model for strain amplification in the actin cytoskeleton of osteocytes due to fluid drag on pericellular matrix", *J. Biomech.* **34**:11 (2001), 1375–1386.

[Zhang and Cowin 1994] D. Zhang and S. C. Cowin, "Oscillatory bending of a poroelastic beam", *J. Mech. Phys. Solids* **42**:10 (1994), 1575–1599.

[Zhang et al. 1998a] D. Zhang, S. Weinbaum, and S. C. Cowin, "Electrical signal transmission in a bone cell network: the influence of a discrete gap junction", *Ann. Biomed. Eng.* **26**:4 (1998), 644–659.

[Zhang et al. 1998b] D. Zhang, S. Weinbaum, and S. C. Cowin, "Estimates of the peak pressures in bone pore water", *J. Biomech. Eng.* (*ASME*) **120**:6 (1998), 697–703.

NATARAJAN CHENNIMALAI KUMAR: natarajan.uiuc@gmail.com
Department of Mechanical Science and Engineering, University of Illinois at Urbana-Champaign, 1206 West Green Street, Urbana, IL 61801, United States

IWONA JASIUK: ijasiuk@illinois.edu
Department of Mechanical Science and Engineering, University of Illinois at Urbana-Champaign, 1206 West Green Street, Urbana, IL 61801, United States
http://www.mechse.uiuc.edu/faculty/ijasiuk

JONATHAN DANTZIG: dantzig@illinois.edu
Department of Mechanical Science and Engineering, University of Illinois at Urbana-Champaign, 1206 West Green Street, Urbana, IL 61801, United States

msp

msp

SIZE-NORMALIZED ROBUSTNESS OF Dpp GRADIENT IN *DROSOPHILA* WING IMAGINAL DISC

Arthur D. Lander, Qing Nie, Benjamin Vargas and Frederic Y. M. Wan

Dedicated to Charles and Marie-Louise Steele for their years of outstanding management of IJSS and JoMMS, and for their more than forty years of warm friendship with the last author.

Exogenous environmental changes are known to affect the intrinsic characteristics of biological organisms. For instance, the synthesis rate of the morphogen decapentaplegic (Dpp) in a *Drosophila* wing imaginal disc has been found to double with an increase of $5.9°C$ in ambient temperature. If not compensated, such a change would alter the signaling Dpp gradient significantly and thereby the development of the wing imaginal disc. To learn how flies continue to develop "normally" under such an exogenous change, we formulate in this paper a spatially two-dimensional reaction-diffusion system of partial differential equations (PDE) that accounts for the biological processes at work in the *Drosophila* wing disc essential for the formation of signaling Dpp gradient. By way of this PDE model, we investigate the effect of the apical-basal thickness and anteroposterior span of the wing on the shape of signaling gradients and the robustness of wing development in an altered environment (including an enhanced morphogen synthesis rate). Our principal result is a delineation of the role of wing disc size change in maintaining the magnitude and shape of the signaling Dpp gradient. The result provides a theoretical basis for the observed robustness of wing development, preserving relative but not absolute tissue pattern, when the morphogen synthesis rate is significantly altered. A similar robustness consideration for simultaneous changes of multiple intrinsic system characteristics is also discussed briefly.

1. Morphogen gradients and temperature change

Morphogens (also known as *ligands* in biochemistry and developmental biology) are molecular substances that bind to cell surface receptors and other molecules. The concentration gradients of different morphogen-receptor complexes [Entchev et al. 2000; Gurdon and Bourillot 2001; Teleman and Cohen 2000] are known to be responsible for cell differentiation and patterning of biological tissues during the developmental phase of a biological organism. For a number of morphogen families, including Dpp in the wing imaginal disc of fruit flies of the genus *Drosophila*, it is well established experimentally and by analyses of appropriate mathematical models that the signaling gradients are formed by morphogens being transported from a localized source and binding to cell surface receptors downstream (see references cited in [Lander et al. 2002; 2005a; 2005b]). However, biochemical processes leading to morphogen gradient formation are influenced by highly cooperative events such as protein folding and membrane fluidity, and

exhibit considerable idiosyncratic effects from exogenous (external environmental) changes. More specifically, it has been observed that the wing size of an adult *Drosophila* may differ by about 15% for substantially different ambient temperature, with larger fly parts in a colder climate and smaller near the equator [de Moed et al. 1997; French et al. 1998; Bitner-Mathé and Klaczko 1999; Azevedo et al. 2002]. However, similar data on *Drosophila* wing imaginal discs relevant to our investigation do not seem to be available. (An *imaginal disc* is a group of undifferentiated cells that develops, at the (later) pupa stage, into a specific adult structure such as eyes, antennae, limbs and wings, with the developmental fate of cells in different zones of such a disc determined by a different combination of morphogen gradients.) At the molecular level, we expect correspondingly that rate constants for diffusion, protein synthesis, binding, internalization, degradation, etc., of fruit flies to change in varying degrees in response to such a temperature change. How significant would such rate constant changes be over the temperature ranges organisms encounter in the wild? Surprisingly little data on this subject (at least for insect cells) can be found in the literature. Preliminary results by metabolic labeling of *Drosophila* S2 cells (through the efforts in A. D. Lander's lab, particularly by his student S. Zhou) showed that overall protein synthesis rate doubles approximately every 5.9°C. In the absence of analogous data on biochemical processes other than protein synthesis rate, we cannot explicitly explore the effects of temperature on any of the known morphogen gradient systems.

The necessity to accommodate temperature (or any other exogenous environmental) change is a major performance objective of morphogen systems. A thorough quantitative analysis of the effects of such a change on the signaling morphogen gradients is imperative. It is therefore important to be clear the reasons why we do not (and cannot) undertake such an endeavor at this time. These include:

- Experimental and field study data are available for adult *Drosophila* while the impact of ambient changes (including signaling morphogen gradient formation) starts at the embryonic stage.

- An ambient change typically affects many systems characteristics including the various rate constants and synthesis rates) but there is a lack of data on most such effects.

Still, we may obtain some insights from exact or approximate analytical solutions, qualitative analyses, and numerical simulations of mathematical models for these systems. For example, we have recently looked at effects of perturbations in which every protein synthesis rate, every endo- and exocytotic rate constant, and every degradation rate constant is doubled, but diffusion and binding constants remain unchanged [Lander et al. 2005c]. We consider such changes to constitute a crude model of the effects due to a 5.9°C temperature increase, and define robustness measures to quantify the sensitivity of the system to the changes made.

In contrast to the work in [Lander et al. 2005c] and in [Khong and Wan 2007] on the effect of a Hill-type feedback (in Dpp synthesis rate) on the signaling Dpp gradient, we pursue a different and more limited objective prompted by the observations in the works cited at the top of the page. Specifically, we determine (by analyzing appropriate mathematical models) whether and how adjusting the size of the *wing disc* (an abbreviation for "*wing imaginal disc*" henceforth) would maintain the morphogen gradient shape and thereby provide optimal, or near-optimal, strategies for meeting the performance objective of ensuring a normal development. For this more limited goal, we take as our starting point the available experimental evidence (from the Lander lab) that the Dpp synthesis rate doubles with a 5.9°C temperature increase. With all other system characteristics fixed, a change in synthesis rate would cause a change in the steady state signaling morphogen gradient, possibly substantial and unacceptable from the view

point of normal development as quantified herein (see Section 8 and also [Lander et al. 2005c]). We determine whether a particular aspect of system architecture, namely the size of wing disc, confers the potential for the biological development to be robust to such change. More specifically, the principal aim of our research is to understand

- the reason for the exaggerated slenderness of the wing disc cells in the apical-basal direction, and
- the role of a size change of the wing disc in its development.

In the process, how the signaling Dpp gradient may be maintained in the face of significant Dpp synthesis rate changes (due to a temperature or any other exogenous environmental change) is delineated.

An interesting, but challenging aspect of robustness of biological development comes from the known interrelationships between temperature, growth, and morphogen signaling. At colder temperature, flies grow slower but end up larger, including having larger but otherwise normal wings; in contrast flies grow faster and are smaller in a hotter climate [de Moed et al. 1997; French et al. 1998; Bitner-Mathé and Klaczko 1999; Azevedo et al. 2002]. The remarkably normal patterning that they display [Bitner-Mathé and Klaczko 1999] is only normal in the context of their altered size. This strongly suggests that the objective of development is to preserve relative, not absolute, pattern. At the very least we need to examine our models to see if they would allow for temperature-dependent scaling of field sizes and *size-normalized* measures of robustness.

The effect of size on the robustness of biological species development has been investigated recently — see [Umulis et al. 2008] and references therein — for general models in the form of partial differential equations of the reaction diffusion type with Neumann or mixed conditions at the boundary of the solution domain. Their main concern is uncovering conditions on the biological system characteristics that would ensure the corresponding model problem to be *scale-invariant*. Within a scale-invariant species, a common structure in individuals of different size develops in proportion to size. (We will refer to such size-mediated developments of a common biological structure as *relative (or size-mediated) robustness* as the development of structural proportion is insensitivity to size changes.) Our concern here is with a specific model of the *Drosophila* wing imaginal disc in the plane of the proximal-distal and apical-basal axes which is *not* scale-invariant and does *not* satisfy the conditions for scale-invariance developed in [Umulis et al. 2008]. Nevertheless, we show how size changes may still be exploited for such a system to maintain *near* relative robustness for size-normalized development in the presence of a significant change in its system characteristics caused by environmental perturbations. Specific quantitative measures of (relative) robustness are adopted for determining (different levels of) robustness. We investigate at first only a change in the Dpp synthesis rate for which we have experimental data, and then also simultaneous changes in several system parameters including degradation and binding rate constants for multifactor robustness. The results are consistent with the observations of smaller *Drosophila melanogaster* flies near the earth's equator and larger one in colder climates away from the equator.

Still, we may obtain some insights from exact or approximate analytical solutions, qualitative analyses, and numerical simulations of mathematical models for these systems. As mentioned earlier, we have recently looked at effects of perturbations in which every protein synthesis rate, every endo- and exocytotic rate constant, and every degradation rate constant is doubled, but diffusion and binding constants remain unchanged [Lander et al. 2005c]. We consider such changes to constitute a crude model of the effects due to a 5.9°C temperature increase, and define robustness measures to quantify the sensitivity of the system to the changes made.

2. A spatially two-dimensional formulation

We take advantage of the approximate symmetry between the anterior compartment and posterior compartment of the *Drosophila* wing imaginal disc and consider here an extracellular model of the posterior compartment. With the Dpp synthesis rate taken to be uniform in the distal-proximal direction, the development is essentially uniform along the distal-proximal axis (*Y*-axis) except possibly for layer phenomena near the edges (see [Vargas 2007]). Our model of the Dpp gradient formation focuses on the uniform development. This allows us to consider a typical cross section of the wing imaginal disc as in Figure 1 idealized as a rectangle orthogonal to the *Y*-axis away from the distal and proximal ends. We take for this rectangle *X*-and *Z*-axis to be in the anteroposterior direction and the apical-proximal direction, respectively.

The new model is essentially an extended version of the one-dimensional model in [Lander et al. 2005b] to allow for variations in the apical-basal direction including biologically realistic Dpp leakage through the basement membrane. In this new two-dimensional model, morphogen is introduced into the extracellular space at a rate V_L through a localized source uniform in the plane perpendicular to the anteroposterior direction. The localized source spans a small interval $(-X_{\min}, 0)$, where $-X_{\min}$ is the
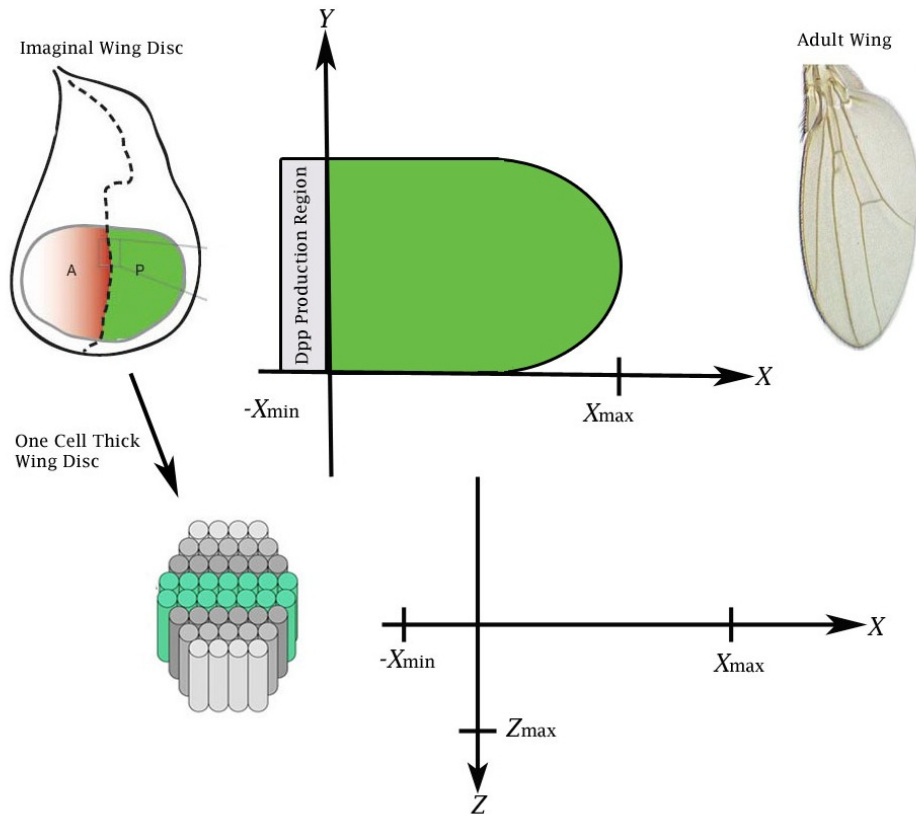
Figure 1. Wing imaginal disc and posterior compartment of *Drosophila*. Courtesy Dr. Oana Marcu of NASA Ames.

location of the border between (and the line of symmetry of) the *anterior* compartment and the *posterior* compartment of the wing disc. The morphogen produced in this localized region diffuses throughout the extracellular space in the posterior compartment (according to Fick's second law), between $Z = 0$ and $Z = Z_{max}$ in the apical-basal direction and from $X = -X_{min}$ toward the *sink* at the edge $X = X_{max}$ of the posterior compartment. Along the way, some morphogen molecules bind themselves with cell surface bound receptors at the binding rate $k_{on}[L][R]$, where $[L(X, Z, T)]$ and $[R(X, Z, T)]$ are, respectively, concentration of free Dpp and unoccupied signaling receptor Thickvein (Tkv) at time T and location (X, Z). The resulting morphogen-receptor complexes of concentration $[LR(X, Z, T)]$ are bound to cell surface membrane since the receptors are. These complexes in turn *dissociate* at the rate $k_{off}[LR]$ and degrade at the rate $k_{deg}[LR]$.

The time evolution of concentrations of free morphogen, morphogen-receptor complexes and unoccupied receptors is then described by the partial differential equations (1)–(3) below governing the rate of change of $[L]$, $[LR]$ and $[R]$, respectively, with k_{on}, k_{off}, and k_{deg} known as the binding rate constant, the dissociation rate constant, and the degradation rate constant, respectively. Altogether, the three (time) rate of changes in (1)–(3) account for the reversible binding, degradation (of both bound and unoccupied receptors), and synthesis of new morphogen and receptors, analogous to the one-dimensional system of [Lou et al. 2004; Lander et al. 2005b]:

$$\frac{\partial[L]}{\partial T} = D\left(\frac{\partial^2[L]}{\partial X^2} + \frac{\partial^2[L]}{\partial Z^2}\right) - k_{on}[L][R] + k_{off}[LR] + V_L(X, Z, T), \tag{1}$$

$$\frac{\partial[LR]}{\partial T} = k_{on}[L][R] - (k_{off} + k_{deg})[LR], \tag{2}$$

$$\frac{\partial[R]}{\partial T} = V_R(X, Z, T) - k_{on}[L][R] + k_{off}[LR] - k_R[R], \tag{3}$$

for $-X_{min} < X < X_{max}$, $0 < Z < Z_{max}$, and $T > 0$, where $V_L(X, Z, T)$ and $V_R(X, Z, T)$ are the rate at which the morphogen Dpp and receptors Tkv are synthesized, respectively, and k_R is the degradation rate constant for unoccupied Tkv. As in [Lander et al. 2005b], we are interested principally in the $[LR]$ gradient in the portion of the wing disc corresponding to $X > 0$ where there is no morphogen production (so that $V_L(X, Z, T) = 0$ for $X > 0$). In this paper, we focus on a time-invariant morphogen and receptor synthesis rates so that $V_L(X, Z, T) = V_L(X, Z)$ and $V_R(X, Z, T) = V_R(X, Z)$.

With $-X_{min}$ being the midpoint of the Dpp production region, we have by symmetry

$$\text{at } X = -X_{min}: \qquad \frac{\partial[L]}{\partial X} = 0, \tag{4}$$

for $T > 0$ and $0 < Z < Z_{max}$. At the far end, the edge of the posterior compartment is taken to be a sink, so that

$$\text{at } X = X_{max}: \qquad [L] = 0 \tag{5}$$

for $T > 0$ and $0 < Z < Z_{max}$. We also investigate the limiting case of $X_{max} = \infty$ since the edge at X_{max} is not strictly absorbing. In the apical-basal direction, we have an essentially sealed wall at the apical face $Z = 0$, so that

$$\text{at } Z = 0: \qquad \frac{\partial[L]}{\partial Z} = 0, \tag{6}$$

for $T > 0$ and $-X_{\min} < X < X_{\max}$, while at the basal face, there is leakage of free morphogens at a rate proportional to its concentration:

$$\text{at } Z = Z_{\max}: \qquad \frac{\partial [L]}{\partial Z} + \frac{\gamma_z}{Z_{\max}} [L] = 0, \qquad (7)$$

for $T > 0$ and $-X_{\min} < X < X_{\max}$. Here, the dimensionless constant $\gamma_z \equiv 1/\sigma_z = Z_{\max}\Gamma_z$ is a prescribed leakage parameter. In one extreme case $\gamma_z = 0$, we have $\partial [L]/\partial Z = 0$ so that the end is sealed without leakage. At the other extreme $\sigma_z = 0$, the end $Z = Z_{\max}$ is absorbing. For a finite nonzero γ_z, the larger the γ_z value the higher is the flux across the end surface.

At the onset of the morphogen synthesis (at $T = 0$), we have the initial conditions

$$[L] = [LR] = 0, \qquad [R] = R_0(X, Z) \qquad (8)$$

for $-X_{\min} \leq X \leq X_{\max}$ and $0 \leq Z \leq Z_{\max}$, reflecting the fact that there was no Dpp in the system and the receptor concentration is in a steady state (as a consequence of a time-invariant receptor synthesis rate).

The extracellular model above is adequate for our purpose. The model can be extended to incorporate the effects of internalization of $[LR]$ complexes through endocytosis prior to degradation as was done in [Lander et al. 2006; Lou et al. 2004] for one-dimensional studies. However, the corresponding system for steady state gradients of interest here has been shown to reduce to the same BVP with modified system rate constants (*loc. cit.*).

To reduce the number of parameters in the problem, we introduce a reference unoccupied receptor concentration level $\bar{R}_0$ (to be specified later) and the normalized quantities

$$t = \frac{D}{Z_{\max}^2} T, \qquad \{x, \ell_M, x_m, z\} = \frac{1}{Z_{\max}} \{X, X_{\max}, X_{\min}, Z\}, \qquad (9)$$

$$\{v_L, v_R\} = \frac{Z_{\max}^2}{D \bar{R}_0} \{V_L, V_R\}, \qquad \{a, b, r, r_0\} = \frac{1}{\bar{R}_0} \{[L], [LR], [R], R_0\}, \qquad (10)$$

$$\{f_z, g_z, g_r, h_z\} = \frac{1}{D/Z_{\max}^2} \{k_{\text{off}}, k_{\text{deg}}, k_R, k_{\text{on}} \bar{R}_0\}. \qquad (11)$$

In terms of these new quantities, we write the initial-boundary value problem (IBVP) (1)–(8) in the normalized form

$$\frac{\partial a}{\partial t} = \nabla^2 a - h_z a r + f_z b + v_L(x, z) \qquad (x, z) \in \Omega, \qquad (12)$$

$$\frac{\partial b}{\partial t} = h_z a r - (f_z + g_z)b, \qquad \frac{\partial r}{\partial t} = v_R(x, z) - h_z a r - g_r r + f_z b, \qquad (x, z) \in \bar{\Omega}, \qquad (13)$$

where Ω is the rectangular domain $\{-x_m < x < \ell_M, 0 < z < 1\}$ and $\bar{\Omega} = \{-x_m \leq x \leq \ell_M, 0 \leq z \leq 1\}$ is its closure. The auxiliary conditions supplementing the differential equations become

$$\text{at } x = -x_m: \quad \frac{\partial a}{\partial x} = 0 \qquad \text{at } x = \ell_M: \quad a = 0 \qquad (0 < z < 1), \qquad (14)$$

$$\text{at } z = 0: \quad \frac{\partial a}{\partial z} = 0 \qquad \text{at } z = 1: \quad \frac{\partial a}{\partial z} + \gamma_z a = 0 \qquad (-x_m < x < \ell_M), \qquad (15)$$

for $t > 0$ and

$$\text{at } t = 0: \qquad a = b = 0, \qquad r = r_0(x, z) \tag{16}$$

for all (x, z) in $\bar{\Omega}$.

3. Time-independent steady state behavior

3.1. *Time-independent synthesis rates.* With both morphogen and the receptor synthesis rates uniform in time, the possibility of a time-independent steady state behavior exists for our model. The two synthesis rates V_L and V_R are to a good approximation uniform in Z so that $V_L = V_L(X)$ and $V_R = V_R(X)$. For the present investigation, we ignore possible feedback effects and, unless indicated otherwise, approximate V_L to be a step function with $V_L(X) = \bar{V}_L H(-X)$ for some constant $\bar{V}_L$. Correspondingly, we have with

$$v_L = v_L(x) = \bar{v}_z H(-x) = \begin{cases} \bar{v}_z \\ 0 \end{cases} \qquad \text{with} \quad \bar{v}_z = \frac{\bar{V}_L / \bar{R}_0}{D/Z_{\max}^2}. \tag{17}$$

We also take the nonnegative receptor synthesis rate to be

$$V_R = \bar{V}_n H(-X) + \bar{V}_p H(X) = \bar{V}_p \{ \rho^2 H(-X) + H(X) \},$$

for $T > 0$ with $0 \le \rho^2 = \bar{V}_n / \bar{V}_p \le 1$, unless indicated otherwise. In that case, we have

$$v_R = v_R(x) = \bar{v}_p \{ \rho^2 H(-X) + H(X) \} \equiv \bar{v}_p r_0(x) \qquad (t > 0). \tag{18}$$

where

$$\bar{v}_p = \frac{\bar{V}_p / \bar{R}_0}{D/Z_{\max}^2}, \qquad r_0(x) = \begin{cases} \rho^2 & (x < 0), \\ 1 & (x > 0). \end{cases} \tag{19}$$

In the extreme case $\rho^2 = 0$, there is no receptor synthesized in the morphogen production region. At the other extreme, $\rho^2 = 1$, the receptor synthesis rate is uniform through out the posterior compartment, i.e., for all (x, z) in $\bar{\Omega}$. With the initial receptor concentration taken to be the steady state receptor distribution prior to the onset of morphogen production, $R_0(x) = V_R(X)/k_R = \bar{V}_p r_0(x)/k_R$, we take

$$\bar{R}_0 = \frac{\bar{V}_p}{k_R}, \tag{20}$$

so that

$$\bar{v}_p = g_r, \qquad R_0(x) = \bar{R}_0 r_0(x) = \bar{R}_0 \{ \rho^2 H(-X) + H(X) \}. \tag{21}$$

We are interested in a time-independent steady state solution $\bar{a}(x, z)$, $\bar{b}(x, z)$, and $\bar{r}(x, z)$ for the system (12)–(16). For such a solution, we may set all time derivatives in these equations to zero to get

$$0 = \nabla^2 \bar{a} - h_z \bar{a} \bar{r} + f_z \bar{b} + v_L(x), \tag{22}$$

and

$$0 = h_z \bar{a} \bar{r} - (f_z + g_z) \bar{b}, \qquad 0 = v_R(x) - h_z \bar{a} \bar{r} - g_r \bar{r} + f_z \bar{b}, \tag{23}$$

for $(x, z) \in \Omega$. The nonlinear system of ODE (22)–(23) is augmented by the boundary conditions (14)–(15). With $v_L(x)$ and $v_R(x)$ both piecewise constant, the form of the (22)–(23) requires that $\bar{a}$ and its first derivative to be continuous at $x = 0$.

3.2. _Reduction to a single equation for $\bar{a}(x)$._ The two equations in (23) may be solved for $\bar{b}$ and $\bar{r}$ in terms of $\bar{a}$ to obtain

$$\bar{r} = \frac{\alpha_z r_0(x)}{\alpha_z + \zeta \bar{a}}, \qquad \bar{b} = \frac{r_0(x)\bar{a}}{\alpha_z + \zeta \bar{a}}, \tag{24}$$

where

$$\zeta = \frac{g_z}{g_r} = \frac{k_{\deg}}{k_R}, \qquad \alpha_z = \frac{g_z + f_z}{h_z}. \tag{25}$$

The expressions in (24) are now used to eliminate $\bar{r}$ and $\bar{b}$ from (22) to get a second-order PDE for $\bar{a}$ alone:

$$\nabla^2 \bar{a} - \frac{g_z r_0(x)\bar{a}}{\alpha_z + \zeta \bar{a}} + v_L(x) = 0 \qquad (x, z) \in \Omega. \tag{26}$$

Equation (26) is supplemented by the four boundary conditions (14)–(15) applied to $\bar{a}$, keeping in mind also the continuity conditions on the unknown and its normal derivative at $x = 0$.

For our choice of synthesis rates V_L and V_R, we have $v_L = 0$ and $r_0(x) = 1$ for the range $0 < x < \ell_M$ so that

$$\nabla^2 \bar{a} = \frac{g_z \bar{a}}{\alpha_z + \zeta \bar{a}} = \frac{g_r \bar{a}}{\alpha_r + \bar{a}}, \qquad \alpha_r = \frac{g_r}{g_z} \alpha_z \tag{27}$$

for (x, z) in $\Omega_0 = \{0 < x < \ell_M, \ 0 < z < 1\}$. In the complementary range $\Omega_m = \{-x_m < x < 0, \ 0 < z < 1\}$, we have $v_L = \bar{v}_z$ and $r_0(x) = \rho^2$ so that

$$\nabla^2 \bar{a} - \frac{g_z \rho^2 \bar{a}}{\alpha_z + \zeta \bar{a}} + \bar{v}_z = 0, \tag{28}$$

for (x, z) in Ω_m and for some known ρ^2 in the range $0 \le \rho^2 \le 1$.

3.3. _Existence and uniqueness of steady state behavior._ The governing PDE (26) for the present extra-cellular model is similar to the corresponding ODE investigated in [Lander et al. 2005b]. This observation effectively allows us to extend the results for the one-dimensional model there to show existence and uniqueness for the two-dimensional model of this paper.

Proposition 1. _There is a unique set of nonnegative steady state concentration gradients $\bar{a}(x, z), \bar{b}(x, z)$ and $\bar{r}(x, z)$ characterized by the two-point boundary value problem (14), (15), (26) and the continuity conditions on $\bar{a}$ and $\partial \bar{a}/\partial x$ at $x = 0$._

To prove existence, we observe that $a_\ell(x, z) \equiv 0$ is a _lower solution_ of the BVP for $\bar{a}(x, z)$ [Sattinger 1972] since it satisfies the inequality

$$-\nabla^2[a_\ell] + \frac{g_z r_0(x) a_\ell}{\alpha_z + \zeta a_\ell} - v_L(x) = -v_L(x) = -\bar{v}_z H(-x) \le 0, \qquad (x, z) \in \Omega$$

and the four relevant boundary conditions, the latter exactly. Also,

$$a_u(x, z) = \bar{v}_z \left\{ (\ell_M - x)x_m + \tfrac{1}{2}(\ell_M^2 - x^2) \right\}$$

with

$$(i) \quad a_u(x) \ge 0, \qquad (ii) \quad \frac{\partial a_u}{\partial x}(x) = -\bar{v}_z(x + x_m) < 0 \qquad (-x_m < x \le \ell_M), \quad .$$

is an *upper solution* [Sattinger 1972]. Note that property (ii) ensures $0 \leq a_u(x) \leq a_u(-x_m)$ in $[-x_m, \ell_M]$. With $a_u(x) > 0$ for $-x_m \leq x < \ell_M$, we have

$$-\nabla^2[a_u] + \frac{g_z r_0(x) a_u}{\alpha_z + \zeta a_u} - v_L(x) = \bar{v}_z + \frac{g_z a_u}{\alpha_z + \zeta a_u} - v_L(x) \geq \frac{g_z a_u}{\alpha_z + \zeta a_u} \geq 0 \qquad (x, z) \in \Omega$$

and

$$\left[\frac{\partial a_u}{\partial x}\right]_{x=-x_m} = 0, \qquad a_u(\ell_M) = 0,$$

$$\left[\frac{\partial a_u}{\partial z} + \gamma_z a_u\right]_{z=1} \geq 0, \qquad \left[\frac{\partial a_u}{\partial z}\right]_{z=0} = 0.$$

There exists then a solution $\bar{a}(x, z)$ of the BVP (26), (14) and (15) with

$$0 = a_\ell(x) \leq \bar{a}(x, z) \leq a_u(x),$$

for $(x, z) \in \Omega$ (see [Amann 1972], [Sattinger 1972], and [Smoller 1983]). It follows that $\bar{a}(x, z)$ must be nonnegative in the whole solution domain.

To show that there is only one solution, suppose $a_1(x, z)$ and $a_2(x, z)$ are two (nonnegative) solutions and $a(x, z) = a_1(x, z) - a_2(x, z)$. Then as a consequence of the differential equation (26) for a_1 and a_2, the difference $a(x, z)$ satisfies the differential equation

$$-\nabla^2 a + \frac{g_z \zeta \alpha_z r_0(x) a}{(\alpha_z + \zeta a_1)(\alpha_z + \zeta a_2)} = 0.$$

Form the following double integral of the PDE above over the solution domain to get

$$\int_0^1 \int_{-x_m}^{\ell_M} \left[-\nabla^2 a + \frac{g_z \zeta \alpha_z r_0 a}{(\alpha_z + \zeta a_1)(\alpha_z + \zeta a_2)}\right] a \, dx \, dz = 0. \tag{29}$$

Upon integration by parts (by way of Green's theorem), observing continuity of $\bar{a}$ and its first derivatives, and application of the boundary conditions in (14) and (15), the relation (29) may be transformed into

$$\int_{-x_m}^{\ell_M} \left[\frac{a^2}{\sigma_z}\right]_{z=1} dx + \int_0^1 \int_{-x_m}^{\ell_M} \left\{|\vec{\nabla} a|^2 + \frac{g_z \zeta \alpha_z r_0 a^2}{(\alpha_z + \zeta a_1)(\alpha_z + \zeta a_2)}\right\} dx \, dz = 0. \tag{30}$$

All terms in the integrands in (30) are nonnegative; therefore we must have $a(x) \equiv 0$ and uniqueness is proved. $\qquad \square$

Note that there is no restriction on the magnitude of the (dimensionless) morphogen production rate $\bar{v}_z$ or the (dimensionless) degradation rate g_z for the existence of steady state concentration gradients. Accordingly, Proposition 1 allows us to obtain exact or approximate solution of the BVP by any choice of analytical or numerical methods.

4. Linear stability

4.1. *A nonlinear eigenvalue problem.* In addition to the existence of unique steady state concentrations $\bar{a}(x, z), \bar{b}(x, z),$ and $\bar{r}(x, z)$, it is important for these concentrations to be asymptotically stable (at least

with respect to small perturbations). To investigate the stability of the steady state solution known to exist from Proposition 1, we consider small perturbations in the form

$$\{a, b, r\} = \{\bar{a}(x, z), \bar{b}(x, z), \bar{r}(x, z)\} + e^{-\omega t}\{\hat{a}(x, z), \hat{b}(x, z), \hat{r}(x, z)\}. \tag{31}$$

After linearization, the differential equations (12)–(13) become

$$-\omega\hat{a} = \nabla^2\hat{a} - h_z(\bar{r}\hat{a} + \bar{a}\hat{r}) + f_z\hat{b}, \tag{32}$$

$$-\omega\hat{b} = h_z(\bar{r}\hat{a} + \bar{a}\hat{r}) - (f_z + g_z)\hat{b}, \tag{33}$$

$$-\omega\hat{r} = -h_z(\bar{r}\hat{a} + \bar{a}\hat{r}) - g_r\hat{r} + f_z\hat{b}. \tag{34}$$

The relations (33) and (34) are then solved for $\hat{b}$ and $\hat{r}$ in terms of $\hat{a}$ making use of $\bar{b} = \dfrac{g_r\bar{a}}{g_z(\bar{a} + \alpha_r)}$ to get

$$\hat{r} = \frac{h_z(\omega - g_z)\bar{r}(x, z)\hat{a}}{(g_r - \omega)(f_z + g_z - \omega) + h_z\bar{a}(x, z)(g_z - \omega)}, \tag{35}$$

$$\hat{b} = \frac{h_z(g_r - \omega)\bar{r}(x, z)\hat{a}}{(g_r - \omega)(f_z + g_z - \omega) + h_z\bar{a}(x, z)(g_z - \omega)}. \tag{36}$$

The expressions (36) and (35) are used to eliminate $\hat{b}$ and $\hat{r}$ from (32) to obtain

$$\nabla^2\hat{a} + [\omega - q_r(x; z, \omega)]\,\hat{a} = 0, \tag{37}$$

where

$$q_r(x; z, \omega) = \frac{h_z\bar{r}(x, z)(g_r - \omega)(g_z - \omega)}{(g_r - \omega)(g_z + f_z - \omega) + h_z\bar{a}(x, z)(g_z - \omega)} \tag{38}$$

$$= \frac{1}{1 + \zeta\bar{\beta}_m A} \frac{h_z r_0(x)(g_r - \omega)(g_z - \omega)}{(g_r - \omega)(g_z + f_z - \omega) + (g_z + f_z)(g_z - \omega)\bar{\beta}_m A} \tag{39}$$

$$\equiv \frac{1}{1 + \zeta\bar{\beta}_m A(x, z)} \frac{N_r(x; z, \omega)}{D_r(x; z, \omega)}, \tag{40}$$

where we have set

$$\bar{a}(x, z) = \alpha_z\bar{\beta}_m A(x, z), \tag{41}$$

with $A(-x_m, 0) = 1$ so that $\bar{a}(-x_m, 0) = \alpha_z\bar{\beta}_m$. Note that $\bar{\beta}_m$ is known to be positive from the solution of the steady state problem of the previous section. Let

$$\bar{\beta}_m = \frac{\beta_m}{\rho^2 - \zeta\beta_m} \qquad \text{or equivalently} \qquad \beta_m = \frac{\rho^2\bar{\beta}_m}{1 + \zeta\bar{\beta}_m}; \tag{42}$$

then $\beta_m = \bar{b}(-x_m, 0)$ is positive.

The PDE (37) is supplemented by the boundary conditions (14)–(15) applied to $\hat{a}(x, z)$. Together, (37), (14) and (15) define an eigenvalue problem with ω as the eigenvalue parameter. Though the PDE for $\hat{a}(x, z)$ is linear, the eigenvalue problem is nonlinear since ω appears nonlinearly in $q_r(x; z, \omega)$ so that the homogeneous boundary value problem defined by (37), (14) and (15) is *not* a Sturm–Liouville problem. Given that $r_0(x)$ (and hence also $\bar{r}(x, z)$ and $\bar{b}(x, z)$) may have at most a simple jump discontinuity at $x = 0$, we expect $\hat{a}$ and $\partial\hat{a}/\partial x$ to be continuous at $x = 0$. In the next subsection, we show that the

eigenvalues of the homogeneous boundary value problem must be positive. The steady state gradients $\bar{a}(x, z)$, $\bar{b}(x, z)$, and $\bar{r}(x, z)$ are therefore asymptotically stable by linear stability theory. Since the proof is based on the same technique as that used for one-dimensional models (see [Lander et al. 2005b] for example), we give it below for the simpler case of uniform receptor synthesis rate which can be easily extended to a discontinuous $v_R(x)$ (leading to a discontinuous $r_0(x)$).

4.2. Positive eigenvalues and asymptotic stability. For $\rho^2 = 1$, so that $r_0(x) = 1$ for $(x, z) \in \bar{\Omega}$, the various gradient concentrations are continuous across $x = 0$.

Proposition 2. *All the eigenvalues of the nonlinear eigenvalue problem* (37), (14) *and* (15) *are real.*

To prove this assertion, suppose ω is a complex eigenvalue and $a_\omega(x)$ an associated nontrivial (generally complex) eigenfunction, then ω^* is also an eigenvalue with eigenfunction $a_\omega^*(x)$, where * denotes complex conjugation. The bilinear relation

$$\int_0^1 \int_{-x_m}^{\ell_M} [a_\omega^* \nabla^2 a_\omega - a_\omega \nabla^2 a_\omega^*] \, dx \, dz = 0$$

(which can be established by integration by parts and applications of the boundary conditions in (14) and (15)) requires

$$\int_0^1 \int_{-x_m}^{\ell_M} \left\{ (\omega - \omega^*) - \left[q_r(x; z, \omega) - q_r(x; z, \omega^*) \right] \right\} (a_\omega^* a_\omega) \, dx \, dz = 0. \tag{43}$$

It is straightforward to verify $q_r(x; z, \omega) - q_r(x; z, \omega^*) = -(\omega - \omega^*) \Phi(x; z, \omega\omega^*)$, where

$$\Phi(x, z; \omega\omega^*) = \frac{h_z \{ f_z Q(g_p, \omega) + (g_z + f_z) \bar{\beta}_m A(x, z) Q(g_z, \omega) \}}{(1 + \zeta \bar{\beta}_m A) D_r(x; z, \omega) D_r(x; z, \omega^*)},$$

with $D_r(x; z, \omega)$ as defined in (40) and

$$Q(y, \omega) = [y - \mathrm{Re}(\omega)]^2 + [\mathrm{Im}(\omega)]^2 > 0.$$

In that case, the condition (43) becomes

$$(\omega - \omega^*) \int_0^1 \int_{-x_m}^{\ell_M} a_\omega a_\omega^* [1 + \Phi(x; z, \omega\omega^*)] \, dx \, dz = 0. \tag{44}$$

Since the double integral is positive for any nontrivial function $a_\omega(x, z; \omega)$, we must have $\omega - \omega^* = 0$. Hence, ω does not have an imaginary part.

Proposition 3. *All eigenvalues of the nonlinear eigenvalue problem* (32)–(34), (14) *and* (15) *are positive and the steady state concentrations* $\bar{a}(x, z)$, $\bar{b}(x, z)$ *and* $\bar{r}(x, z)$ *are asymptotically stable by a linear stability analysis.*

If the assertion is false and $\omega \leq 0$, let $\hat{a}_\omega(x)$ be a nontrivial eigenfunction of the homogeneous BVP (37), (14) and (15) for the nonpositive eigenvalue $\omega = -|\omega|$. Multiply (37) by $\hat{a}_\omega$ and integrate over the solution domain to get

$$\int_0^1 \int_{-x_m}^{\ell_M} \left\{ \hat{a}_\omega \nabla^2 \hat{a}_\omega - q_r(x; z, \omega)(\hat{a}_\omega)^2 \right\} \, dx \, dz = -\omega \int_0^1 \int_{-x_m}^{\ell_M} (\hat{a}_\omega)^2 \, dx \, dz.$$

After integration by parts and applications of the homogeneous boundary conditions (14) and (15), we obtain

$$\omega \int_0^1 \int_{-x_m}^{\ell_M} (\hat{a}_\omega)^2 \, dx \, dz = \int_0^1 \int_{-x_m}^{\ell_M} \left| \vec{\nabla} \hat{a}_\omega \right|^2 \, dx \, dz + \int_0^1 \int_{-x_m}^{\ell_M} q_r(x; z, \omega)(\hat{a}_\omega)^2 \, dx \, dz. \qquad (45)$$

With $\omega = -|\omega| \le 0$, we have

$$q_r(x; z, -|\omega|) = \frac{\bar{r}(x, z) h_z (g_z + |\omega|)(g_p + |\omega|)}{(g_r + |\omega|)(g_z + f_z + |\omega|) + h_z \bar{a}(x, z)(g_z + |\omega|)} > 0$$

in Ω. For any nontrivial solution of the eigenvalue problem under the assumption $\omega \le 0$, the right-hand side of (45) is positive, which contradicts the assumption $\omega = -|\omega| \le 0$. Hence the eigenvalues of the eigenvalue problem for $\hat{a}$ must be positive and the proposition is proved.

5. Perturbation solution for $\zeta < 1$ and gradient robustness

For Dpp gradients in *Drosophila* wing disc, k_{deg} is typically smaller than the degradation rate constant k_R of the signaling receptor Tkv so that $\zeta < 1$. For this case, a perturbation solution in ζ is appropriate for moderate Dpp synthesis rate resulting in low receptor occupancy (see [Bender and Orszag 1999]):

$$\bar{a}(x, z; \zeta) = \sum_{k=0}^{\infty} \bar{a}_k(x, z) \zeta^k. \qquad (46)$$

For sufficiently small values of ζ so that $\zeta \bar{a} \ll \alpha_z$, the leading term $\bar{a}_0(x, z)$, determined by the linear PDE

$$\nabla^2 \bar{a}_0 - \mu_z^2 r_0(x) \bar{a}_0 + v_L(x) = 0, \qquad \mu_z^2 = \frac{g_z}{\alpha_z} \qquad (47)$$

and the four boundary condition (14)–(15) applied to $\bar{a}_0$, is an adequate approximation of the exact solution. Here, we have, in terms of the Heaviside unit step function $H(\cdot)$, $r_0(x) = \{H(x) + \rho^2 H(-x)\}$ and $v_L(x) = \bar{v}_z H(-x)$ with $\bar{v}_z = (\bar{V}_L / \bar{R}_0)/(D/Z_{\max}^2)$. The omission of the $\zeta \bar{a}$ term in (26) to get the leading term approximation (47) may also be viewed as a case of *low receptor occupancy* resulting from a sufficiently high receptor synthesis rate (or a sufficiently low morphogen synthesis rate). With plenty of unoccupied receptors available to capture any free Dpp, the normalized morphogen concentration $\bar{a}$ would be sufficiently low for $\zeta \bar{a}$ to be negligible compared to α_z.

The linear BVP (47), (14)–(15) can be solved by Fourier cosine series in the z variable:

$$\{\bar{a}_0(x, z), v_L(x)\} = \sum_{n=1}^{\infty} \{A_n(x), v_n \bar{v}_z H(-x)\} \cos(\lambda_n z) \qquad (48)$$

where $\{\lambda_n\}$ are roots of

$$\cot(\lambda) = \sigma_z \lambda, \qquad \sigma_z = \frac{1}{\gamma_z}, \qquad (49)$$

so that $\bar{a}_0(x, z)$ satisfies the boundary condition at both $z = 0$ and $z = 1$. Orthogonality of the eigenfunctions $\{\cos(\lambda_n z)\}$ requires

$$v_n = \frac{4 \sin(\lambda_n)}{2\lambda_n + \sin(2\lambda_n)}$$

and

$$A_n'' - \Lambda_n^2 A_n + v_n \bar{v}_z H(-x) = 0, \qquad \Lambda_n^2 = \lambda_n^2 + r_0(x)\mu_z^2, \tag{50}$$

where $\mu_z^2 = g_z/\alpha_z$, with

$$A_n'(-x_m) = 0, \qquad A_n(\ell_M) = 0 \tag{51}$$

for $n = 1, 2, 3, \ldots$. When $\rho^2 = 1$, we expect $A_n(x)$ and $A_n'(x)$ to be continuous at $x = 0$.

Remark 1. The leading term perturbation solution in the small parameter ζ is generally an accurate characterization of the actual nonlinear phenomenon. If $\zeta \bar{a}(x, z)$ should not be small compared to α_z, the gradient [LR] for $\rho = 1$ would be nearly uniform in the anteroposterior direction except for a boundary layer near the edge of the wing disc. Such [LR] gradients are not biologically realistic for patterning. We may therefore focus our attention on the low receptor occupancy case (with $\zeta \bar{a}(x, z) \ll \alpha_z$) independent of the magnitude of ζ to investigate the effects of size on the signaling morphogen gradient.

To extract useful information from the Fourier cosine series solution (48)–(51), we deduce below a simple but adequate approximation of the leading term perturbation solution, focusing on the special case $\rho(x) = 1$ so that $r_0(x) = 1$ and give special attention to the limiting case of $\ell_M = \infty$.

5.1. *Finite ℓ_M*. For special case $\rho(x) = 1$, the exact solution for $A_j(x)$ is

$$A_j(x) = \begin{cases} \dfrac{\bar{v}_z v_j}{\Lambda_j^2}\left(1 - \dfrac{\cosh(\Lambda_j \ell_M)}{\cosh(\Lambda_j(\ell_M + x_m))}\cosh(\Lambda_j(x_M + x))\right) & (-x_m \leq x \leq 0), \\[3mm] \dfrac{\bar{v}_z v_j \sinh(\Lambda_j x_m)}{\Lambda_j^2 \cosh(\Lambda_j(\ell_M + x_m))}\sinh(\Lambda_j(\ell_M - x)) & (0 \leq x \leq \ell_M). \end{cases} \tag{52}$$

Correspondingly, we have

$$\bar{a}(0, z) \sim \bar{a}_0(0, z) = \bar{v}_z \sum_{j=1}^{\infty} \frac{v_j}{\Lambda_j^2}\frac{\sinh(\Lambda_j x_m)\sinh(\Lambda_j \ell_M)}{\cosh(\Lambda_j(\ell_M + x_m))}\cos(\lambda_j z),$$

$$\bar{a}(-x_m, z) \sim \bar{a}_0(-x_m, z) = \bar{v}_z \sum_{j=1}^{\infty} \frac{v_j}{\Lambda_j^2}\left(1 - \frac{\cosh(\Lambda_j \ell_M)}{\cosh(\Lambda_j(\ell_M + x_m))}\right)\cos(\lambda_j z).$$

With $(v_j/\Lambda_j^2)/(v_1/\Lambda_1^2) \ll 1$ for $j > 1$, we have as a leading term approximation

$$b(x, z) \sim \frac{1}{\alpha_z}\bar{a}(x, z) \sim \frac{1}{\alpha_z}\bar{a}_0(x, z) \approx \frac{1}{\alpha_z}A_1(x)\cos(\lambda_1 z)$$

$$= \frac{\bar{v}_z v_1}{\alpha_z \Lambda_1^2}\frac{\sinh(\Lambda_1 x_m)}{\cosh(\Lambda_1(\ell_M + x_m))}\sinh(\Lambda_1(\ell_M - x))\cos(\lambda_1 z). \tag{53}$$

5.2. *The limiting case of $\ell_M = \infty$*. For the wing imaginal disc of *Drosophila* species, $X_{max} \gg Z_{max}$ so $\ell_M \gg 1$. It often suffices for our purpose to consider the limiting case of $\ell_M = \infty$ (and $\rho = 1$) for which

$$A_j(x) = \begin{cases} \dfrac{\bar{v}_z v_j}{\Lambda_j^2}\{1 - e^{-\Lambda_j x_m}\cosh(\Lambda_j(x + x_m))\} & (-x_m \leq x \leq 0), \\[3mm] \dfrac{\bar{v}_z v_j}{\Lambda_j^2}\sinh(\Lambda_j x_m)e^{-\Lambda_j(x + x_m)} & (0 \leq x < \infty). \end{cases} \tag{54}$$

with

$$\bar{a}(0, z) \sim \bar{a}_0(0, z) = \bar{v}_z \sum_{j=1}^{\infty} \frac{v_j}{\Lambda_j^2} e^{-\Lambda_j x_m} \sinh(\Lambda_j x_m) \cos(\lambda_j z), \tag{55}$$

$$\bar{a}(-x_m, z) \sim \bar{a}_0(-x_m, z) = \bar{v}_z \sum_{j=1}^{\infty} \frac{v_j}{\Lambda_j^2} \left(1 - e^{-\Lambda_j x_m}\right) \cos(\lambda_j z). \tag{56}$$

To a leading term approximation, we have for $\ell_M = \infty$

$$b(x, z) \sim \frac{1}{\alpha_z} \bar{a}(x, z) \sim \frac{1}{\alpha_z} \bar{a}_0(x, z) \approx \frac{1}{\alpha_z} A_1(x) \cos(\lambda_1 z)$$

$$= \frac{\bar{v}_z v_1}{\alpha_z \Lambda_1^2} \sinh(\Lambda_1 x_m) e^{-\Lambda_1(x+x_m)} \cos(\lambda_j z) \quad (0 \le x < \infty). \tag{57}$$

For an [LR] gradient to be biologically useful in developing tissue patterns, its graph must be neither nearly uniform nor a boundary layer phenomenon near the source. It follows that the concentration of signaling Dpp–Tkv complexes [LR] should be in a state of low receptor occupancy throughout the wing disc with both free and bound Dpp approximately in a state of simple exponential decay from the source end to the sink at the edge of wing disc. With the free Dpp expression exponentially small away from the source, the actual location of the absorbing edge should not have a significant effect on the signaling Dpp gradient and may be taken to be far away at infinity.

6. Approximate expressions for the eigenvalues $\{\lambda_n\}$

6.1. *A sealed basement membrane* ($\gamma_z = 0$). For the limiting case of $\gamma_z = 0$, both apical and basal faces are sealed (see (15)) given that

$$\left[\frac{\partial \bar{a}(x, z)}{\partial z}\right]_{z=1} = \left[\frac{\partial \bar{a}(x, z)}{\partial z}\right]_{z=0} = 0. \tag{58}$$

Upon writing the equation (49) for the eigenvalues as

$$\gamma_z \cos(\lambda) = \lambda \sin(\lambda),$$

we have for $\gamma_z = 0$

$$\lambda_n = (n-1)\pi \qquad (n = 1, 2, \dots). \tag{59}$$

It follows that

$$v_1 = 1, \qquad v_n = 0 \quad (n \ge 2).$$

The concentration gradient is therefore a uniform distribution in z. In that case, the solution for the (normalized) free morphogen concentration $\bar{a}$ is uniform in z so that

$$\bar{a}(x, z) \sim \bar{a}_0(x)\{1 + O(\zeta)\}$$

and $\bar{a}_0(x)$ is just the corresponding solution for the spatially one-dimensional problem previously treated in [Lander et al. 2005b].

For the case $\ell_M = \infty$ (and $\rho = 1$), we have from (57)

$$\frac{[LR(x,z)]}{\bar{R}_0} \sim \frac{\bar{a}_0(x)}{\alpha_z} = \begin{cases} (\bar{v}_z/g_z)\{1 - e^{-\mu_z x_m}\cosh(\mu_z(x+x_m))\} & (-x_m \le x \le 0), \\ (\bar{v}_z/g_z)\sinh(\mu_z x_m)e^{-\mu_z(x+x_m)} & (0 \le x < \infty). \end{cases} \tag{60}$$

with

$$\mu_z^2 = \frac{g_z}{\alpha_z} \simeq h_z, \quad \bar{v}_z = \frac{\bar{V}_L/\bar{R}_0}{D/Z_{\max}^2}, \quad \frac{\bar{v}_z}{g_z} = \frac{\bar{V}_L/\bar{R}_0}{k_{\deg}}. \tag{61}$$

With $h_z = O(10^{-1})$ for typical wing disc parameter values, we have $\mu_z x_m \ll 1$ and therewith

$$e^{-\mu_z x_m}\sinh(\mu_z x_m) \simeq \mu_z x_m\{1 + O(\mu_z^2 x_m^2)\},$$

so

$$[LR(x,z)] \sim \frac{\bar{R}_0}{\alpha_z}\bar{a}_0(x) \approx \frac{\bar{V}_L}{k_{\deg}}\mu_z x_m e^{-\mu_z x} = \frac{\bar{V}_L X_{\min}}{k_{\deg}}\bar{k}_{\mathrm{on}}e^{-\bar{k}_{\mathrm{on}}X} \tag{62}$$

for the signaling region $0 \le x < \infty$, with $\bar{k}_{\mathrm{on}} = \sqrt{k_{\mathrm{on}}\bar{R}_0/D}$.

We know from biological evidence that the basal end is not sealed, so that the limiting case of $\gamma_z = 0$ only demonstrates the validity and consistency of our more general solution, but is otherwise biologically unrealistic and of no relevance to the actual problem. The analytical consequences of a sealed end, as seen from (62), are shown in the next section to be also unacceptable from the view point of size adjustment for robustness with respect to a substantial change in the Dpp synthesis rate.

6.2. *An absorbing basement membrane* ($\sigma_z = 0$).

At the other extreme, Equation (49) in the case of $\sigma_z = 1/\gamma_z = 0$ becomes $\cot(\lambda) = 0$, so that

$$\lambda_n = (n - \tfrac{1}{2})\pi \qquad (n = 1, 2, 3, \dots). \tag{63}$$

It follows that, for the limiting case of $\ell_M = \infty$ and $\rho = 1$,

$$A_j(x) = \begin{cases} (\bar{v}_z v_j/\Lambda_j^2)\{1 - e^{-\Lambda_j x_m}\cosh(\Lambda_j(x+x_m))\} & (-x_m < x < 0), \\ (\bar{v}_z v_j/\Lambda_j^2)\sinh(\Lambda_j x_m)e^{-\Lambda_j(x+x_m)} & (0 < x < \infty). \end{cases} \tag{64}$$

where

$$\Lambda_n^2 = \lambda_n^2 + \mu_z^2 = (n - \tfrac{1}{2})^2\pi^2 + \mu_z^2, \quad v_n = (-)^{n-1}\frac{2}{\lambda_n}, \quad (n = 1, 2, 3, \dots).$$

Correspondingly, we have for the signaling region $0 < x < \infty$

$$\frac{[L]}{R_0} \sim A_1(x)\cos(\lambda_1 z) = \frac{\bar{v}_z v_1}{\Lambda_1^2}\sinh(\Lambda_1 x_m)e^{-\Lambda_1(x+x_m)}\cos(\lambda_1 z), \tag{65}$$

$$\frac{[LR]}{R_0} \sim \frac{A_1(x)}{\alpha_z}\cos(\lambda_1 z) = \frac{\bar{v}_z v_1}{\alpha_z \Lambda_1^2}\sinh(\Lambda_1 x_m)e^{-\Lambda_1(x+x_m)}\cos(\lambda_1 z). \tag{66}$$

With $f_z \ll g_z$, the expression for μ_z^2 is accurately approximated by $h_z = k_{\mathrm{on}}\bar{R}_0/(D/Z_{\max}^2) = O(10^{-1})$. Since μ_z^2 is small compared to $\pi^2/4$ (and much smaller than $(n - \tfrac{1}{2})^2\pi^2$ for $n \ge 2$) and $\lambda_1 x_m \ll 1$, we have the following accurate approximation for the various Fourier components of the Dpp concentration:

$$\Lambda_n^2 = \lambda_n^2 + \mu_z^2 \simeq \lambda_n^2 = (n - \tfrac{1}{2})^2 \pi^2, \qquad \sinh(\lambda_1 x_m) e^{-\lambda_1 x_m} \approx \lambda_1 x_m, \tag{67}$$

$$A_j(x) \simeq \frac{\bar{v}_z v_j}{\lambda_j^2} \sinh(\lambda_j x_m) e^{-\lambda_j (x_m + x)} \qquad (0 < x < \infty) \tag{68}$$

so that

$$[L] \sim \bar{R}_0 \bar{a}_0(x, z) \approx \bar{R}_0 A_1(x) \cos(\lambda_1 z) \simeq \frac{\bar{R}_0 \bar{v}_z v_1 x_m}{\lambda_1} e^{-\lambda_1 x} \cos(\lambda_1 z), \tag{69}$$

$$[LR] \sim \frac{\bar{R}_0}{\alpha_z} \bar{a}_0(x, z) \simeq \frac{\bar{R}_0 \bar{v}_z v_1 x_m}{\alpha_z \lambda_1} e^{-\lambda_1 x} \cos(\lambda_1 z), \tag{70}$$

for $0 < x < \infty$. Upon observing (67) and setting $\Lambda_n^2 \approx \lambda_n^2$, (69) and (70) become

$$[L] \sim \frac{\bar{V}_L X_{\min} Z_{\max}}{D} \left[\frac{8}{\pi^2} e^{-\pi x/2} \cos(\tfrac{1}{2} \pi z) \right], \tag{71}$$

$$[LR] \sim \frac{\bar{V}_L X_{\min} Z_{\max}}{D} \left[\frac{8}{\alpha_z \pi^2} e^{-\pi x/2} \cos(\tfrac{1}{2} \pi z) \right], \tag{72}$$

for $0 < x < \infty$. The signaling Dpp gradient given in (70) and (72) is qualitatively different from that in (62) and (72). While the limiting case of $\sigma_z = 0$ is also not biologically realistic as the basement membrane is neither completely sealed nor absorbing [Dowd et al. 1999], the consequences of (70) are more characteristic of the $[LR]$ gradient than those of (62) in the actual range of γ_z as we shall see in the next subsection.

It may seem rather remarkable that (the leading term asymptotic behavior of) $[LR]$ does not depend on the receptor synthesis rate. However the inherent assumption of low receptor occupancy in effect corresponds to (an abundance of receptors resulting from) a sufficiently high receptor synthesis rate to make its magnitude inconsequential in a first approximation theory.

6.3. *Eigenvalues for small and large σ_z.* We now turn to the biologically more realistic case of a leaky basement membrane at $Z = Z_{\max}$ with a finite $\gamma_z = 1/\sigma_z$. For the wing disc problem, we have $0 < \sigma_z < 1$ and a perturbation solution for λ in a power series in σ_z gives

$$\lambda_n = \left(n - \tfrac{1}{2}\right) \pi \left[1 - \sigma_z + O(\sigma_z^2) \right], \tag{73}$$

for $n = 1, 2, 3, \ldots$. No new eigenpairs arise from nonpositive integers n. We conclude (as in the limiting case of an absorbing basal end):

Proposition 4. *For $0 < \sigma_z = 1/\gamma_z < 1$, the expression (70) is an accurate leading term approximation for the signaling Dpp gradient outside the Dpp production region where the $[LR]$ gradient is instrumental for the wing disc development. Thus, with a permeable (leaky) basement membrane with $0 < \sigma_z < 1$, the slope and convexity (but not the magnitude) of the signaling Dpp gradient in both X and Z direction depends only on the parameter σ_z for the biologically realistic case of low receptor occupancy.*

It appears that a leaky basement membrane serves the purpose of regulating the availability of Dpp at a level that maintain the signaling Dpp gradient *shape*. That is, the *slope* and *convexity* of the $[LR]$ concentration gradient are not sensitive to the synthesis rate $\bar{V}_L$ as it changes with significant environmental perturbations. On the other hand, the magnitude of the bound morphogen concentration is seen from (72) to be proportional to $\bar{V}_L$ (as well as the two length quantities $X_{\min}$ and $Z_{\max}$ and inversely

proportional to the diffusion coefficient D and the composite parameter $\alpha_z = (k_{\text{off}} + k_{\text{deg}})/\bar{R}_0 k_{\text{on}}$ of the binding, dissociation and degradation rate constants).

For completeness, consider also the low leakage case of $\sigma_z \gg 1$ or $0 < \gamma_z \ll 1$. For this case, we may seek a singular perturbation solution for λ in γ_z to obtain

$$\lambda_1 \sim \sqrt{\gamma_z}\left[1 - \tfrac{1}{6}\gamma_z + O(\gamma_z^2)\right], \qquad \lambda_{k+1} \sim k\pi\left[1 + \frac{1}{(k\pi)^2}\gamma_z + O(\gamma_z^2)\right] \tag{74}$$

for $k = 1, 2, \ldots$, with no new eigenpair arising from negative square roots and negative k. The results reduce to those of the limiting case of $\gamma_z = 0$ given in (59) with leading term approximation for $[LR(x, z)]$ for $\ell_M = \infty$ and $\rho = 1$ as previously given in (62). For this (biologically unrealistic) low leakage case, the signaling gradient *shape* is much more sensitive to the leakage parameter γ_z and the degradation-to-binding rate ratio (but not the Dpp synthesis rate $\bar{V}_L$) since we have now

$$\Lambda_1^2 = \lambda_1^2 + \mu_z^2 \sim \gamma_z + \mu_z^2.$$

7. Signaling gradient and wing disc size change

The need to accommodate exogenous environmental changes is a major performance objective of morphogen systems. Given a lack of information on the change in system characteristics caused by such changes, we pursue in this and the next section a more limited objective by investigating the change of signaling morphogen gradient in response to a doubling of Dpp synthesis rate (whatever the cause may be) while all system characteristics remain unchanged. We determine here whether a specific aspect of system architecture, namely, the wing disc size, offers the potential of meeting the performance objective by maintaining the $[LR]$ concentration magnitude and its gradient shape (and thereby preserving tissue pattern) relative to the new size. The concept of robustness is quantified in the next section (see also [Lander et al. 2005c; Vargas 2007]) and used to analyze the sensitivity of the signaling Dpp gradient to the morphogen synthesis rate change and how wing disc size changes may ameliorate this sensitivity. In this section, robustness is taken informally to mean no biologically or functionally significant change in the magnitude and shape of the signaling Dpp gradient $[LR]$ under a significant change in Dpp synthesis rate. (More quantitative measures of robustness will be discussed in the next section.)

We begin by focusing on the case $\ell_M = \infty$ and recalling the following four observations from the results of the previous sections:

1. The apical-basal height of wing disc cells of *Drosophila*s is considerably larger than the lineal dimension of its cross section in the plane of the wing disc. Typically, we have of $x_m = X_{\min}/Z_{\max} = O(10^{-1}) \ll 1$ or smaller so that $0 < \lambda_1 x_m \ll 1$, resulting in $\sinh(\lambda_1 x_m)e^{-\lambda_1 x_m} \approx \lambda_1 x_m$.

2. The basal membrane is in reality neither sealed nor completely absorbing; instead there is a significant amount of leakage with $\gamma_z = 1/\sigma_z > 1$. Consequently, the approximate expression (73) for the eigenvalues $\{\lambda_j\}$ and Proposition 4 for the signaling gradient apply.

3. With μ_z^2 accurately given by $h_z = k_{\text{on}}\bar{R}_0/(D/Z_{\max}^2) = O(10^{-1}) \ll 1 < \pi^2/4$ for a wing imaginal disc with a relatively high leakage through the basal membrane, it follows from (67) that

$$\Lambda_1^2 = \lambda_1^2 + \mu_z^2 \simeq \lambda_1^2 = O\left(\frac{\pi^2}{4}\right) \tag{75}$$

for $\gamma_z = 1/\sigma_z > 1$ so that $\lambda_1 x_m \ll 1$ and $\sinh(\Lambda_1 x_m)e^{-\Lambda_1 x_m} \simeq \lambda_1 x_m$.

4. Anticipating the need to reduce wing disc span ($X_{\max}$ and $X_{\min}$) and cell height ($Z_{\max}$) to maintain robustness, we note that the approximation $\Lambda_1^2 \approx \lambda_1^2$ for the relevant range $\gamma_z = 1/\sigma_z > 1$ improves with smaller wing disc size. It follows from the fact that a reduction in $Z_{\max}$ further reduces the effect of μ_z^2 (which was already minimal according to Proposition 4). This allows us to make the same approximation for smaller $Z_{\max}$ (or $X_{\min}$ and $X_{\max}$). For $\tilde{Z}_{\max} = Z_{\max}/\sqrt{2}$, we have

$$\frac{\tilde{\Lambda}_1}{\Lambda_1} = \frac{\sqrt{1 + \tilde{\mu}_z^2/\lambda_1^2}}{\sqrt{1 + \mu_z^2/\lambda_1^2}} = \frac{\sqrt{1 + \mu_z^2/2\lambda_1^2}}{\sqrt{1 + \mu_z^2/\lambda_1^2}} \approx 1 - \frac{\mu_z^2}{4\lambda_1^2} \tag{76}$$

so that the approximation $\tilde{\Lambda}_1^2 \approx \Lambda_1^2$ is accurate to within than 4%.

Given that the higher harmonic terms in the eigenfunction expansion for $\bar{a}_0(x, z)$ are negligibly small compared to the leading term, the four observations above lead to the following accurate approximate expression for the signaling gradient

$$[LR] \approx L_{R0}\frac{v_1}{\alpha_z\lambda_1}e^{-\lambda_1 x}\cos(\lambda_1 z), \qquad L_{R0} = \frac{\bar{V}_L X_{\min} Z_{\max}}{D} \tag{77}$$

in the range $0 < x < \infty$ as was found earlier in (70). Averaging $[LR]$ over the interval $[0, Z_{\max}]$ gives

$$[\overline{LR}] \equiv \frac{1}{Z_{\max}}\int_0^{Z_{\max}}[LR]dZ = \int_0^1[LR]dz \approx L_{R0}\frac{v_1}{\alpha_z\lambda_1}e^{-\lambda_1 x}\frac{\sin(\lambda_1)}{\lambda_1} \tag{78}$$

with $\lambda_1 \approx \pi(1 - \sigma_z)/2$ for $0 \leq \sigma_z < \frac{1}{2}$.

We now arrive at the key development of this paper. Suppose the *Dpp* synthesis rate is doubled from $\bar{V}_L$ to $\tilde{V}_L = 2\bar{V}_L$. The maximum magnitude of the signal morphogen gradient (which, for the signaling range $[0, \infty)$, occurs at $x = 0$) would then be doubled since the signaling gradient is proportional to $\bar{V}_L$. The factor $(v_1/\alpha_z\lambda_1)e^{-\lambda_1 x}$ in (77) does not depend on $X_{\min}$ and $Z_{\max}$ explicitly, so the magnitude of $[LR]_{z=0}$ (at least the leading term approximation) may be brought back down to the same level prior to the synthesis rate doubling by reducing either $X_{\min}$ or $Z_{\max}$ by half. Either change would lead to a significant distortion of the developed wing disc. There is however the biologically more realistic alternative of reducing both size parameters by a factor of $\sqrt{2}$. Given $\tilde{X}_{\min} = X_{\min}/\sqrt{2}$ and $\tilde{Z}_{\max} = Z_{\max}/\sqrt{2}$ along with observation (4) above, the new signaling gradient with the modified parameters of this option becomes

$$[\widetilde{LR}]_{z=0} \approx \tilde{L}_{R0}\frac{v_1}{\alpha_z\lambda_1}e^{-\lambda_1 x} = L_{R0}\frac{v_1}{\alpha_z\lambda_1}e^{-\lambda_1 x} \qquad (0 < x < \infty), \tag{79}$$

with $\tilde{L}_{R0} = \tilde{V}_L\tilde{X}_{\min}\tilde{Z}_{\max}/D = \bar{V}_L X_{\min} Z_{\max}/D = L_{R0}$. The right-hand side is just the expression for $[LR]_{z=0}$ in (77). Since the signaling gradient $[LR]$ remains a decaying exponential with the same magnitude at the same scaled proximal-distal location, the *size-normalized* signaling gradient is identical to the corresponding gradient prior to morphogen synthesis rate doubling (though the wing disc size has been reduced). We refer to such preservation of signal gradient shape as *size-normalized robustness* in subsequent developments. Note that the apical-basal average of $[LR]$ as given in (78) is clearly also size–normalized robust.

Among the three options for maintaining the signaling gradient shape after Dpp synthesis rate doubling, halving the anteroposterior span (and hence $X_{\min}$ and $X_{\max}$) alone would mean a more drastic reduction

in the wing disc span in the distal-proximal direction. The resulting new wing disc would be skewed in one direction. At the other extreme, halving $Z_{\max}$ alone would change the shape of the gradient for $x > 0$, given that the exponential factor $e^{-\lambda_1 x}$ becomes $e^{-\lambda_1 X/\tilde{Z}_{\max}} = e^{-\lambda_1(X/Z_{\max})/2}$. Neither is consistent with the often observed consequence of a substantial increase in ambient temperature: a smaller wing disc that is essentially similar to the normal wing imaginal disc in tissue patterning [de Moed et al. 1997; French et al. 1998; Bitner-Mathé and Klaczko 1999; Azevedo et al. 2002]. Hence, we are led to the following proposition (for $\ell_M = \infty$):

Proposition 5. *When there is a significant increase in the Dpp synthesis rate by a factor M with $\tilde{V}_L = M\bar{V}_L$ and the approximation (77) for [LR] holds, the Drosophila wing imaginal disc patterning is preserved by the size reduction $\tilde{X}_{\min} = X_{\min}/\sqrt{M}$ and $\tilde{Z}_{\max} = Z_{\max}/\sqrt{M}$.*

The reduction of $X_{\min}$ and $Z_{\max}$ (as well as $X_{\max}$ in the case of a finite ℓ_M) by the same amplification factor $1/\sqrt{M}$ is known as *self-similar size reduction*. With such a size reduction, the development of the wing disc is insensitive to an M fold increase in Dpp synthesis rate and is therefore *size-normalized robustness*.

It is worth digressing to comment briefly on the biologically unrealistic case $\gamma_z = 0$. Had the basal end been sealed so that (62) applies with

$$[LR(x, z)] \sim \frac{\bar{V}_L X_{\min}}{k_{\deg}} \bar{k}_{\mathrm{on}} e^{-\bar{k}_{\mathrm{on}} X}, \tag{80}$$

the only way to maintain the concentration magnitude $[LR(0, z)]$ by a size change when $\bar{V}_L$ is doubled would be to reduce $X_{\min}$ by half (instead of by $\sqrt{2}$ as in the case of a leaking basal membrane). The size reduction would not be self-similar unless $Z_{\max}$ is also reduced by half. In this latter case, the size reduction would be self-similar size but substantially more drastic than those observed. Thus a porous basement membrane appears to serve an important function in the robust development of the wing disc.

For the case of a large but finite $X_{\max}$, the expression (53) simplifies to

$$\begin{aligned} [LR(x, z)] &\sim \frac{\bar{R}_0}{\alpha_z} \bar{a}(x, z) \sim \frac{\bar{R}_0}{\alpha_z} \bar{a}_0(x, z) \approx \frac{\bar{R}_0}{\alpha_z} A_1(x) \cos(\lambda_1 z) \\ &\simeq L_{R0} \frac{v_1 e^{-\lambda_1 x}}{\alpha_z \lambda_1} \left(1 - e^{-2\lambda_1(\ell_M - x)}\right) \cos(\lambda_1 z), \end{aligned} \tag{81}$$

for $\Lambda_1 x_m \simeq \lambda_1 x_m \ll 1$ (keeping in mind that $L_{R0} = \bar{V}_L X_{\min} Z_{\max}/D$). We have from this the following extension of Proposition 5:

Proposition 6. *For a finite (dimensionless) anteroposterior span ℓ_M, the signaling Dpp gradient is (size-normalized) robust after a Dpp synthesis rate doubling provided that $X_{\max}$ is also reduced by the same factor $\sqrt{2}$ (as $X_{\min}$ and $Z_{\max}$).*

We emphasize that there are at least three advantages in changing both $X_{\min}$ and $Z_{\max}$ (as well as $X_{\max}$) by the same factor to maintain the shape (that is, slope and convexity) of the signaling gradient $[LR]$ when the Dpp synthesis rate is doubled:

1. The size of the wing disc is reduced in all dimensions but the physical shape remains geometrically similar before and after the reduction.

2. The reduction is less drastic, by only a factor of $\sqrt{2}$ instead of half.

3. Most importantly, it is consistent with the observation that fly wings are smaller physically in all directions under the higher temperature, not just in the direction of the anteroposterior axis.

8. Robustness of signaling gradients

We saw from the eigenfunction expansion for the leading term perturbation solution that the signaling [*LR*] gradient is generally sensitive to system parameter changes. Yet actual biological systems are generally robust to such changes (up to a self-similar size change [de Moed et al. 1997; French et al. 1998; Bitner-Mathé and Klaczko 1999; Azevedo et al. 2002]). The analysis of the previous section showed that the magnitude and shape of the signaling gradient [*LR*] can be maintained even when the Dpp synthesis rate increases substantially by changing the anteroposterior span (including the width of the localized Dpp synthesis region) and apical-basal cell height of the wing disc. In this section we quantify this observation by way of either of the two robustness indices to be introduced below and generalize the mechanism of self-similar size reduction for more flexible applications. To be concrete, we focus on robustness with respect to a two-fold change in the Dpp synthesis rate in our model problem as in [Lander et al. 2005c]. The general methodology developed for this parameter change is extended to allow for multifactor changes and robustness with respect to other parameter changes in Section 9.

With a doubling of the morphogen synthesis rate $\tilde{V}_L = 2\bar{V}_L$, we have chosen in the previous section an "amplification" (or a "diminution") factor κ (with $\tilde{X}_{\min} = \kappa X_{\min}$, $\tilde{Z}_{\max} = \kappa Z_{\max}$ as well as $\tilde{X}_{\max} = \kappa X_{\max}$) to be $1/\sqrt{M} = 1/\sqrt{2}$ to maintain the order of magnitude of $A_1(0)$. However, the shape of the new signaling gradient is preserved by such a change only after the approximation $\tilde{\Lambda}_n^2 \approx \Lambda_n^2$. For $\gamma_z = 2$ however, the error incurred for the dominant $n = 1$ term in the eigenfunction expansion is about 15%. We explore in this section how we may limit the effects on Λ_1^2 (and more generally Λ_n^2) resulting from more general size changes of $Z_{\max}$ and $X_{\min}$ when $\gamma_z = O(1)$ (but < 2).

For the particular case of Dpp synthesis rate doubling, there is nothing canonical about reducing $Z_{\max}$ and $X_{\min}$ by a factor $\sqrt{2}$ to minimize the change in [*LR*(0, *z*)]. If we should reduce $Z_{\max}$ (as well as $X_{\min}$ and $X_{\max}$) by a different factor κ instead,

$$\{\tilde{X}_{\min}, \tilde{Z}_{\max}, \tilde{X}_{\max}\} = \kappa \{X_{\min}, Z_{\max}, X_{\max}\}, \tag{82}$$

(e.g., $\kappa = 1/\sqrt{3}$), the new concentration parameter $\tilde{L}_{R0}$ would generally not be the same as L_{R0} as desired, smaller if $\kappa^2 < \frac{1}{2}$ (e.g., $\tilde{L}_{R0} = 2L_{R0}/3$ when $\kappa = 1/\sqrt{3}$) and greater if $\kappa^2 > \frac{1}{2}$. While neither appears ideal for maintaining the magnitude of [*LR*] at $x = 0$ after synthesis rate doubling, the added flexibility may offer alternative benefits for $x > 0$ given that (75) is only an approximate relation and (73) only holds for the biologically realistic range $\sigma_z \ll 1$. This flexibility is explored in the next two subsections. For those developments, it should be noted that the ratio (76) describing the change in the gradient shape factor Λ_1 is modified to become

$$\frac{\tilde{\Lambda}_1}{\Lambda_1} = \frac{\sqrt{1 + \kappa^2 \mu_z^2 / \lambda_1^2}}{\sqrt{1 + \mu_z^2 / \lambda_1^2}} \approx 1 + (\kappa^2 - 1)\frac{\mu_z^2}{2\lambda_1^2}, \tag{83}$$

which remains nearly 1 for a basal membrane with high leakage. The possibility of specifying κ enables us to choose it to minimize the effect of size changes on two robustness indices that measures the relative change in the magnitude and shape of the signaling $[LR]$ gradient downstream.

8.1. *Root-mean-square differential signaling.* To quantify the advantage of a smaller κ (than $1/\sqrt{2}$), we let $\bar{b}(x, z)$ and $\tilde{b}(x, z)$ be the normalized signaling morphogen-receptor gradients for morphogen synthesis rate $\bar{V}_L$ and $\tilde{V}_L = 2\bar{V}_L$, respectively. We have, for the low receptor occupancy case,

$$\{\bar{b}(x, z),\ \tilde{b}(x, z)\} \sim \frac{1}{\alpha_z} \{\bar{a}_0(x, z),\ \tilde{a}_0(x, z)\} \approx \frac{\cos(\lambda_1 z)}{\alpha_z}\{A_1(x),\ \tilde{A}_1(x)\}$$

where $A_1(x)$ (for $\rho = 1$ and $\ell_M = \infty$) is as in (54) and $\tilde{A}_1(x, z)$ is $A_1(x)$ with $\bar{v}_z = (\bar{V}_L/\bar{R}_0)/(D/Z_{\max}^2)$ replaced by $\tilde{v}_z = (\tilde{V}_L/\bar{R}_0)/(D/\tilde{Z}_{\max}^2)$, for $\tilde{Z}_{\max} = \kappa Z_{\max}$ (and $\{\tilde{X}_{\min}, \tilde{X}_{\max}\} = \kappa\{X_{\min}, X_{\max}\}$). Consider the following measure of deviation from the signaling Dpp concentration $[LR]$, namely, the *signal robustness index* R_b defined by

$$R_b(\kappa) = \frac{(\Delta b)_{\text{rms}}}{b_h - b_\ell} = \frac{1}{b_h - b_\ell}\sqrt{\frac{1}{x_\ell - x_h}\int_{x_h}^{x_\ell}[\tilde{b}(x, 0) - \bar{b}(x, 0)]^2 dx} \tag{84}$$

where $0 < b_\ell = \bar{b}(x_\ell, 0) < b_h = \bar{b}(x_h, 0) \le \bar{b}(0, 0)$, with $0 \le x_h < x_\ell \le \ell_M$. It is just the root-mean-square deviation from $\bar{b}(x, 0)$, the normalized signaling gradient $[LR]/\bar{R}_0$ along the apical face over a relevant span of the wing disc. The quantities x_ℓ and x_h (or b_ℓ and b_h) may be chosen away from the extremities to minimize the effects of outliers. (A more encompassing measure would be to use the average value of $\tilde{b}(x, z)$ and $\bar{b}(x, z)$ over the interval $0 < z < 1$ instead of their respective value at $z = 0$. However, it suffices to use the simpler expression (84) in a proof of concept discussion.)

With the doubling of the morphogen synthesis rate from $\bar{V}_L$ to $\tilde{V}_L = 2\bar{V}_L$, we reduce both $X_{\min}$ and $Z_{\max}$ (as well as $X_{\max}$) by an amplification factor κ to make $R_b(\kappa)$ as small as possible to minimize the sensitivity of the signaling gradient to the synthesis rate doubling (and size changes). Alternatively, we may choose instead a value for κ so that $R_b(\kappa)$ is in an acceptable range of values, say $R_b(\kappa) < 0.1$, in order for the root mean square deviation to be effectively less than 10% of the gradient prior to the Dpp synthesis rate doubling.

For the low receptor occupancy case, an explicit expression can be obtained for R_b. For $x_h = 0$ and $x_\ell = 2$ (for which $e^{-2\tilde{\Lambda}_1 x_\ell}$ is negligibly small), we have

$$R_b(\kappa) \simeq \frac{1}{\sqrt{10}}\left(\frac{4\kappa^4}{\tilde{\Lambda}_1^3} + \frac{1}{\Lambda_1^3} - \frac{8\kappa^2}{\Lambda_1\tilde{\Lambda}_1(\tilde{\Lambda}_1 + \Lambda_1)}\right)^{1/2}. \tag{85}$$

The minimum value for this R_b expression is attained at

$$\kappa_{\min}^2 = \frac{\xi^2}{1 + \xi}, \qquad \xi = \frac{\tilde{\Lambda}_1}{\Lambda_1}. \tag{86}$$

Table 1 on the next page shows some results for $\kappa_{\min}$ and several other values of κ for $x_\ell = 2$, $x_h = 0$ and the typical set of system parameter values used in [Lander et al. 2005c] (with the dimensionless parameter values $g_z = 0.002$, $f_z = 0.00001$, $h_z = 0.1$, $g_r = 0.01$ and $x_m = X_{\min}/Z_{\max} = 0.1$). For the three values of the leakage parameters used, $\sigma_z = 0.2$, 0.5, and 1.0, the corresponding first eigenvalue being $\lambda_1 \simeq 1.31$,

$1/\kappa^2$	$\sigma_z = 1/\gamma_z$		
	1.0	0.5	0.2
1	0.5223	0.4720	0.4301
2	0.0258	0.0153	0.0095
2.025	0.0198	0.0099	0.0048
2.05	0.0144	0.0057	0.0034
2.10	0.0084	0.0094	0.0119
2.15	0.0147	0.0191	0.0213
3	0.1522	0.1444	0.1354

Table 1. Values of R_b.

1.08, and 0.86, respectively. These results are calculated from the leading term approximation of the eigenfunction expansion for $[LR]$ obtained in a previous subsection. Their accuracy has been confirmed by numerical simulations of the nonlinear IBVP in [Vargas 2007].

If the organism size remains unchanged ($\kappa = 1$), the signal robustness index well into the unacceptable range as specified in [Lander et al. 2005c]. A reduction factor of $\kappa = 1/\sqrt{2}$ reduces R_b by a factor greater than 20 to well below the acceptable range of 10%. A slightly smaller κ would reduce R_b toward the minimum point but not practical. Beyond $\kappa_{\min}$ given in (86), further reduction would only worsen the robustness index (and hence the gradient differential).

8.2. *Root-mean-square signal displacement.* The signal robustness index R_b is not the only measure of the deviation of the modified signaling gradient from the original one prior to Dpp synthesis rate doubling. Given an existing genetic program for individual cells, a more relevant measure of robustness may be the displacement of the same level of morphogen-receptor complex concentration due to a change of morphogen synthesis rate. Let $\bar{b}(x, 0)$ and $\tilde{b}(x, 0)$ again be the normalized signaling morphogen-receptor gradients at location x on the apical cell surface, $z = 0$, for morphogen synthesis rate $\bar{V}_L$ and $\tilde{V}_L = 2\bar{V}_L$, respectively. Let x and $\tilde{x}$ be the corresponding location where they attain the value b, i.e., $\bar{b}(x, 0) = \tilde{b}(\tilde{x}, 0) = b$. With a change of ligand synthesis rate, $\tilde{x}$ is generally different from x with $\tilde{x} - x = \Delta x$. The root-mean-square of Δx over the range of b would be another meaningful measure of robustness:

$$R_x = \frac{(\Delta x)_{\text{rms}}}{x_\ell - x_u} = \frac{1}{x_\ell - x_u} \sqrt{\frac{1}{b_u - b_\ell} \int_{b_\ell}^{b_u} (\Delta x)^2 db}. \tag{87}$$

To minimize the effects of outliers, we may limit the range of b to be the interval (b_ℓ, b_u) with $0 \le b_\ell < b_u \le \bar{b}(0)$. (We may take $b_\ell = \frac{1}{10}\bar{b}(0)$ and $b_u = \frac{9}{10}\bar{b}(0)$, for instance.)

In general, the dependence of displacement Δx on κ is through the expression

$$\bar{b}(x, 0) \sim \frac{1}{\alpha_z} A_1(x) \approx \frac{\bar{v}_z x_m v_1}{\alpha_z \Lambda_1} \exp(-\Lambda_1 x)$$

for $x \ge 0$ and the corresponding expression for $\tilde{b}(x, 0)$. Inverting these relations, we obtain

$$x \approx -\frac{1}{\Lambda_1} \ln(\eta \Lambda_1 b), \quad \tilde{x} \approx -\frac{1}{\tilde{\Lambda}_1} \ln(\tilde{\eta} \tilde{\Lambda}_1 b), \tag{88}$$

where

$$\eta = \frac{\alpha_z}{\bar{v}_z x_m v_1} = \frac{1}{\beta_z x_m v_1 \mu_z^2}, \qquad \tilde{\eta} = \frac{\eta}{2\kappa^2}, \qquad \beta_z = \frac{\bar{v}_z}{g_z}. \tag{89}$$

From these we obtain

$$\Delta x = \frac{1}{\Lambda_1} \ln(\eta \Lambda_1 b) - \frac{1}{\tilde{\Lambda}_1} \ln(\tilde{\eta} \tilde{\Lambda}_1 b) \equiv c_0(\Lambda_1, \tilde{\Lambda}_1) + c_1(\Lambda_1, \tilde{\Lambda}_1) \ln b, \tag{90}$$

with

$$c_0(\Lambda_1, \tilde{\Lambda}_1) = \frac{1}{\Lambda_1} \ln(\Lambda_1 \eta) - \frac{1}{\tilde{\Lambda}_1} \ln(\tilde{\Lambda}_1 \tilde{\eta}), \qquad c_1(\Lambda_1, \tilde{\Lambda}_1) = \left(\frac{1}{\Lambda_1} - \frac{1}{\tilde{\Lambda}_1} \right). \tag{91}$$

It follows that

$$R_x = \frac{(\Delta x)_{\text{rms}}}{x_\ell - x_u} = \frac{1}{x_\ell - x_u} \sqrt{\frac{1}{b_u - b_\ell} \int_{b_\ell}^{b_u} \left[c_0(\Lambda_1, \tilde{\Lambda}_1) + c_1(\Lambda_1, \tilde{\Lambda}_1) \ln b \right]^2 db}$$

$$= \frac{1}{x_\ell - x_u} \sqrt{\frac{1}{b_u - b_\ell} \left[c_0^2 b + 2c_0 c_1 b(\ln b - 1) + c_1^2 \left(b(\ln b)^2 - 2b(\ln b - 1) \right) \right]_{b_\ell}^{b_u}}$$

For sample calculations, we take $b_u = b(0)$ and $b_\ell = b(5) \approx 0$, so that

$$R_x = \tfrac{1}{5} \sqrt{c_0^2 + 2c_0 c_1 (\ln b_u - 1) + c_1^2 \left((\ln b_u)^2 - 2(\ln b_u - 1) \right)} \tag{92}$$

Evidently, R_x should be as small as possible to minimize the deviation from the normal gradient with $R_x = 0$ being no change. Alternatively, we may choose κ so that $R_x(\kappa)$ is in an acceptable range of values. Below are some results for $R_x(\kappa)$ for several choice values of κ for the same typical set system parameter values used for Table 1. However, unlike R_b, the parameter $\bar{v}_z$ now appears explicitly (through η in c_0, see (92), (89) and (91); it was taken so that $\beta_z = \bar{v}_z/g_z = \tfrac{1}{4}$ for the results in Table 2. These results are calculated from the leading term approximation of the eigenfunction expansion for $[LR]$ obtained in a previous subsection. Again, their accuracy has been confirmed by numerical simulations [Strikwerda 1989] of the nonlinear IBVP in [Vargas 2007].

If the organism size remains unchanged, the value of the signal robustness index is well above the acceptable range (see also [Lander et al. 2005c]). An amplification factor $\kappa = 1/\sqrt{2}$ reduces R_x by at least an order of magnitude to well within the acceptable range for (size-normalized) robustness. Beyond the

	$\sigma_z = 1/\gamma_z$		
$1/\kappa^2$	1.0	0.5	0.2
1	0.7564	0.6177	0.5130
1/2	0.0761	0.0407	0.0231
1/2.025	0.0649	0.0316	0.0156
1/2.05	0.0546	0.0239	0.0109
1/2.10	0.0392	0.0200	0.0185
1/2.15	0.0373	0.0331	0.0340
1/2.16	0.0388	0.0365	0.0373
1/3	0.3713	0.3229	0.2783

Table 2. Values of R_x.

	exact		linear	
$1/\kappa^2$	R_b	R_x	R_b	R_x
1	0.3833	0.2488	0.3873	0.2460
2	0.1477	0.1570	0.1422	0.1652
2.75	0.0627	–	0.0543	–
3	0.0503	0.1017	0.0467	0.939
3.75	–	0.0687	–	0.0757
4	0.0661	0.0736	0.0701	0.1206

Table 3. Comparison of exact and approximate values of R_b and R_x (linear approximation).

minimum point (slightly smaller than $1/\sqrt{2}$), a further reduction in κ does not improve R_x but worsens it instead.

8.3. *General receptor occupancy.* The analysis and numerical results presented in the previous subsections are based on the leading term perturbation solution for the steady state problem though confirmed by numerical simulation of the corresponding nonlinear IBVP for the low receptor occupancy case. Without the assumption of low receptor occupancy, numerical simulations for the nonlinear BVP (14), (15), and (26) have been carried out in [Vargas 2007] to show that

- size–normalized robustness persists for typical sets of biologically realistic system parameter values, and

- linearized theory is adequate in estimating appropriate self-similar size reduction for size-normalized robustness for the same sets of system parameter values.

Some sample results are shown in Table 3 for $\bar{v}_z = 10^{-3}$ and $\ell_M = 5$ with the remaining parameter values identical to those used for Tables 1 and 2.

9. Multifactor robustness

An exogenous change affects more than just one intrinsic characteristic of a biological organism. While we do not have experimental data to document the effects of a temperature change on other system parameters beside the Dpp synthesis rate, we generally expect the rate constants k_{deg}, k_{off}, k_g and $\bar{V}_R$ to increase with temperature, but are less certain about the effect on the remaining system parameters such as D and k_{on}. We examine here how simultaneous (hypothetical) changes in several system parameters (in response to an external change) may alter the wing disc gradient, at least as dictated by our model. For simplicity, we assume the effect of the external change on the diffusion coefficient D to be negligible in the range of the change considered. Typically, we proceed with the analytical results for the biologically realistic case of low receptor occupancy which enable us to explore how size changes may ameliorate the net effects of the changes in several system characteristics on the signaling gradient. Such an investigation constitutes a first attempt to address theoretically the *multifactor robustness* of temperature change beyond its effects through the single parameter of Dpp synthesis rate.

Some observations can be made immediately about the effect of changing two of the system parameters. First, the dissociation rate constant k_{off} is known to be much smaller than k_{deg} so that $\alpha_z = (g_z + f_z)/h_z \simeq g_z/h_z$. The following is an immediate consequence of this observation and (26):

Proposition 7. *Changes in the dissociation rate constant k_{off} have negligible effect on the signaling Dpp gradient* [*LR*].

This conclusion is not restricted to the low receptor saturation case, since (26) holds in general.

Next, with $\mu_z^2 = g_z h_z/(g_z + f_z) \simeq h_z = O(10^{-1})$, the approximate expression (67) applies as long as the change in h_z is not larger by an order of magnitude. In that case, Proposition 4 continues to hold and, upon rewriting (70) as

$$[LR] \sim \frac{\bar{V}_L X_{\min} Z_{\max}}{D(k_{\text{deg}} + k_{\text{off}})/k_{\text{on}} \bar{R}_0} \frac{v_1}{\lambda_1} e^{-\lambda_1 x} \cos(\lambda_1 z) \simeq \frac{\bar{V}_L}{k_{\text{deg}}} \frac{\bar{V}_R}{k_g} \frac{k_{\text{on}}}{D/(X_{\min} Z_{\max})} \frac{v_1}{\lambda_1} e^{-\lambda_1 x} \cos(\lambda_1 z) \quad (93)$$

for $0 < x < \infty$, which shows the magnitude of $[LR]$ to increase with the binding rate constant k_{on}, we have the following observation:

Proposition 8. *A substantial change in the binding rate* (*such as doubling its magnitude*) *would have an effect only on the magnitude of the signaling* [*LR*] *gradient*.

With many rate constants playing a substantive role in the magnitude of the signaling Dpp gradient [*LR*], many scenarios are possible depending on relative effects of a particular kind of exogenous perturbation (such as a temperature change) on these constants, resulting in the new rate constants $\tilde{V}_L$, $\tilde{V}_R$, $\tilde{k}_{\text{deg}}$, $\tilde{k}_g$, etc., with $m_L = \tilde{V}_L/V_L$, $m_{\text{deg}} = \tilde{k}_{\text{deg}}/k_{\text{deg}}$, and so on. An attempt will be made below to identify some reasonable scenarios and their consequences.

One possible scenario involves the degradation rate constants k_{deg} and k_g changing at the same rate as the corresponding protein synthesis rate $\bar{V}_L$ and $\bar{V}_R$, respectively. In that case, the following is an immediate consequence of (93), with the diffusion rate constant D assumed to be unaffected by the exogenous change:

Proposition 9. *For the low receptor occupancy case*, *if*

$$\frac{m_L}{m_{\text{deg}}} \equiv \frac{\tilde{V}_L/\bar{V}_L}{\tilde{k}_{\text{deg}}/k_{\text{deg}}} = 1, \qquad \frac{m_R}{m_g} \equiv \frac{\tilde{V}_R/\bar{V}_R}{\tilde{k}_g/k_g} = 1,$$

then a size-normalized robust gradient is attained by a self-similar reduction (*of all spatial lengths*, $X_{\min}$, $Z_{\max}$ *and* $X_{\max}$) *by a factor of* $\sqrt{2}$ *when the binding rate doubles* (*in response to a substantial exogenous change*).

More generally, when $m_L \neq m_{\text{deg}}$ and $m_R \neq m_g$, the situation is more complicated. Upon making use of (93) to form the ratio

$$\frac{[\widetilde{LR}]}{[LR]} \simeq \frac{m_L}{m_{\text{deg}}} \frac{m_R}{m_g} \frac{m_{\text{on}}}{\kappa^2} \quad (94)$$

where $m_{\text{on}} = \tilde{k}_{\text{on}}/k_{\text{on}}$ and where all wing dimensions are changed by the same factor κ so that $\tilde{X}_{\min} = \kappa X_{\min}$, etc., the next proposition also follows from (93):

Proposition 10. *With $m_D = \tilde{D}/D = 1$, size-normalized robustness is attained by changing all wing disc dimension by*

$$\kappa^2 = m_{\text{on}} \frac{m_L}{m_{\text{deg}}} \frac{m_R}{m_g}. \quad (95)$$

Evidently, a more optimal κ can be found with the help of the robustness indices R_b and R_x as was done in a previous section on the change for the rate constant $\bar{V}_L$ alone. While the observations in the last two propositions are based on the analytical results, numerical simulations show that they continue to hold for the general model for typical sets of realistic system parameter values.

10. Patterning of a smaller wing disc

It is of some interest to determine also how a higher Dpp synthesis rate and a self-similar size-reduction of the wing imaginal disc may affect the elapsed time to a steady state gradient. Note that the lowest eigenvalue, $\omega^{(0)}$, of the nonlinear eigenvalue problem (37)–(41) provides a good first estimate of this half-life measure. An asymptotic solution for this lowest eigenvalue can be obtained by a regular perturbation solution for $\omega^{(0)}$ in powers of the parameter $\zeta = g_z/g_r < 1$,

$$\omega^{(0)} = \omega_0 + \zeta \omega_1 + \zeta^2 \omega_2 + \cdots .$$

Consistent with the low receptor occupancy assumption of this paper (and similar to the development in [Lander et al. 2005b] for the one-dimensional model), the leading term ω_0 is determined by the simpler eigenvalue problem

$$\nabla^2 a_0 + [\omega_0 - q_0(x; z, \omega_0)] a_0 = 0, \tag{96}$$

where

$$q_0(x; z, \omega_0) = \frac{h_z r_0(x, z)(g_z - \omega_0)}{g_z + f_z - \omega_0}, \quad r_0(x, z) = \rho^2 H(-x) + H(x), \tag{97}$$

with the four boundary conditions (14)–(15) applied to $a_0(x, z; \omega_0)$.

To solve this simpler eigenvalue problem, we note that $q_0(x; z, \omega_0)$ is independent of z given that the receptor synthesis rate is chosen so that r_0 depends only on x. We can then separate out the z-dependent part by setting

$$a_0(x, z) = \sum_{j=1}^{\infty} A_j(x) \cos(\lambda_j z),$$

where $\{\lambda_j\}$ are the solutions of (49). Orthogonality of the functions $\{\cos(\lambda_j z)\}$ leads to the eigenvalue problem

$$A_j'' + \left(\omega_0 - \lambda_j^2 - q_0(x; \omega_0)\right) A_j = 0, \qquad A_j'(-x_m) = 0, \qquad A_j(\ell_M) = 0, \tag{98}$$

where a prime $'$ indicates differentiation with respect to x and $j = 1, 2, 3, \ldots$. For $\rho^2 = 1$ so that $r_0(x, z) = 1$ for all $(x, z) \in \bar{\Omega}$, the solutions of the eigenvalue problem (98) are the eigenpairs

$$A_{kj}(x) = a_{kj} \cos(\pi_k(x + x_m)), \qquad \pi_k = \frac{\pi(2k - 1)}{2(\ell_M + x_m)} \tag{99}$$

for $k, j = 1, 2, \ldots$, with

$$\omega_0 = \pi_k^2 + \lambda_j^2 + \frac{h_z(g_z - \omega_0)}{g_z + f_z - \omega_0}. \tag{100}$$

The relation (100) is a quadratic equation for ω_0; finding its two roots, denoted by $\omega_{kj}^{(1)}$ and $\omega_{kj}^{(2)}$, is straightforward. Given $f_z \ll g_z \ll h_z < 1$, these roots are approximated by

$$\omega_{kj}^{(1)} \approx \xi_{kj} \approx \pi_k^2 + \lambda_j^2 + h_z, \qquad \omega_{kj}^{(2)} \approx \frac{\delta_{kj}}{\xi_{kj}} \approx g_z. \tag{101}$$

with $g_z \approx \omega_{kj}^{(2)} \ll \omega_{kj}^{(1)} \approx \pi_k^2 + \lambda_j^2 + h_z$. The decay rate is dominated by the smallest (lowest) $\omega_{kj}^{(2)}$, namely $\omega_{kj}^{(2)} \approx \omega_{11}^{(2)} \approx g_z$. With

$$\omega_{kj}^{(2)} t \approx g_z t = k_{\text{deg}} T,$$

the expression $e^{-g_z t} = e^{-k_{\text{deg}} T}$ gives a half life of $1/k_{\text{deg}}$ which does not vary with size changes in $Z_{\max}$, $X_{\min}$ or $X_{\max}$. We have then the following result on time to steady state gradient:

Proposition 11. *Starting with some small initial (perturbation from the steady) state, the time to a time-independent steady state is determined by the lowest member of the second family of frequencies, which is approximately g_z. Moreover, the corresponding actual half-life is independent of size changes and depends only on k_{deg}.*

It follows that the half life of transients would vary only with a change in k_{deg} (caused by some exogenous change) and a size change (self-similar or not) would have no effects on this change. In particular, doubling the magnitude of k_{deg} (caused by a temperature rise) would reduce by half the half life of transients independent of any size changes.

We note that the situation is not the same for the other family $\{\omega_{kj}^{(1)}\}$. With

$$\omega_{kj}^{(1)} \approx \tilde{\omega}_{kj}^{(1)} \equiv \pi_k^2 + \lambda_j^2$$

(which is also independent of the size parameters $\{Z_{\max}, X_{\min}, X_{\max}\}$) and with $t = DT/Z_{\max}^2$, a size change via $Z_{\max}$ by an amplification factor κ leads to

$$T = \frac{t}{D} \tilde{Z}_{\max}^2, \qquad \omega_{kj}^{(1)} t = \frac{D\omega_{kj}^{(1)}}{\tilde{Z}_{\max}^2} T = \frac{D}{\kappa^2 Z_{\max}^2} T (\pi_k^2 + \lambda_j^2). \tag{102}$$

Hence, it would take less time for all harmonics of a smaller disc associated with the first family of (faster) decay rates $\omega_{kj}^{(1)}$ to reach the same level of concentration of [*LR*]. As previously mentioned, size changes have no effect on the actual time to steady state given (102). The main purpose in discussing the family $\{\omega_{kj}^{(1)}\}$ at all is to show that its effect on half life is in the same direction as that of $\omega_{kj}^{(2)} \simeq g_z$ and hence preserving the size-normalized robust nature of the latter when there is a significant exogenous change.

While a reduction of half life of transients is not synonymous with the speed of growth observed in nature, e.g., faster growth of fruit flies in hotter temperature [de Moed et al. 1997; French et al. 1998; Bitner-Mathé and Klaczko 1999; Azevedo et al. 2002], it certainly is not inconsistent with such observations and may very well contribute to its realization.

11. Concluding remarks

Exogenous (environmental) changes, such as a substantial change in ambient temperature, are expected to affect multiple intrinsic system characteristics such as morphogen and receptor synthesis rates, various

binding, dissociation and degradation rate constants, and diffusion rates, of a biological organism. If the effects are substantial as in the case of a 6°C change in ambient temperature on the Dpp synthesis rate in the *Drosophila* wing imaginal disc, they could alter the biological development of the organism. The temperature–Dpp synthesis rate relation, being the only known data of this type, prompted us to initiate an investigation, of at first limited scope, on the effects of exogenous changes on biological development by examining the effect of the synthesis rate change on the signaling Dpp–Tkv concentration gradient. For biologically realistic gradients occurring in the low receptor occupancy range, we show that a signaling gradient is approximately proportional to the Dpp synthesis rate. It is then shown that the effect of a synthesis rate change may be ameliorated by self-similar size changes to result in a (size-normalized) robust development as measured by several different robustness indices introduced in Section 8.

To the extent that a single exogenous perturbation affects generally more than a single system characteristic, the paper also investigates briefly, in the absence of experimental data, how the effects of multiple system characteristic changes on the signaling gradient can also be ameliorated by self-similar size changes, resulting again in size-normalized robustness.

Changes in morphogen gradient shape with temperature and changes in wing disc size with temperature are probably not independent phenomena, given that Dpp and its receptor Tkv control not just cell identity, but also proliferation [Martin et al. 2004; Martin-Castellanos and Edgar 2002]. Indeed, we do not know whether the primary effect of temperature is to change the size of a morphogen field, and then the morphogen gradient responds by re-scaling to fit the new size, or if the primary effect of temperature is on the morphogen gradient, and the patterned field changes its size in response to the morphogen gradient. Most likely reality lies somewhere between these extremes. (We do know, however, that the wing disc BMP gradient can re-scale to fit alterations in disc size imposed by perturbations other than temperature [Teleman and Cohen 2000].)

Clearly, there is much we do not know about the effects of temperature on biological development or, more fundamentally, its interaction with morphogen gradients. Given how important temperature effects are likely to have been in evolution, there is a great need to gather some of the basic information that will be required to understand these effects. The effort of this paper is a step in this direction. Our results have shown that, whether synthesis rate change is a direct or downstream effect of a temperature change, a (self-similar) change of wing disc size would nearly offset the effects of a doubling of morphogen synthesis rate on the signaling gradient to orchestrate a size-normalized development of wing discs.

At the organismic level, flies grow slower at low temperature and to larger size including having larger wings; they have smaller wings and grow faster (but smaller) at higher temperatures [de Moed et al. 1997; French et al. 1998; Bitner-Mathé and Klaczko 1999; Azevedo et al. 2002]. The faster growth rate of the flies at higher temperature is probably related to faster biochemical reactions not addressed by the mathematical model of the present investigation. While the molecular mechanisms that lead to a speed up of the relevant biochemical reactions are unknown to the authors, faster growth may simply be a consequence of (self-similar) size reduction (to maintain size-normalized development robustness) requiring less biological constituent parts for a smaller fly. Time to a steady-state signaling gradient is not synonymous with biological growth rate. Nevertheless, it seems reasonable that the half life to steady state signaling gradient should be consistent with the different growth rate. Given the result of the previous section, it would appear that degradation rate constants k_{deg} should also change

with exogenous temperature perturbations in order to affect the time to a steady state signaling Dpp–Tkv complex gradient (which should contribute to but is unlikely to be completely responsible for the speed of general development and growth). As there are no measurements available to validate the size-normalized multifactor robustness of development (and growth) predicted by our analysis and numerical simulations, it is hoped that observations made herein would stimulate experimental work for measuring the effects of a significant change in temperature on the various rate constants as was done for the Dpp synthesis rate $\bar{V}_L$ (in the Lander lab) mentioned in the first section of this paper.

References

[Amann 1972] H. Amann, "On the existence of positive solutions of nonlinear elliptic boundary value problems", *Indiana Univ. Math. J.* **21**:2 (1972), 125–146.

[Azevedo et al. 2002] R. B. R. Azevedo, V. French, and L. Partridge, "Temperature modulates epidermal cell size in *Drosophila melanogaster*", *J. Insect. Physiol.* **48**:2 (2002), 231–237.

[Bender and Orszag 1999] C. M. Bender and S. A. Orszag, *Advanced mathematical methods for scientists and engineers, I: Asymptotic methods and perturbation theory*, Springer, New York, 1999.

[Bitner-Mathé and Klaczko 1999] B. C. Bitner-Mathé and L. B. Klaczko, "Plasticity of *Drosophila melanogaster* wing morphology: effects of sex, temperature and density", *Genet.* **105**:2 (1999), 203–210.

[Dowd et al. 1999] C. J. Dowd, C. L. Cooney, and M. A. Nugent, "Heparan sulfate mediates bFGF transport through basement membrane by diffusion with rapid reversible binding", *J. Biol. Chem.* **274**:8 (1999), 5236–5244.

[Entchev et al. 2000] E. V. Entchev, A. Schwabedissen, and M. González-Gaitán, "Gradient formation of the TGF-β homolog Dpp", *Cell* **103**:6 (2000), 981–991.

[French et al. 1998] V. French, M. Feast, and L. Partridge, "Body size and cell size in *Drosophila*: the developmental response to temperature", *J. Insect. Physiol.* **44**:11 (1998), 1081–1089.

[Gurdon and Bourillot 2001] J. B. Gurdon and P.-Y. Bourillot, "Morphogen gradient interpretation", *Nature* **413**:6858 (2001), 797–803.

[Khong and Wan 2007] M. Khong and F. Y. M. Wan, "Negative feedback in morphogen gradients", pp. 29–51 in *Frontiers of applied mathematics* (Proc. of 2nd International Symposium) (Beijing, 2006), edited by D.-Y. Hsieh et al., World Scientific, Singapore, 2007.

[Lander et al. 2002] A. D. Lander, Q. Nie, and F. Y. M. Wan, "Do morphogen gradients arise by diffusion?", *Dev. Cell* **2**:6 (2002), 785–796.

[Lander et al. 2005a] A. D. Lander, Q. Nie, B. Vargas, and F. Y. M. Wan, "Aggregation of a distributed source in morphogen gradient formation", *Stud. Appl. Math.* **114**:4 (2005), 343–374.

[Lander et al. 2005b] A. D. Lander, Q. Nie, and F. Y. M. Wan, "Spatially distributed morphogen production and morphogen gradient formation", *Math. Biosci. Eng.* **2**:2 (2005), 239–262.

[Lander et al. 2005c] A. D. Lander, F. Y. M. Wan, and Q. Nie, "Multiple paths to morphogen gradient robustness", preprint, Center for Complex Biological Systems, University of California, Irvine, CA, 2005, available at https://eee.uci.edu/11s/45869/home/Draft5.pdf.

[Lander et al. 2006] A. D. Lander, Q. Nie, and F. Y. M. Wan, "Internalization and end flux in morphogen gradient formation", *J. Comput. Appl. Math.* **190**:1-2 (2006), 232–251.

[Lou et al. 2004] Y. Lou, Q. Nie, and F. Y. M. Wan, "Nonlinear eigenvalue problems in the stability analysis of morphogen gradients", *Stud. Appl. Math.* **113**:2 (2004), 183–215.

[Martin-Castellanos and Edgar 2002] C. Martin-Castellanos and B. A. Edgar, "A characterization of the effects of Dpp signaling on cell growth and proliferation in the *Drosophila* wing", *Dev.* **129**:4 (2002), 1003–1013.

[Martin et al. 2004] F. A. Martin, A. Pérez-Garijo, E. Moreno, and G. Morata, "The *brinker* gradient controls wing growth in *Drosophila*", *Dev.* **131**:20 (2004), 4921–4930.

[de Moed et al. 1997] G. H. de Moed, G. de Jong, and W. Scharloo, "Environmental effects on body size variation in *Drosophila melanogaster* and its cellular basis", *Genet. Res.* **70**:1 (1997), 35–43.

[Sattinger 1972] D. H. Sattinger, "Monotone methods in nonlinear elliptic and parabolic boundary value problems", *Indiana Univ. Math. J.* **21**:11 (1972), 979–1000.

[Smoller 1983] J. Smoller, *Shock waves and reaction-diffusion equations*, Grundlehren der Mathematischen Wissenschaften **258**, Springer, New York, 1983. 2nd ed. published in 1994.

[Strikwerda 1989] J. C. Strikwerda, *Finite difference schemes and partial differential equations*, Wadsworth & Brooks/Cole, Pacific Grove, CA, 1989.

[Teleman and Cohen 2000] A. A. Teleman and S. M. Cohen, "Dpp gradient formation in the *Drosophila* wing imaginal disc", *Cell* **103**:6 (2000), 971–980.

[Umulis et al. 2008] D. M. Umulis, M. B. O'Connor, and H. G. Othmer, "Robustness of embryonic spatial patterning in *Drosophila melanogaster*", *Curr. Top. in Dev. Biol.* **81** (2008), 65–111.

[Vargas 2007] B. Vargas, *Leaky boundaries and morphogen gradient formation*, Ph.D. thesis, University of California, Irvine, CA, 2007.

ARTHUR D. LANDER: adlander@uci.edu
*Department of Developmental and Cell Biology, University of California, Irvine, 2638 Biological Science III,
Irvine, CA 92697-2300, United States*
http://lander-office.bio.uci.edu/

QING NIE: qnie@uci.edu
Department of Mathematics, University of California, Irvine, 510B Rowland Hall, Irvine, CA 92697-3875, United States
http://www.math.uci.edu/~qnie/

BENJAMIN VARGAS: paryvargas@hotmail.com
Department of Mathematics, Irvine Valle College, 5500 Irvine Center Drive, Irvine, CA 92618-0300, United States
and
Department of Mathematics, University of California, Irvine, 510B Rowland Hall, Irvine, CA 92697-3875, United States

FREDERIC Y. M. WAN: fwan@uci.edu
Department of Mathematics, University of California, Irvine, 510B Rowland Hall, Irvine, CA 92697-3875, United States
http://www.math.uci.edu/~fwan/

■msp

FRACTALS IN THERMOELASTOPLASTIC MATERIALS

Jun Li and Martin Ostoja-Starzewski

Fractal patterns are observed in computer simulations of elastic-plastic transitions in linear, locally isotropic thermoelastic-hardening plastic heterogeneous materials. The models involve 2D aggregates of homogeneous grains with weak random fluctuations in thermal expansion coefficients, equivalent to modeling the effects of random residual strains. The spatial assignment of material randomness is a nonfractal, strict-white-noise random field on a 256×256 square lattice. The flow rule of each grain follows associated plasticity with loading applied through either one of three macroscopically uniform boundary conditions admitted by the Hill–Mandel condition. Upon following the evolution of a set of grains that become plastic (*plastic set*), we find that it has a fractal dimension increasing smoothly from 0 towards 2. Transitions under various types of model randomness and combinations of material constants are examined. While the grains possess sharp elastic-plastic stress-strain curves, the overall stress-strain responses are smoothly curved and asymptote toward plastic flows of reference homogeneous media. As the randomness decreases to zero, they turn into conventional curves with sharp kinks of homogeneous materials. Overall, the fractal dimension D of the plastic set is a readily accessible parameter to investigate transition patterns in many materials.

1. Background

Given that many materials display fractal characteristics (see [Mandelbrot 1982; Feder 1988], for instance), fractal concepts have been used in the geometric characterization as well as morphogenesis models of spatial patterns. Numerous such phenomena, both in natural and artificial materials, include phase transitions and accretion, fracture surfaces and dislocation patterns; see, e.g., [Sornette 2006]. However, very little work was done on fractals in elasto-plasticity, except for the research on plastic ridges in ice fields [Ostoja-Starzewski 1990; Overland et al. 1995] and on shear bands in rocks of Mohr–Coulomb type [Poliakov et al. 1994]. More recently we observed fractal patterns under two-dimensional plane strain loadings at elastic-to-plastic transitions in random, linear elastic-perfectly (or hardening) plastic materials [Li and Ostoja-Starzewski 2010a; 2010b] made of homogeneous, isotropic or anisotropic grains. This paper focuses on thermoelastic-hardening plastic homogeneous materials, where the thermal expansion coefficients are assigned from a (nonfractal) strict-white-noise random field. The reason for this type of a random field assumption is that the evolution of plastic zones would very likely be fractal should the material properties be fractally distributed at the outset. In all the cases we study, it turns out that the elastic-to-plastic transition occurs through a fractal set of plastic grains, gradually plane-filling the entire material domain under monotonic loading. We also study several related cases by varying model configurations.

2. Model formulation

As conventionally done in mechanics of random media, we consider the random heterogeneous material to be a set $\mathcal{B} = \{B(\omega); \omega \in \Omega\}$ of realizations $B(\omega)$, defined over the sample space Ω, each one evolving in a deterministic fashion [Ostoja-Starzewski 2008]. That is, for an elementary event $\omega \in \Omega$ we have a realization of deterministic media $B(\omega)$, each taken as an aggregate of crystals (or grains). With $\mathcal{B}$ embedded in a physical space, the aggregate is essentially modeled by a random field. Any material property, say $\boldsymbol{G}$, is required to be mean-ergodic, that is

$$\overline{\boldsymbol{G}(\omega)} \equiv \lim_{L \to \infty} \frac{1}{V} \int_V \boldsymbol{G}(\omega, \boldsymbol{x}) \, dV = \int_\Omega \boldsymbol{G}(\omega, \boldsymbol{x}) \, dP(\omega) \equiv \langle \boldsymbol{G}(\boldsymbol{x}) \rangle \tag{2-1}$$

where the overbar means the volume average and $\langle \ \rangle$ indicates the ensemble average. $P(\omega)$ is the probability measure assigned to the ensemble $\{\boldsymbol{G}(\omega, \boldsymbol{x}); \omega \in \Omega, \boldsymbol{x} \in V\}$ and its σ-algebra. In general, the grains are homogeneous, isotropic, linear (thermo)elastic-hardening-plastic materials, where the randomness just resides in either the moduli, or plastic limits, or thermal expansion coefficients. Thus, the constitutive response of each grain is

- when $f < c$ (thermoelastic region):

$$\epsilon_{ij} = S_{ijkl}\sigma_{kl} + \alpha_{ij}(\omega, \boldsymbol{x})\theta; \tag{2-2}$$

- when $f \geq c$ and $df \geq 0$ (plastic region):

$$d\epsilon'_{ij} = \frac{d\sigma'_{ij}}{2G} + \lambda \cdot \frac{\partial f}{\partial \sigma_{ij}}, \quad d\epsilon = \frac{d\sigma}{K} \quad \left(d\epsilon = \tfrac{1}{3}d\epsilon_{ii}, d\sigma = \tfrac{1}{3}d\sigma_{ii}\right). \tag{2-3}$$

Here primes indicate deviatoric tensor components, S_{ijkl} is the compliance tensor, $\alpha_{ij}(\omega, \boldsymbol{x})$ is the thermal expansion coefficient (randomly specified in each grain), $\theta = (T - T_0)$ is the temperature change, f is the yield function following the associated J_2 flow theory, c is the yield limit, λ is a consistency parameter, and G and K are the shear and bulk moduli, respectively. Clearly, the randomness in thermal expansion coefficients effectively models random residual strains ($\epsilon_{ij}^R(\omega, \boldsymbol{x}) = \alpha_{ij}(\omega, \boldsymbol{x})\theta$).

Regarding the loading of $\mathcal{B}$, we recall the Hill–Mandel condition, which guarantees the equivalence of energetically and mechanically defined effective responses:

$$\overline{\boldsymbol{\sigma} : d\boldsymbol{\epsilon}} = \bar{\boldsymbol{\sigma}} : d\bar{\boldsymbol{\epsilon}} \iff \int_{\partial B_\delta} (\boldsymbol{t} - \bar{\boldsymbol{\sigma}} \cdot \boldsymbol{n}) \cdot (d\boldsymbol{u} - d\bar{\boldsymbol{\epsilon}} \cdot \boldsymbol{x}) \, dS = 0, \tag{2-4}$$

where ∂B_δ is the boundary of a given specimen B_δ of size δ, see also [Hazanov 1998]. This equation suggests three special types of uniform boundary conditions (BCs):

(i) uniform displacement BC:

$$d\boldsymbol{u} = d\bar{\boldsymbol{\epsilon}} \cdot \boldsymbol{x} \tag{2-5}$$

(ii) uniform traction BC:

$$\boldsymbol{t} = \bar{\boldsymbol{\sigma}} \cdot \boldsymbol{n} \tag{2-6}$$

(iii) uniform mixed-orthogonal BC:

$$(\boldsymbol{t} - \bar{\boldsymbol{\sigma}} \cdot \boldsymbol{n}) \cdot (d\boldsymbol{u} - d\bar{\boldsymbol{\epsilon}} \cdot \boldsymbol{x}) = 0 \tag{2-7}$$

3. Fractal patterns from computational mechanics

Given the lack of analytical solutions for study of patterns in random (thermo)inelastic materials, a numerical simulation of the elastic-plastic transition, in plane strain, is carried out with the ABAQUS FEM software. The domain comprises 256×256 square-shaped grains, i.e., the domain is sufficiently large to compute fractal dimensions. Each grain is homogeneous and isotropic, its thermal expansion coefficient α being a uniform random variable (r.v.) from a range up to $\pm 2.5\%$ about the mean and other material parameters being constant. The mean values are taken from the *ABAQUS Example Manual 5.1.2*: $E = 93.5$ GPa, $h = 76.5$ GPa, $c = 153$ MPa, $\alpha = 11.7 \times 10^{-6}$ /K, $\nu = 0.27$. The temperature change is set to be $\theta = 20$ K. We apply shear loading through one of three types of uniform BC consistent with (2-5)–(2-7):

$$\text{Displacement}: \quad d\bar{\epsilon}_{11} = -d\bar{\epsilon}_{22} = d\epsilon^0, \qquad d\bar{\epsilon}_{12} = 0, \tag{3-1}$$

$$\text{Traction}: \quad \bar{\sigma}_{11} = -\bar{\sigma}_{22} = \sigma^0, \qquad \bar{\sigma}_{12} = 0, \tag{3-2}$$

$$\text{Mixed}: \quad d\bar{\epsilon}_{11} = d\epsilon^0, \qquad \bar{\sigma}_{22} = -\sigma^0, \qquad d\bar{\epsilon}_{12} = \bar{\sigma}_{12} = 0. \tag{3-3}$$

In the following, for the sake of clarity in the figures, we do not show results from loading under mixed-orthogonal boundary conditions because they are bounded by those from (3-1) and (3-2), which already provide very tight bounds. As these two loadings are applied, the material domains evolve from fully elastic to fully plastic by exhibiting gradually growing sets of plastic grains. We call such a set the plastic set. Furthermore, define the elastic set as the set of all the remaining grains, i.e., those that have not yet gone through the elastic-plastic transition.

Figure 1 shows elastic-plastic transition patterns for increasing stress σ^0 under traction BC. The figures use a binary format in the sense that elastic grains are white, while the plastic ones are black. As the loading increases, the plastic set grows with an apparently disordered geometry. Its fractal dimension D is estimated using the box-counting method [Feder 1988]. The fractal character of sets of plastic grains is
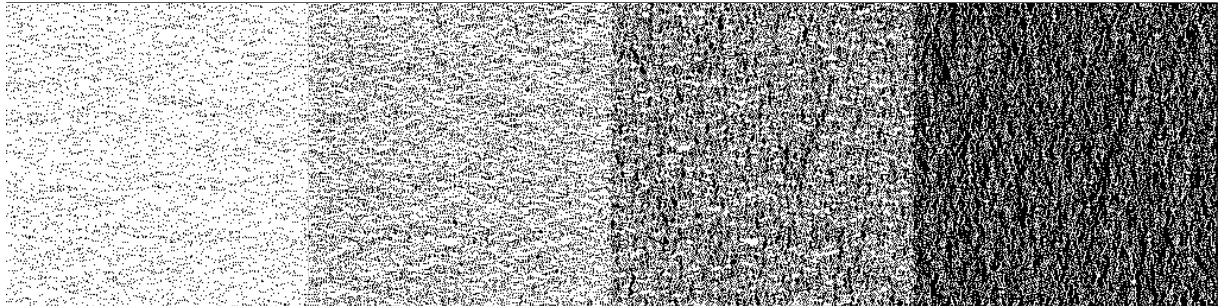
$\bar{\epsilon}_p = 0.0001$	$\bar{\epsilon}_p = 0.00025$	$\bar{\epsilon}_p = 0.00125$	$\bar{\epsilon}_p = 0.0025$
$D = 1.8502$ $R = 0.9970$	$D = 1.9477$ $R = 0.9995$	$D = 1.9984$ $R = 0.9999$	$D = 1.9999$ $R = 1.0000$

Figure 1. Field images of sets of grains that have become plastic (black) at the elastic-to-plastic transition in a 256×256 domain of square-shaped grains under uniform traction BC at four consecutive levels of normalized average plastic strain $\bar{\epsilon}_p$. Each white pixel represents one elastic grain, and each black pixel a plastic grain. The values of the fractal dimension D and the correlation coefficient R for linear fits of $\log N_r \sim \log r$, where N_r is the number of boxes with size r required to cover the object, are also given.
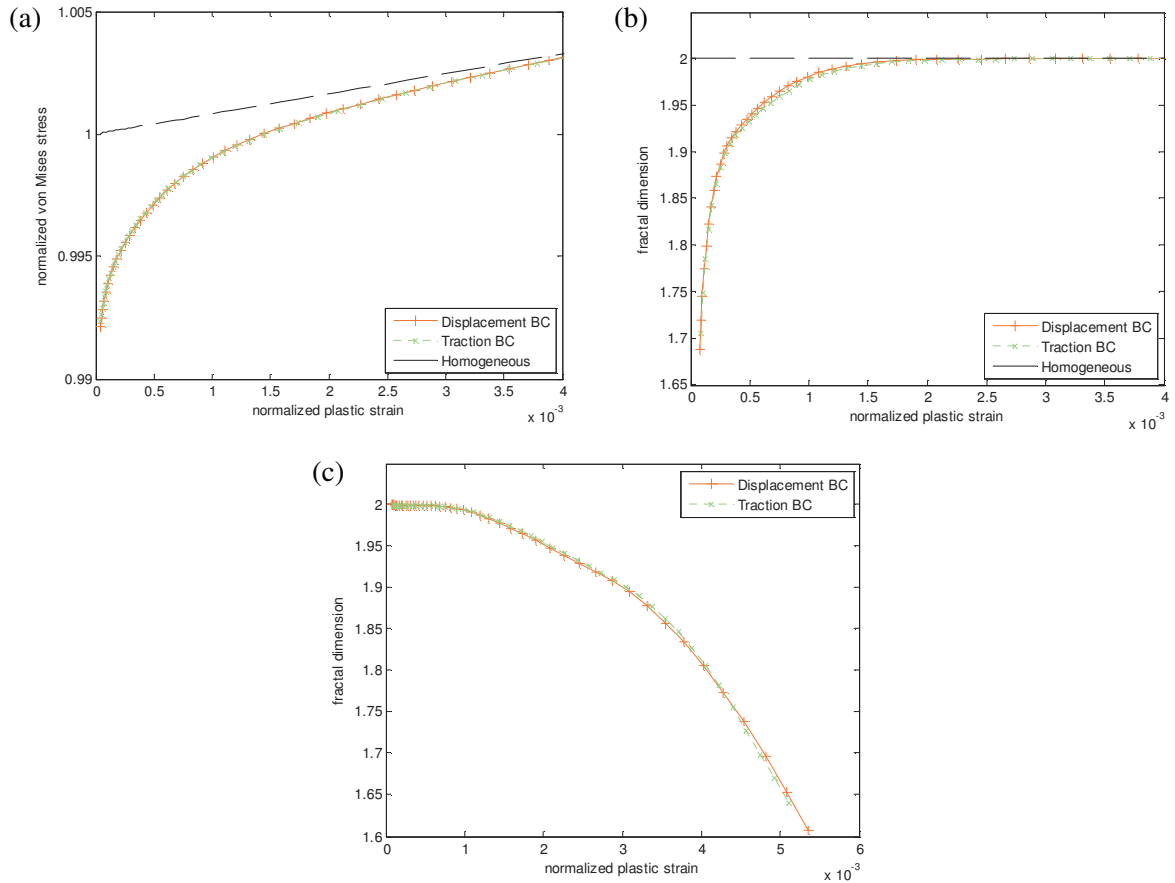
Figure 2. Response curves for monotonic loading under different BCs: (a) averaged stress~strain; (b) fractal dimension of the plastic set versus strain; (c) fractal dimension of the elastic set versus strain.

evident. The same type of results is obtained for each BC (2-5)–(2-7) and each particular material model, whereby the spread of plastic grains is always fastest under (2-5), slower under (2-7), and slowest under (2-6). Furthermore, as the noise in the material coefficient decreases from the range ±2.5% to, say, ±1%, the transition simply occurs more rapidly. i.e., over a shorter interval of the loading parameter such as applied stress. As the noise tends towards 0, the transition occurs instantaneously, in kink-like fashion.

Figure 2a shows response curves under these two BCs in terms of volume-averaged stress versus strain. They have been nondimensionized as normalized stress or strain (divided by yield stress or yield strain) for later comparative studies among materials. The responses of single grain homogeneous phases are also given for reference. Clearly, the responses of random heterogeneous materials all display smooth curves tending towards the line of homogeneous phases, which, in fact, is more realistic, since in real materials (always possessing small scale randomness) the elastic-plastic transition develops smoothly rather than through a kink-transition. Also note that the constitutive response is bounded from above (resp. below) by that under displacement (traction) BCs. This is consistent with the scale-dependent hierarchies of bounds for elastic-inelastic composites reviewed in [Ostoja-Starzewski 2008].
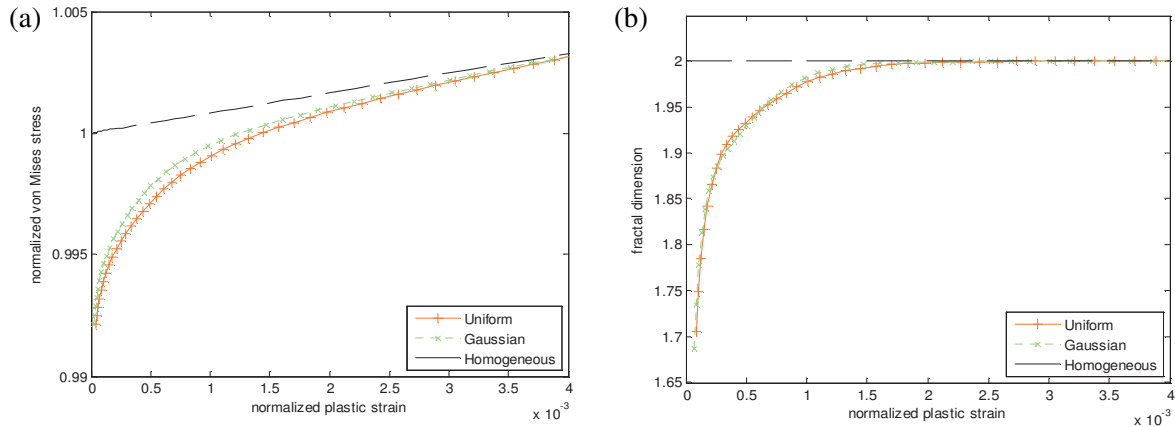
Figure 3. Comparison of elastic-plastic transitions for uniform and Gaussian random material distributions: (a) averaged stress~strain; (b) fractal dimension of the plastic set versus strain.

Figures 2b and 2c show evolutions of fractal dimensions of the plastic and elastic sets versus strain, respectively. As expected, the first of these grows from 0 towards 2, while the second decreases from 2 towards 0 although we do not show the entire range since as the box-counting estimation of fractal dimension is not reliable for very sparse set. In general, these two fractal dimensions do not add up to 2.

Next, the sensitivity of the model to various types of randomness is studied through a comparison of a uniform as opposed to a Gaussian noise; the latter is truncated at $\pm 6\sigma$. Figure 3 shows that this is a secondary effect only both, in terms of stress-strain curves and in terms of fractal dimension evolution. Since uniform randomness is effectively equivalent to Gaussian and it is simpler to implement, we examine in Figure 4 two further cases in uniform distribution with different variances. Note that, according to (2-2),
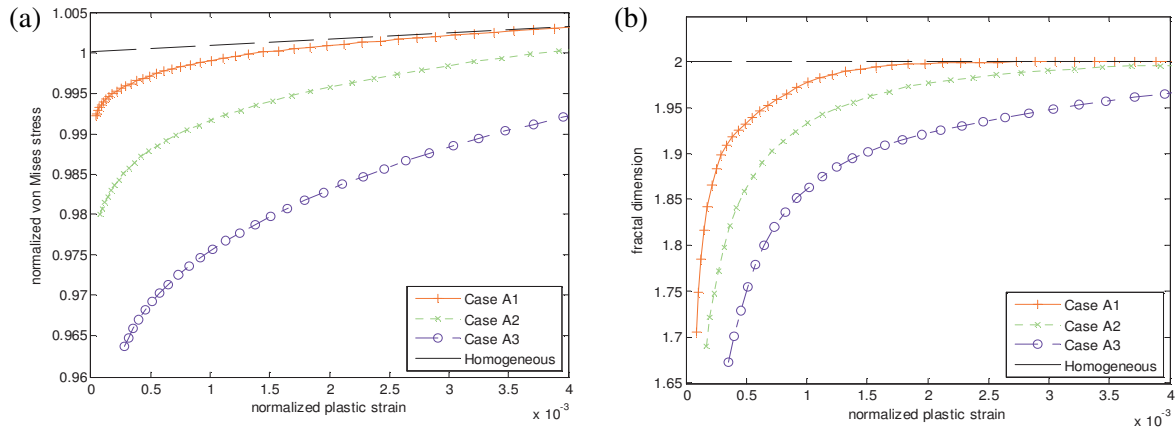
Figure 4. Comparison of elastic-plastic transitions for uniform distribution with different variances: Case A1 has uniform r.v. α up to $\pm 2.5\%$ about the mean, and $\theta = 20\,\mathrm{K}$; Case A2 has uniform r.v. α up to $\pm 12.5\%$ about the mean, and $\theta = 20\,\mathrm{K}$. Case A3 has the same variance of α as in A1, but $\theta = 100\,\mathrm{K}$. For A2 and A3, $\Delta(\alpha\theta)$ is the same but the mean value of $\alpha\theta$ is higher in A3.
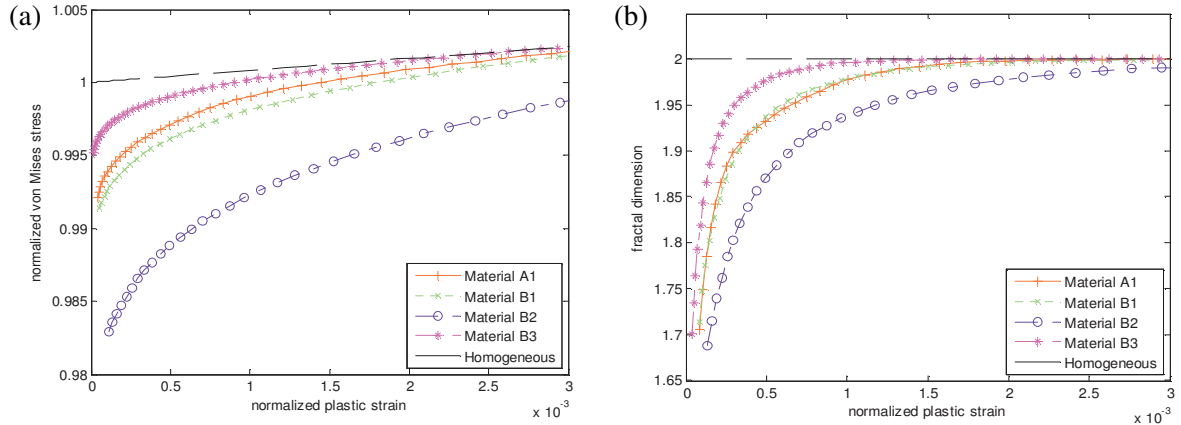
Figure 5. Comparison of elastic-plastic transitions with different material yield limits: material A1 is the same as before; material B1 has $E = 207\,\text{GPa}$, $\alpha = 13.5 \times 10^{-6}\,/\text{K}$ (from *ABAQUS Benchmark 4.7.2*), $E/h, \alpha/(c/E)$ are the same as in A1, i.e., $h = 169.36\,\text{GPa}$, $c = 390.84\,\text{MPa}$; material B2 is the same as B1 but with c reduced by half; material B3 is the same as B1 but with c doubled.

the response is affected by the multiplicity $\alpha\theta$ as a whole. Cases A2 and A3 are thus assigned the same variance $\Delta(\alpha\theta)$ but for the latter the mean $\alpha\theta$ is higher. We find that different random variances in the model configuration lead to quantitatively, but not qualitatively different transition patterns. Basically, a lower randomness results in a narrower elastic-plastic transition, and the mean value of $\alpha\theta$ takes a stronger effect when the absolute variance is fixed, both in curves of the average stress as well as the fractal dimension versus the average strain.

We next examine the elastic-to-plastic transition under differing material parameters. First, we study the effect of material yield limits on (a) the stress-strain curves and (b) the fractal dimension-strain curves. This involves a comparison of the original material A1 with three other hypothetical materials (B1, B2, B3) defined in the caption of Figure 5. Overall, we see that higher $\alpha/(c/E)$ result in a slower elastic-plastic transition, a fact which is understandable, since under these circumstances the thermal fluctuation has a stronger influence on the elasto-plastic response (ratio of residual strain versus yield strain). Our investigation culminates in Figure 6 which shows the influence of plastic hardening on the stress and fractal dimension as functions of the volume averaged plastic strain. In general, the larger is the E/h (ratio of elastic moduli to plastic moduli), i.e., the weaker are the relative hardening effects, the slower is the transition. Also, note that the homogeneous responses in stress-strain curves are distinct for materials B1, B4 and B5. The trends to approaching homogeneous response curves in conventional stress-strain calibrations are not easy to discern among different materials. On the other hand, the fractal dimension always increases from 0 to 2 during the transition, thus providing a practical parameter to assess the transition process.

4. Mesh dependence

One more issue which we address is that of mesh dependence. Namely, how would the results change if we used a different resolution of a single grain than by modeling up until now as one finite element?
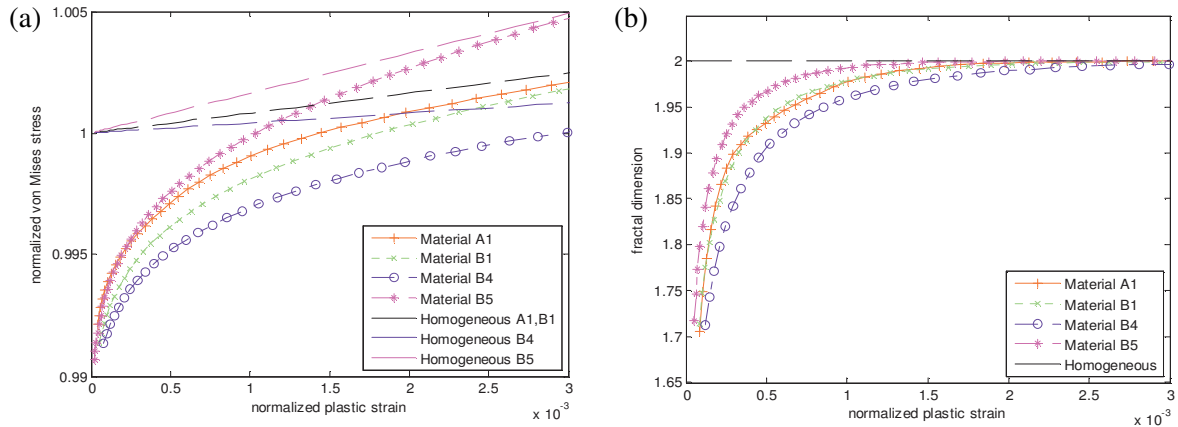
Figure 6. Comparison of elastic-plastic transitions with different hardening properties: materials A1 and B1 are the same as before; material B4 is the same as B1, but with h reduced by half; material B5 is the same as B1 but with h doubled.
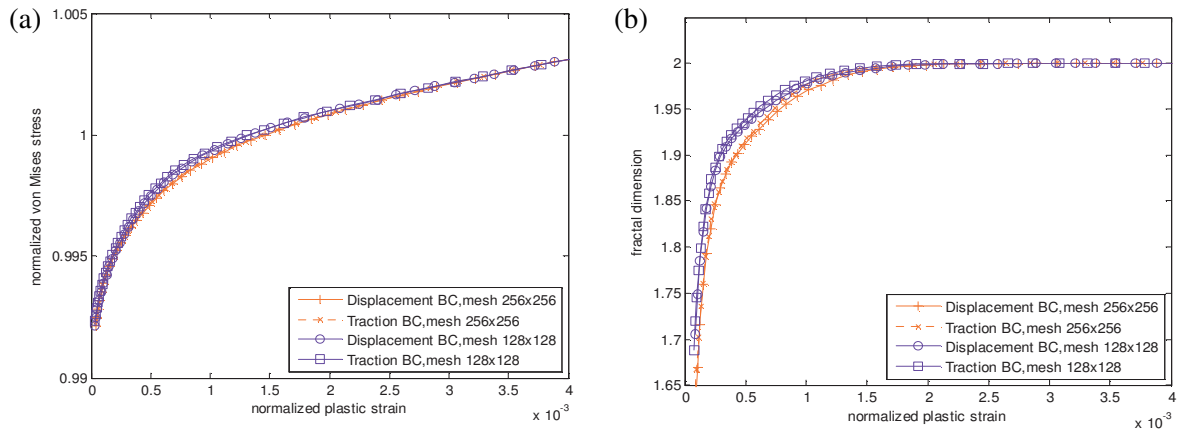
Figure 7. (a) Comparison of effective, normalized stress-strain curves in a 128×128 lattice with one grain = one finite element or one grain = 2×2 elements. (b) A corresponding comparison for the fractal dimension versus the normalized plastic strain.

Given the computer limitations, we can only use a twice finer finite element mesh, thus using a 256×256 mesh to represent a 128×128 grain lattice. As shown in the resulting Figure 7, the stress-strain curves display a bit softer response, while the fractal dimension seems to be lower in bigger mesh. The first result is explained by noting that a finer mesh offers more degrees of freedom to the given grain microstructure, whereas the second observation is understood by noting that a larger mesh leads to the possibility of partial plasticity in one grain, now modeled by four finite elements as opposed to one element equal one grain which may be either fully elastic or plastic.

5. Conclusion

The key result of this study is that a nonfractal random field of thermal expansion coefficients results in the set of plastic grains growing as a fractal through the elastic-plastic transition, and gradually filling the entire material domain. Parallel to this, the set of elastic grains evolves as another fractal, gradually diminishing to a set of zero measure. These results are obtained under three types of monotonic loadings (either displacement, traction, or mixed-orthogonal) consistent with the Hill–Mandel condition. Notably, a gradual transition of the material from elastic type to plastic type, where plasticity spreads in a plane-filling fashion, is far more realistic than the idealized homogeneous medium model in which the transition is an immediate process, characterized by a kink in the stress-strain curve. With the fractal dimension parameter we explore transition patterns by varying material constants' means and/or types of randomness. In the particular thermoelastoplastic model studied here, when the effect of random thermal fluctuations increases, as represented by a higher coefficient of variation of $\alpha/(c/E)$, the elastic-plastic transition process becomes slower. Considering that the magnitude of plastic strain is reflected in the strength of slip-lines and shear bands, we see that even very weak material randomness in material parameters of elastic-ductile materials causes plastic slip-lines and shear bands to evolve as fractals. Given the need to simulate very large lattices so as to obtain as reliable estimates of fractal dimensions as possible, all the results have been obtained by representing each grain by one finite element; a twice finer resolution has been found to have very minor effects only. Finally, we argue that the fractal dimension D, being a much more easily observable parameter than the stress state, can conveniently be used to grasp the state of loading in any given material and to describe the transition patterns in a unified way for many different materials.

Acknowledgement

References

[Feder 1988] J. Feder, *Fractals*, Plenum Press, New York, 1988.

[Hazanov 1998] S. Hazanov, "Hill condition and overall properties of composites", *Arch. Appl. Mech.* **68**:6 (1998), 385–394.

[Li and Ostoja-Starzewski 2010a] J. Li and M. Ostoja-Starzewski, "Fractal pattern formation at elastic-plastic transition in heterogeneous materials", *J. Appl. Mech.* (*ASME*) **77**:2 (2010), 021005.

[Li and Ostoja-Starzewski 2010b] J. Li and M. Ostoja-Starzewski, "Fractals in elastic-hardening plastic materials", *Proc. R. Soc. A* **466**:2114 (2010), 603–621.

[Mandelbrot 1982] B. B. Mandelbrot, *The fractal geometry of nature*, W. H. Freeman, San Francisco, 1982.

[Ostoja-Starzewski 1990] M. Ostoja-Starzewski, "Micromechanics model of ice fields, II: Monte Carlo simulation", *Pure Appl. Geophys.* **133**:2 (1990), 229–249.

[Ostoja-Starzewski 2008] M. Ostoja-Starzewski, *Microstructural randomness and scaling in mechanics of materials*, Chapman & Hall/CRC, Boca Raton, FL, 2008.

[Overland et al. 1995] J. E. Overland, B. A. Walter, T. B. Curtin, and P. Turet, "Hierarchy and sea ice mechanics: a case study from the Beaufort sea", *J. Geophys. Res.* **100**:C3 (1995), 4559–4571.

[Poliakov et al. 1994] A. N. B. Poliakov, H. J. Herrmann, Y. Y. Podladchikov, and S. Roux, "Fractal plastic shear bands", *Fractals* **2**:4 (1994), 567–581.

[Sornette 2006] D. Sornette, *Critical phenomena in natural sciences. Chaos, fractals, selforganization and disorder: concepts and tools*, 2nd ed., Springer, Berlin, 2006.

JUN LI: junli3@illinois.edu
Department of Mechanical Science and Engineering, University of Illinois at Urbana–Champaign, 1206 W. Green Street, Urbana, IL 61801-2906, United States

MARTIN OSTOJA-STARZEWSKI: martinos@uiuc.edu
Department of Mechanical Science and Engineering, Institute for Condensed Matter Theory, and Beckman Institute, University of Illinois at Urbana–Champaign, Urbana, IL 61801, United States

RATE-TYPE ELASTICITY AND VISCOELASTICITY
OF AN ERYTHROCYTE MEMBRANE

Vlado A. Lubarda

The rate-type constitutive theory of elastic and viscoelastic response of an erythrocyte membrane is presented. The results are obtained for an arbitrary isotropic strain energy function, and for its particular Evans–Skalak form. The explicit representations of the corresponding fourth-order tensors of elastic moduli are derived, with respect to principal axes of stress, and an arbitrary set of orthogonal axes. The objective rate-type viscoelastic constitutive equations of the cell membrane are then derived, based on the Maxwell and Kelvin models of viscoelastic behavior.

1. Introduction

The study of the mechanical response of a red blood cell (erythrocyte) membrane has received a great amount of attention from both classical mechanics and bio-chemo-mechanical points of view [Evans and Skalak 1980; Fung 1993; Agre and Parker 1989; Boal 2002]. Recent studies have addressed the evolution of mechanical properties of the cell membrane, associated with dynamic remodeling and reorganization of the membrane structure during large deformations of the cell, in response to mechanical, thermal, and chemical forces [Gov 2007; Li et al. 2007; Park et al. 2010]. Once the evolution equations for such structural changes are constructed, their incorporation in the constitutive theory can be accomplished through the rate-type constitutive equations, analogous to the existing theories of biological remodeling or growth of soft tissues [Lubarda and Hoger 2002; Garikipati et al. 2006; Taber 2009]. Since the rate-type constitutive theories relate the rates of stress and deformation, an essential ingredient of such theories is the elastic constitutive expression in the rate-type (incremental) form. The objective of this paper is to develop such rate-type constitutive expression, corresponding to an arbitrary isotropic strain energy function, and its Evans–Skalak particular form. The explicit expressions are derived for the components of the fourth-order tensor of elastic moduli, which relate the objective rates of the conjugate stress and strain tensors. The derived rate-type elasticity equations are then used to construct the objective rate-type viscoelastic constitutive equations for erythrocyte membrane, based on the Maxwell and Kelvin viscoelastic models.

Since the red blood cell membrane is only about 100 Å thick, it is assumed that the model of continuum mechanics applies only within the plane of the membrane. The applied forces are thus considered to be distributed along the length, with the membrane stresses defined by the force/length ratios. The upper bound on the in-plane Poisson ratio is equal to 1, rather than $\frac{1}{2}$, as in the three-dimensional isotropic elasticity. When modeled as infinitesimally thin, the erythrocyte membrane has no buckling resistance, and thus can support only noncompressive loadings, which give rise to nonnegative principal stresses

[Lubarda and Marzani 2009; Lubarda 2010]. The incorporation of small, but finite bending stiffness (proportional to the square of the membrane thickness) is needed to address other features of the membrane response, such as the size and shape of membrane wrinkles which may form during large deformations [Haughton 2001; Géminard et al. 2004], the membrane adhesive interactions [Agrawal and Steigmann 2009], or the resting shape of the cell and the skeleton/bilayer interactions in optical tweezer stretching and micropipette aspiration tests [Peng et al. 2010].

The content of the paper is as follows. Section 2 contains the kinematics and kinetics of the membrane in-plane deformation. The Evans–Skalak form of the elastic strain energy is adopted, as in most other recent work on the mechanics of red blood cell, (e.g., [Mills et al. 2004; Zhu et al. 2007]). The rate-type elasticity equations, which relate the Jaumann rate of the Kirchhoff stress to the rate of deformation tensor, are derived in Section 3. The explicit representation of the elastic moduli tensor is given for an arbitrary isotropic strain energy function, and for its special Evans–Skalak form. The results are given for the coordinate axes parallel to the principal axes of stress, and for the arbitrary set of orthogonal axes. Since the areal modulus of the erythrocyte cell is several orders of magnitude higher than the shear modulus, for some applications it may be appropriate to model the membrane as infinitely stiff to its area change; in Section 4 we accordingly present the rate-type elastic constitutive analysis for isoareal membranes. The derived rate-type elasticity equations are used in Section 5 to derive the objective rate-type viscoelasticity equations, corresponding to the Maxwell and Kelvin viscoelastic models. The results are given explicitly for an arbitrary isotropic strain energy function and for its particular Evans–Skalak form. Concluding remarks are given in Section 6. For completeness of the analysis, the appendices contain the results of the rate-type elastic analysis based on the Lagrangian strain and its conjugate Piola–Kirchhoff stress (Appendix A), and the deformation gradient and its conjugate nominal stress (Appendix B).

2. Kinematic and kinetic preliminaries

The deformation gradient associated with in-plane deformation of a thin plane membrane is

$$\boldsymbol{F} = \lambda_1 \, \boldsymbol{n}_1 \otimes \boldsymbol{N}_1 + \lambda_2 \, \boldsymbol{n}_2 \otimes \boldsymbol{N}_2, \tag{1}$$

where λ_1 and λ_2 are the principal stretches, $\boldsymbol{N}_i$ are the unit vectors along the principal directions of the right stretch tensor $\boldsymbol{U}$ in the undeformed configuration, and $\boldsymbol{n}_i = \boldsymbol{R} \cdot \boldsymbol{N}_i$ are the unit vectors in the deformed configuration, along the principal directions of the left stretch tensor $\boldsymbol{V}$. The polar decomposition of the deformation gradient is $\boldsymbol{F} = \boldsymbol{V} \cdot \boldsymbol{R} = \boldsymbol{R} \cdot \boldsymbol{U}$, where $\boldsymbol{R} = \boldsymbol{n}_1 \otimes \boldsymbol{N}_1 + \boldsymbol{n}_2 \otimes \boldsymbol{N}_2$ is the rotation tensor. The dyadic product of two vectors is denoted by $\otimes$.

The strain energy (per unit initial area) of an isotropic membrane is a function of the principal stretches, $\phi = \phi(\lambda_1, \lambda_2)$. An infinitesimal change of the strain energy is equal to the work done by the true (Cauchy) stress components σ_1 and σ_2 on the incremental stretches $d\lambda_1$ and $d\lambda_2$, which gives

$$d\phi = \sigma_1 \lambda_2 \, d\lambda_1 + \sigma_2 \lambda_1 \, d\lambda_2. \tag{2}$$

By taking the differential of $\phi = \phi(\lambda_1, \lambda_2)$, the comparison with (2) establishes the constitutive expressions

$$\sigma_1 = \frac{1}{\lambda_2} \frac{\partial \phi}{\partial \lambda_1}, \quad \sigma_2 = \frac{1}{\lambda_1} \frac{\partial \phi}{\partial \lambda_2}. \tag{3}$$

The corresponding Kirchhoff stress components, $\tau_i = (\det \boldsymbol{F}) \sigma_i = (\lambda_1 \lambda_2) \sigma_i$, are

$$\tau_1 = \lambda_1 \frac{\partial \phi}{\partial \lambda_1}, \quad \tau_2 = \lambda_2 \frac{\partial \phi}{\partial \lambda_2}. \tag{4}$$

2.1. *Evans–Skalak form of the strain energy.* The strain energy of an isotropic membrane must be a symmetric function of λ_1 and λ_2. Evans and Skalak [1980] proposed that $\phi = \phi(\alpha, \beta)$, where the deformation measures

$$\alpha = \lambda_1 \lambda_2 - 1, \quad \beta = \frac{1}{2} \left(\frac{\lambda_1}{\lambda_2} + \frac{\lambda_2}{\lambda_1} \right) - 1 \tag{5}$$

are constructed so that $\alpha = \beta = 0$ in the undeformed configuration. The parameter α is the area change per unit undeformed area, while β is a measure of the distortional deformation. The Evans–Skalak model of a nonlinear elastic membrane corresponds to $\phi = \frac{1}{2} \kappa \alpha^2 + \mu \beta$, i.e.,

$$\phi = \frac{1}{2} \kappa (\lambda_1 \lambda_2 - 1)^2 + \mu \left[\frac{1}{2} \left(\frac{\lambda_1}{\lambda_2} + \frac{\lambda_2}{\lambda_1} \right) - 1 \right], \tag{6}$$

where κ and μ are the areal (bulk) modulus and the shear modulus of the membrane, respectively. From (3), the stress components are

$$\sigma_{1,2} = \kappa (\lambda_1 \lambda_2 - 1) \pm \frac{1}{2} \mu (\lambda_2^{-2} - \lambda_1^{-2}), \tag{7}$$

$$\sigma_1 + \sigma_2 = 2\kappa (\lambda_1 \lambda_2 - 1), \quad \sigma_1 - \sigma_2 = \mu (\lambda_2^{-2} - \lambda_1^{-2}). \tag{8}$$

Numerous experiments have been conducted in the past to measure the elastic and viscous properties of the human erythrocyte membrane. For example, from the micropipette aspiration tests of red blood cells, it has been estimated that the shear modulus (μ) of the erythrocyte membrane is in the range $4 - 10$ μN/m, while the areal modulus (κ) is on the order $10^3 - 10^4$ higher than that [Evans and Skalak 1980; Boal 2002]. Such a large difference in two stiffnesses is because the areal modulus of the cell membrane is controlled mostly by the phospholipidic bilayer, while the shear modulus is determined by the elastic properties of the cytoskeleton, a two-dimensional network of spectrin strands bound to the bilayer. The viscous properties of the cell are due to glycoproteins, lipid integral and peripheral membrane proteins, lipid rafts, and transmembrane cholesterol. Based on the measurements of the characteristic time for relaxation, the shear viscosity has been estimated to be in the range 0.6–2.7 μN·s/m [Hochmuth 1987].

If the Kirchhoff stress components are used, the expressions (7)–(8) can be rewritten as

$$\tau_1 + \tau_2 = 2\kappa \lambda_1 \lambda_2 (\lambda_1 \lambda_2 - 1), \quad \tau_1 - \tau_2 = \mu \left(\frac{\lambda_1}{\lambda_2} - \frac{\lambda_2}{\lambda_1} \right). \tag{9}$$

For the later purposes of the rate-type theory of viscoelasticity and the incremental solution of the boundary value problems, it is important to invert (9), and express $\lambda_1 \lambda_2$ and λ_1 / λ_2 in terms of stress components. The resulting expressions are

$$\lambda_1 \lambda_2 = \frac{1}{2} \left[1 + \left(1 + 2 \frac{\tau_1 + \tau_2}{\kappa} \right)^{1/2} \right],$$

$$\frac{\lambda_1}{\lambda_2} = \frac{\tau_1 - \tau_2}{2\mu} + \left[1 + \left(\frac{\tau_1 - \tau_2}{2\mu} \right)^2 \right]^{1/2}, \quad \frac{\lambda_2}{\lambda_1} = -\frac{\tau_1 - \tau_2}{2\mu} + \left[1 + \left(\frac{\tau_1 - \tau_2}{2\mu} \right)^2 \right]^{1/2}. \tag{10}$$

The sum and difference of the Kirchhoff stress components can be expressed in terms of the Cauchy stress components as

$$\tau_1 \pm \tau_2 = \left(1 + \frac{\sigma_1 + \sigma_2}{2\kappa}\right)(\sigma_1 \pm \sigma_2), \tag{11}$$

because, from (9), $\lambda_1\lambda_2 = 1 + (\sigma_1 + \sigma_2)/2\kappa$, and $\tau_i = (\lambda_1\lambda_2)\sigma_i$, for $i = 1, 2$.

2.2. *Stress response on arbitrary axes.*

The strain energy of an isotropic membrane can be expressed as a function of two independent invariants of the left Cauchy–Green deformation tensor $\boldsymbol{B} = \boldsymbol{V}^2$, i.e.,

$$\phi = \phi(I_B, II_B), \quad I_B = \operatorname{tr}\boldsymbol{B} = \lambda_1^2 + \lambda_2^2, \quad II_B = \det\boldsymbol{B} = \lambda_1^2\lambda_2^2. \tag{12}$$

where I_B is the trace, and II_B the determinant of $\boldsymbol{B}$. Since $\dot{I}_B = \dot{B}_{kk} = 2B_{ij}D_{ij}$, $\dot{II}_B = 2II_B D_{kk}$, and $\dot\phi = \tau_{ij}D_{ij}$, by taking the time-rate of (12), if follows that

$$\tau_{ij} = 2\left(\frac{\partial\phi}{\partial I_B}B_{ij} + II_B\frac{\partial\phi}{\partial II_B}\delta_{ij}\right), \tag{13}$$

where δ_{ij} is the Kronecker delta. The invariants I_B and II_B are related to the invariants α and β by

$$\alpha = II_B^{1/2} - 1, \quad \beta = \frac{I_B}{2II_B^{1/2}} - 1; \quad I_B = 2(1+\alpha)(1+\beta), \quad II_B = (1+\alpha)^2. \tag{14}$$

Consequently,

$$\frac{\partial\phi}{\partial I_B} = \frac{1}{2}II_B^{-1/2}\frac{\partial\phi}{\partial\beta}, \quad \frac{\partial\phi}{\partial II_B} = \frac{1}{4}II_B^{-1/2}\left(2\frac{\partial\phi}{\partial\beta} - I_B II_B^{-1}\frac{\partial\phi}{\partial\beta}\right), \tag{15}$$

and the substitution into (13) gives an alternative representation of the stress response

$$\sigma_{ij} = \frac{\partial\phi}{\partial\alpha}\delta_{ij} + \frac{1}{\det\boldsymbol{B}}\frac{\partial\phi}{\partial\beta}B'_{ij}, \tag{16}$$

where $B'_{ij} = B_{ij} - B_{kk}\delta_{ij}/2$ is the deviatoric part of B_{ij}.

If ϕ is given by the Evans–Skalak form $\phi = \frac{1}{2}\kappa\alpha^2 + \mu\beta$, (16) reduces to

$$\sigma_{ij} = \kappa\alpha\,\delta_{ij} + \frac{\mu}{\det\boldsymbol{B}}B'_{ij}. \tag{17}$$

The first term on the right-hand side of (16) and (17) is the spherical part, $\sigma_{kk}\delta_{ij}/2$, and the second term is the deviatoric part of the stress, σ'_{ij}. If the coordinate axes are parallel to the principal axes of $\boldsymbol{B}$, (17) reduces to (7).

3. Rate-type constitutive analysis

In this section we derive the rate-type form of the elastic constitutive equations, expressed with respect to the principal axes of stress and the constitutive structure (4), and with respect to arbitrary set of orthogonal axes and the constitutive structure (13) or (16).

3.1. *Rate of deformation and spin tensors.* For the rate-type constitutive analysis, the rates of deformation need to be considered. To that goal, we first introduce the spin tensors

$$\boldsymbol{\Omega} = \Omega(\boldsymbol{N}_1 \otimes \boldsymbol{N}_2 - \boldsymbol{N}_2 \otimes \boldsymbol{N}_1), \quad \boldsymbol{\omega} = \omega(\boldsymbol{n}_1 \otimes \boldsymbol{n}_2 - \boldsymbol{n}_2 \otimes \boldsymbol{n}_1), \tag{18}$$

such that $\dot{\boldsymbol{N}}_i = \boldsymbol{\Omega} \cdot \boldsymbol{N}_i$ and $\dot{\boldsymbol{n}}_i = \boldsymbol{\omega} \cdot \boldsymbol{n}_i$ $(i = 1, 2)$. In particular,

$$\dot{\boldsymbol{N}}_1 = -\Omega \boldsymbol{N}_2, \quad \dot{\boldsymbol{N}}_2 = \Omega \boldsymbol{N}_1, \quad \dot{\boldsymbol{n}}_1 = -\omega \boldsymbol{n}_2, \quad \dot{\boldsymbol{n}}_2 = \omega \boldsymbol{n}_1, \tag{19}$$

where $-\Omega$ is the rate of the counterclockwise rotation of the dyad $(\boldsymbol{N}_1, \boldsymbol{N}_2)$, while $-\omega$ is the rate of the counterclockwise rotation of the dyad $(\boldsymbol{n}_1, \boldsymbol{n}_2)$. Consequently, by differentiating (1), and by using (19), the rate of the deformation gradient becomes

$$\dot{\boldsymbol{F}} = \dot{\lambda}_1 \boldsymbol{n}_1 \otimes \boldsymbol{N}_1 + \dot{\lambda}_2 \boldsymbol{n}_2 \otimes \boldsymbol{N}_2 + (\lambda_2 \omega - \lambda_1 \Omega) \boldsymbol{n}_1 \otimes \boldsymbol{N}_2 - (\lambda_1 \omega - \lambda_2 \Omega) \boldsymbol{n}_2 \otimes \boldsymbol{N}_1. \tag{20}$$

Since the inverse of the deformation gradient is

$$\boldsymbol{F}^{-1} = \frac{1}{\lambda_1} \boldsymbol{N}_1 \otimes \boldsymbol{n}_1 + \frac{1}{\lambda_2} \boldsymbol{N}_2 \otimes \boldsymbol{n}_2, \tag{21}$$

the substitution of (20) and (21) into the expression for the velocity gradient $\boldsymbol{L} = \dot{\boldsymbol{F}} \cdot \boldsymbol{F}^{-1}$ gives

$$\boldsymbol{L} = \frac{\dot{\lambda}_1}{\lambda_1} \boldsymbol{n}_1 \otimes \boldsymbol{n}_1 + \frac{\dot{\lambda}_2}{\lambda_2} \boldsymbol{n}_2 \otimes \boldsymbol{n}_2 + \omega(\boldsymbol{n}_1 \otimes \boldsymbol{n}_2 - \boldsymbol{n}_2 \otimes \boldsymbol{n}_1) - \Omega\left(\frac{\lambda_1}{\lambda_2} \boldsymbol{n}_1 \otimes \boldsymbol{n}_2 - \frac{\lambda_2}{\lambda_1} \boldsymbol{n}_2 \otimes \boldsymbol{n}_1\right). \tag{22}$$

Its symmetric and antisymmetric parts are the rate of deformation tensor

$$\boldsymbol{D} = \frac{\dot{\lambda}_1}{\lambda_1} \boldsymbol{n}_1 \otimes \boldsymbol{n}_1 + \frac{\dot{\lambda}_2}{\lambda_2} \boldsymbol{n}_2 \otimes \boldsymbol{n}_2 + \Omega \frac{\lambda_2^2 - \lambda_1^2}{2\lambda_1\lambda_2}(\boldsymbol{n}_1 \otimes \boldsymbol{n}_2 + \boldsymbol{n}_2 \otimes \boldsymbol{n}_1), \tag{23}$$

and the spin tensor

$$\boldsymbol{W} = \left(\omega - \Omega \frac{\lambda_1^2 + \lambda_2^2}{2\lambda_1\lambda_2}\right)(\boldsymbol{n}_1 \otimes \boldsymbol{n}_2 - \boldsymbol{n}_2 \otimes \boldsymbol{n}_1). \tag{24}$$

The components of the rate of deformation and spin tensors, on the current axes $\boldsymbol{n}_i$, are

$$D_{11} = \frac{\dot{\lambda}_1}{\lambda_1}, \quad D_{22} = \frac{\dot{\lambda}_2}{\lambda_2}, \quad D_{12} = \Omega \frac{\lambda_2^2 - \lambda_1^2}{2\lambda_1\lambda_2} \tag{25}$$

and

$$W = \omega - \Omega \frac{\lambda_1^2 + \lambda_2^2}{2\lambda_1\lambda_2}, \tag{26}$$

so that

$$\boldsymbol{D} = \sum_{i,j=1}^{2} = D_{ij} \boldsymbol{n}_i \otimes \boldsymbol{n}_j, \quad \boldsymbol{W} = W(\boldsymbol{n}_1 \otimes \boldsymbol{n}_2 - \boldsymbol{n}_2 \otimes \boldsymbol{n}_1). \tag{27}$$

3.2. *Rate-type constitutive equations.* A comprehensive treatment of the three-dimensional rate-type (incremental) elasticity can be found in Haughton and Ogden [1978] and Ogden [1984]. Some of the results in this section can be deduced from this theory directly, but some require a separate analysis. The rate of the Kirchhoff stress $\boldsymbol{\tau} = \tau_1 \boldsymbol{n}_1 \otimes \boldsymbol{n}_1 + \tau_2 \boldsymbol{n}_2 \otimes \boldsymbol{n}_2$ is

$$\dot{\boldsymbol{\tau}} = \dot{\tau}_1 \boldsymbol{n}_1 \otimes \boldsymbol{n}_1 + \dot{\tau}_2 \boldsymbol{n}_2 \otimes \boldsymbol{n}_2 - \omega(\tau_1 - \tau_2)(\boldsymbol{n}_1 \otimes \boldsymbol{n}_2 + \boldsymbol{n}_2 \otimes \boldsymbol{n}_1), \tag{28}$$

from which we conclude that

$$\dot{\tau}_{12} = \dot{\tau}_{21} = -\omega(\tau_1 - \tau_2).$$ (29)

Observing, from (25) and (26), that

$$\omega = W - \frac{\lambda_1^2 + \lambda_2^2}{\lambda_1^2 - \lambda_2^2} D_{12},$$ (30)

the expression (29) can be rewritten as

$$\overset{\circ}{\tau}_{12} = (\tau_1 - \tau_2) \frac{\lambda_1^2 + \lambda_2^2}{\lambda_1^2 - \lambda_2^2} D_{12}.$$ (31)

The Jaumann rate $\overset{\circ}{\boldsymbol{\tau}}$ is defined by $\overset{\circ}{\boldsymbol{\tau}} = \dot{\boldsymbol{\tau}} - \boldsymbol{W} \cdot \boldsymbol{\tau} + \boldsymbol{\tau} \cdot \boldsymbol{W}$. Since, instantaneously, on the axes $\boldsymbol{n}_i$, the shear stress $\tau_{12} = 0$, the components of the Jaumann rate of the Kirchhoff stress are

$$\overset{\circ}{\tau}_{11} = \dot{\tau}_{11}, \quad \overset{\circ}{\tau}_{22} = \dot{\tau}_{22}, \quad \overset{\circ}{\tau}_{12} = \dot{\tau}_{12} + W(\tau_1 - \tau_2),$$ (32)

which was used in the transition from (29) to (31).

Having regard to constitutive expressions (4), the rates of the Kirchhoff stress components are

$$\dot{\tau}_1 = \left(\frac{\partial \phi}{\partial \lambda_1} + \lambda_1 \frac{\partial^2 \phi}{\partial \lambda_1^2}\right)\dot{\lambda}_1 + \left(\lambda_1 \frac{\partial^2 \phi}{\partial \lambda_1 \partial \lambda_2}\right)\dot{\lambda}_2, \quad \dot{\tau}_2 = \left(\lambda_2 \frac{\partial^2 \phi}{\partial \lambda_1 \partial \lambda_2}\right)\dot{\lambda}_1 + \left(\frac{\partial \phi}{\partial \lambda_2} + \lambda_2 \frac{\partial^2 \phi}{\partial \lambda_2^2}\right)\dot{\lambda}_2.$$ (33)

Consequently, by using (25), the objective rate-type constitutive expressions for elastic deformations of a thin membrane are

$$\overset{\circ}{\tau}_{11} = \lambda_1 \left(\frac{\partial \phi}{\partial \lambda_1} + \lambda_1 \frac{\partial^2 \phi}{\partial \lambda_1^2}\right) D_{11} + \left(\lambda_1 \lambda_2 \frac{\partial^2 \phi}{\partial \lambda_1 \partial \lambda_2}\right) D_{22},$$

$$\overset{\circ}{\tau}_{22} = \left(\lambda_1 \lambda_2 \frac{\partial^2 \phi}{\partial \lambda_1 \partial \lambda_2}\right) D_{11} + \lambda_2 \left(\frac{\partial \phi}{\partial \lambda_2} + \lambda_2 \frac{\partial^2 \phi}{\partial \lambda_2^2}\right) D_{22},$$ (34)

$$\overset{\circ}{\tau}_{12} = \left(\lambda_1 \frac{\partial \phi}{\partial \lambda_1} - \lambda_2 \frac{\partial \phi}{\partial \lambda_2}\right) \frac{\lambda_1^2 + \lambda_2^2}{\lambda_1^2 - \lambda_2^2} D_{12}.$$

3.3. *Elastic moduli tensor.* The tensor representation of the constitutive expressions (34) is

$$\overset{\circ}{\boldsymbol{\tau}} = \boldsymbol{\mathcal{L}} : \boldsymbol{D}, \quad \overset{\circ}{\tau}_{ij} = \mathcal{L}_{ijkl} D_{kl},$$ (35)

where the components of the fourth-order tensor of the elastic moduli

$$\boldsymbol{\mathcal{L}} = \mathcal{L}_{1111} \, \boldsymbol{n}_1 \otimes \boldsymbol{n}_1 \otimes \boldsymbol{n}_1 \otimes \boldsymbol{n}_1 + \mathcal{L}_{2222} \, \boldsymbol{n}_2 \otimes \boldsymbol{n}_2 \otimes \boldsymbol{n}_2 \otimes \boldsymbol{n}_2$$

$$+ \mathcal{L}_{1122} \, (\boldsymbol{n}_1 \otimes \boldsymbol{n}_1 \otimes \boldsymbol{n}_2 \otimes \boldsymbol{n}_2 + \boldsymbol{n}_2 \otimes \boldsymbol{n}_2 \otimes \boldsymbol{n}_1 \otimes \boldsymbol{n}_1)$$

$$+ \mathcal{L}_{1212} \, (\boldsymbol{n}_1 \otimes \boldsymbol{n}_2 + \boldsymbol{n}_2 \otimes \boldsymbol{n}_1) \otimes (\boldsymbol{n}_1 \otimes \boldsymbol{n}_2 + \boldsymbol{n}_2 \otimes \boldsymbol{n}_1)$$ (36)

are

$$\mathcal{L}_{1111} = \lambda_1 \left(\frac{\partial \phi}{\partial \lambda_1} + \lambda_1 \frac{\partial^2 \phi}{\partial \lambda_1^2} \right), \quad \mathcal{L}_{2222} = \lambda_2 \left(\frac{\partial \phi}{\partial \lambda_2} + \lambda_2 \frac{\partial^2 \phi}{\partial \lambda_2^2} \right),$$

$$\mathcal{L}_{1122} = \lambda_1 \lambda_2 \frac{\partial^2 \phi}{\partial \lambda_1 \partial \lambda_2}, \quad \mathcal{L}_{1212} = \frac{1}{2} \left(\lambda_1 \frac{\partial \phi}{\partial \lambda_1} - \lambda_2 \frac{\partial \phi}{\partial \lambda_2} \right) \frac{\lambda_1^2 + \lambda_2^2}{\lambda_1^2 - \lambda_2^2}, \tag{37}$$

with the obvious symmetry properties

$$\mathcal{L}_{1122} = \mathcal{L}_{2211}, \quad \mathcal{L}_{1212} = \mathcal{L}_{2121} = \mathcal{L}_{1221} = \mathcal{L}_{2112}. \tag{38}$$

The components of the tensor $\mathcal{L}$, as well as the components of the tensors $\overset{\circ}{\tau}$ and $\boldsymbol{D}$, relative to the fixed base vectors, can be easily obtained by the tensor transformation rules. An alternative, direct derivation is presented in section 3.4.

If the strain energy is given by (6), the elastic moduli become

$$\mathcal{L}_{1111} = \mathcal{L}_{2222} = \kappa \lambda_1 \lambda_2 (2\lambda_1 \lambda_2 - 1) + \mu \frac{\lambda_1^2 + \lambda_2^2}{2\lambda_1 \lambda_2},$$

$$\mathcal{L}_{1122} = \kappa \lambda_1 \lambda_2 (2\lambda_1 \lambda_2 - 1) - \mu \frac{\lambda_1^2 + \lambda_2^2}{2\lambda_1 \lambda_2}, \quad \mathcal{L}_{1212} = \mu \frac{\lambda_1^2 + \lambda_2^2}{2\lambda_1 \lambda_2}. \tag{39}$$

3.4. *Elastic compliances tensor.* The inverted form of (35), giving the rate of deformation tensor in terms of the Jaumann rate of the Kirchhoff stress, is

$$D_{ij} = \mathcal{M}_{ijkl} \overset{\circ}{\tau}_{kl}. \tag{40}$$

The nonvanishing components of the elastic compliances tensor are

$$\mathcal{M}_{1111} = \frac{\mathcal{L}_{2222}}{\Delta}, \quad \mathcal{M}_{2222} = \frac{\mathcal{L}_{1111}}{\Delta}, \quad \mathcal{M}_{1122} = \mathcal{M}_{2211} = -\frac{\mathcal{L}_{1122}}{\Delta}, \quad \mathcal{M}_{1212} = \frac{1}{4\mathcal{L}_{1212}}, \tag{41}$$

where

$$\Delta = \mathcal{L}_{1111} \mathcal{L}_{2222} - \mathcal{L}_{1212}^2. \tag{42}$$

If the strain energy is given by the Evans–Skalak form (6), the determinant Δ is

$$\Delta = 2\kappa \mu (2\lambda_1 \lambda_2 - 1)(\lambda_1^2 + \lambda_2^2), \tag{43}$$

while the elastic moduli $\mathcal{L}_{ijkl}$ (on the principal axes of stress) are given by (39).

3.5. *Rate-type elasticity on arbitrary axes.* The objective rate form of the constitutive expression (13) can be obtained directly by applying to it the Jaumann derivative. Since $\overset{\circ}{\boldsymbol{B}} = \boldsymbol{B} \cdot \boldsymbol{D} + \boldsymbol{D} \cdot \boldsymbol{B}$, $\dot{I}_B = 2\boldsymbol{B} : \boldsymbol{D}$, and $\dot{II}_B = 2II_B \operatorname{tr} \boldsymbol{D}$, it readily follows that

$$\overset{\circ}{\tau}_{ij} = \mathcal{L}_{ijkl} D_{kl}, \tag{44}$$

where

$$\mathcal{L}_{ijkl} = c_1 \delta_{ij} \delta_{kl} + c_2 (\delta_{ij} B_{kl} + B_{ij} \delta_{kl}) + c_3 B_{ij} B_{kl} + c_4 (\delta_{ik} B_{jl} + B_{ik} \delta_{jl} + \delta_{jk} B_{il} + B_{jk} \delta_{il}) \tag{45}$$

are the components of the fourth-order tensor of elastic moduli with respect to arbitrary orthogonal axes, and

$$c_1 = 4II_B\left(\frac{\partial\phi}{\partial II_B} + II_B\frac{\partial^2\phi}{\partial II_B^2}\right), \quad c_2 = 4II_B\frac{\partial^2\phi}{\partial I_B\partial II_B}, \quad c_3 = 4\frac{\partial^2\phi}{\partial I_B^2}, \quad c_4 = \frac{\partial\phi}{\partial I_B}. \quad (46)$$

If the strain energy is expressed in terms of the invariants α and β, rather than I_B and II_B, the parameters c_i become

$$c_1 = (1+\alpha)^2\frac{\partial^2\phi}{\partial\alpha^2} + (1+\beta)^2\frac{\partial^2\phi}{\partial\beta^2} - 2(1+\alpha)(1+\beta)\frac{\partial^2\phi}{\partial\alpha\partial\beta} + (1+\alpha)\frac{\partial\phi}{\partial\alpha} + (1+\beta)\frac{\partial\phi}{\partial\beta},$$

$$c_2 = \frac{\partial^2\phi}{\partial\alpha\partial\beta} - \frac{1}{1+\alpha}\left[(1+\beta)\frac{\partial^2\phi}{\partial\beta^2} + \frac{\partial\phi}{\partial\beta}\right], \quad (47)$$

$$c_3 = \frac{1}{(1+\alpha)^2}\frac{\partial^2\phi}{\partial\beta^2}, \quad c_4 = \frac{1}{2(1+\alpha)}\frac{\partial\phi}{\partial\beta}.$$

In the case of the Evans–Skalak strain energy, this simplifies to

$$c_1 = \kappa(1+\alpha)(1+2\alpha) + \mu(1+\beta), \quad c_2 = -\mu(1+\alpha)^{-1}, \quad c_3 = 0, \quad c_4 = -c_2/2. \quad (48)$$

If the coordinate axes are parallel to the principal axes of $\boldsymbol{B}$, (45) with (48) reduces to (39). In this case, $\mathcal{L}_{1111} = \mathcal{L}_{2222} = c_1$, $\mathcal{L}_{1122} = c_1 - 2\mathcal{L}_{1212}$, and $\mathcal{L}_{1212} = \mu(1+\beta)$.

4. Isoareal membranes

As discussed in the introduction, the areal modulus of the red blood cell is several orders of magnitude higher than the shear modulus, and for some applications it may be appropriate to model the membrane as infinitely stiff to the area change ($\kappa \to \infty$, $\nu = 1$). In this case, there is a deformation constraint $\lambda_1\lambda_2 - 1 = 0$, so that $D_{11} + D_{22} = 0$. The rate of work is $\dot{\phi} = (\sigma_1 - \sigma_2)\dot{\lambda}_1/\lambda_1$, and by differentiation $\phi = \phi(\lambda_1)$ with respect to time, there follows

$$\sigma_1 - \sigma_2 = \lambda_1\frac{d\phi}{d\lambda_1}. \quad (49)$$

The average normal stress is undetermined by the constitutive analysis, and denoting it by $-p_0$, we can write

$$\sigma_1 + \sigma_2 = -2p_0. \quad (50)$$

Therefore, the principal stresses are

$$\sigma_1 = \frac{1}{2}\lambda_1\frac{d\phi}{d\lambda_1} - p_0, \quad \sigma_2 = -\frac{1}{2}\lambda_1\frac{d\phi}{d\lambda_1} - p_0. \quad (51)$$

The function $p_0 = p_0(x_1, x_2)$ is determined by solving a specific boundary-value problem under consideration. If the strain energy is

$$\phi = \mu\beta = \mu\left(\tfrac{1}{2}(\lambda_1^2 + \lambda_1^{-2}) - 1\right), \quad (52)$$

the stresses become

$$\sigma_1 = \tfrac{1}{2}\mu(\lambda_1^2 - \lambda_1^{-2}) - p_0, \quad \sigma_2 = -\tfrac{1}{2}\mu(\lambda_1^2 - \lambda_1^{-2}) - p_0. \quad (53)$$

4.1. *Rate theory for isoareal membranes.* For isoareal membranes the Cauchy and Kirchhoff stress are equal to each other, so that (31) becomes

$$\overset{\circ}{\sigma}_{12} = \lambda_1 \frac{d\phi}{d\lambda_1} \frac{\lambda_1^2 + \lambda_1^{-2}}{\lambda_1^2 - \lambda_1^{-2}} D_{12}. \tag{54}$$

Furthermore, by differentiating (50) and (51),

$$\overset{\circ}{\sigma}_{11} + \overset{\circ}{\sigma}_{22} = -2\dot{p}_0, \quad \overset{\circ}{\sigma}_{11} - \overset{\circ}{\sigma}_{22} = \left(\lambda_1 \frac{d\phi}{d\lambda_1} + \lambda_1^2 \frac{d^2\phi}{d\lambda_1^2} \right) D_{11}. \tag{55}$$

If ϕ is given by (52), the rate-type constitutive expressions, on the principal stress axes, become

$$\overset{\circ}{\sigma}_{ij} = -\dot{p}_0 \delta_{ij} + \mu(\lambda_1^2 + \lambda_1^{-2}) D_{ij}. \tag{56}$$

4.2. *Expressions on arbitrary axes.* The strain energy of an isotropic isoareal membrane is a function of only one strain of stretch invariant, e.g., $\phi = \phi(I_B)$. Its rate is

$$\dot{\phi} = 2 \frac{d\phi}{dI_B} B'_{ij} D_{ij}, \tag{57}$$

because $\dot{I}_B = 2B'_{ij} D_{ij}$, the rate of deformation being deviatoric, so that $B_{ij} D_{ij} = B'_{ij} D_{ij}$. Since the rate of work is $\dot{\phi} = \sigma_{ij} D_{ij} = \sigma'_{ij} D_{ij}$, again because $D_{kk} = 0$, the comparison with (57) gives

$$\sigma'_{ij} = 2 \frac{d\phi}{dI_B} B'_{ij}, \tag{58}$$

and thus the stress response is

$$\sigma_{ij} = -p_0 \delta_{ij} + 2 \frac{d\phi}{dI_B} B'_{ij}, \tag{59}$$

To derive the rate-type form of (59), it is convenient to first rewrite (59) as

$$\sigma_{ij} = -\left(p_0 + I_B \frac{d\phi}{dI_B} \right) \delta_{ij} + 2 \frac{d\phi}{dI_B} B_{ij}. \tag{60}$$

The application of the Jaumann derivative to (60), having regard to $\dot{I}_B = 2B_{ij} D_{ij}$ and $\overset{\circ}{B}_{ij} = B_{ik} D_{kj} + D_{ik} B_{kj}$, then yields

$$\overset{\circ}{\sigma}_{ij} = -\dot{p}_0 \,\delta_{ij} + \left[c_0 \delta_{ij} B_{kl} + c_3 B_{ij} B_{kl} + c_4 (\delta_{ik} B_{jl} + B_{ik} \delta_{jl} + \delta_{jk} B_{il} + B_{jk} \delta_{il}) \right] D_{kl}. \tag{61}$$

The parameters c_0, c_2, and c_4 are

$$c_0 = -2 \left(\frac{d\phi}{dI_B} + I_B \frac{d^2\phi}{dI_B^2} \right), \quad c_3 = 4 \frac{d^2\phi}{dI_B^2}, \quad c_4 = \frac{d\phi}{dI_B}. \tag{62}$$

If the strain energy is of the Evans–Skalak form

$$\phi = \mu\beta = \mu\left(\tfrac{1}{2} I_B - 1 \right), \tag{63}$$

the parameters (62) reduce to $c_0 = -\mu$, $c_3 = 0$, and $c_4 = \mu/2$, while (61) simplifies to

$$\overset{\circ}{\sigma}_{ij} = -\dot{p}_0 \,\delta_{ij} + \tfrac{1}{2} \mu \left(\delta_{ik} B_{jl} + B_{ik} \delta_{jl} + \delta_{jk} B_{il} + B_{jk} \delta_{il} - 2\delta_{ij} B_{kl} \right) D_{kl}. \tag{64}$$

When coordinate axes coincident with the principal directions of $\boldsymbol{B}$ are used, (64) reduces to (56).

5. Rate theory of viscoelasticity

To study the time-dependent aspiration of an erythrocyte cell membrane into a micropipette [Evans and Hochmuth 1976], or the cell recovery upon its deformation by optical tweezers [Dao et al. 2003; Mills et al. 2004], viscoelastic constitutive models have been employed. In this section, we extend the constitutive models their considered by constructing the objective rate type viscoelastic constitutive equations of both Maxwell and Kelvin type, based on an arbitrary isotropic strain energy function, and its special Evans–Skalak form. For an integral type of viscoelastic constitutive equations for thin membranes, we refer to [Tözeren et al. 1982; Wineman 2007].

5.1. *Maxwell viscoelastic model.* In the Maxwell viscoelastic model, the rate of deformation is assumed to be the sum of the elastic and viscous parts, i.e.,

$$D_{ij} = D_{ij}^{\mathrm{e}} + D_{ij}^{\mathrm{v}}. \tag{65}$$

The elastic part is related to the objective Jaumann rate of the Kirchhoff stress by the constitutive expression (36), in which $\phi = \phi^{\mathrm{e}}(I_B^{\mathrm{e}}, II_B^{\mathrm{e}})$, so that

$$\overset{\circ}{\tau}_{ij} = \mathcal{L}_{ijkl}^{\mathrm{e}} D_{kl}^{\mathrm{e}}, \quad D_{ij}^{\mathrm{e}} = \mathcal{M}_{ijkl}^{\mathrm{e}} \overset{\circ}{\tau}_{kl}. \tag{66}$$

The viscous part is assumed to be deviatoric and governed by the Newton viscosity law

$$D_{ij}^{\mathrm{v}} = \frac{1}{2\eta} \sigma_{ij}', \quad \sigma_{ij}' = \sigma_{ij} - \frac{1}{2} \sigma_{kk} \delta_{ij}. \tag{67}$$

The coefficient of membrane viscosity η is assumed to be constant, although it could vary with the amount of stretching [Evans and Hochmuth 1976]. The experimentally reported values for η are in the range between 0.6 and 2.7 μN·s/m, being strongly influenced by the concentration of hemoglobin in the cytoplasm, which binds to the membrane of the cell [Hochmuth 1987; Hochmuth and Waugh 1987]. By substituting (66) and (67) into (65), there follows

$$\overset{\circ}{\tau}_{ij} = \mathcal{L}_{ijkl}^{\mathrm{e}} \left(D_{kl} - \frac{1}{2\eta} \sigma_{kl}' \right), \quad D_{ij} = \mathcal{M}_{ijkl}^{\mathrm{e}} \overset{\circ}{\tau}_{kl} + \frac{1}{2\eta} \sigma_{ij}'. \tag{68}$$

When expressed on the principal axes of the current stress, $\mathcal{L}_{ijkl}^{\mathrm{e}}$ is given by (37), and (68) becomes

$$\overset{\circ}{\tau}_{ij} = \mathcal{L}_{ijkl}^{\mathrm{e}} D_{kl} - \frac{1}{4\eta} (\mathcal{L}_{ij11}^{\mathrm{e}} - \mathcal{L}_{ij22}^{\mathrm{e}})(\sigma_1 - \sigma_2). \tag{69}$$

If the elastic response is governed by the Evans–Skalak form of the elastic strain energy, the elastic moduli $\mathcal{L}_{ijkl}^{\mathrm{e}}$ are given by (39), in which λ_1 and λ_2 are replaced by λ_1^{e} and λ_2^{e}. Since the viscous part of the rate of deformation is assumed to be deviatoric, the membrane area change is entirely due to elastic

deformation, so $\lambda_1 \lambda_2 = \lambda_1^e \lambda_2^e$. Consequently, from Equations (10)–(11),

$$
\lambda_1^e \lambda_2^e = \frac{1}{2} \left[1 + \left(1 + 2 \frac{\tau_1 + \tau_2}{\kappa} \right)^{1/2} \right] = 1 + \frac{\sigma_1 + \sigma_2}{2\kappa},
$$

$$
\frac{\lambda_1^e}{\lambda_2^e} = \frac{\tau_1 - \tau_2}{2\mu} + \left[1 + \left(\frac{\tau_1 - \tau_2}{2\mu} \right)^2 \right]^{1/2}, \quad \frac{\lambda_2^e}{\lambda_1^e} = -\frac{\tau_1 - \tau_2}{2\mu} + \left[1 + \left(\frac{\tau_1 - \tau_2}{2\mu} \right)^2 \right]^{1/2},
$$

(70)

where $\tau_1 \pm \tau_2$ is given in terms of stress by (11). This specifies the components of the elastic moduli tensor $\mathcal{L}_{ijkl}^e$ in terms of the current stress state, as needed in the incremental procedure of solving the path-dependent viscoelastic boundary value problems.

If the membrane is modeled as an isoareal membrane, by using (56), we obtain (on the current principal axes of stress)

$$
\mathring{\sigma}_{ij} = -\dot{p}_0 \delta_{ij} + 2\mu \left[1 + \left(\frac{\sigma_1 - \sigma_2}{2\mu} \right)^2 \right]^{1/2} \left(D_{ij} - \frac{1}{2\eta} \sigma_{ij}' \right)
$$

(71)

and

$$
D_{ij}' = \frac{1}{2\mu} \left[1 + \left(\frac{\sigma_1 - \sigma_2}{2\mu} \right)^2 \right]^{-1/2} \mathring{\sigma}_{ij}' + \frac{1}{2\eta} \sigma_{ij}'.
$$

(72)

5.2. *Kelvin viscoelastic model.* An alternative, Kelvin-type viscoelastic model can be constructed by assuming that the deviatoric part of stress is the sum of the elastic and viscous contributions, $\sigma_{ij}' = \sigma_{ij}'^e + \sigma_{ij}'^v$. By using the elastic constitutive expression (16) and the Newton viscosity law, this gives

$$
\sigma_{ij}' = \frac{1}{\det \boldsymbol{B}} \frac{\partial \phi}{\partial \beta} B_{ij}' + 2\eta D_{ij}'.
$$

(73)

The average normal stress is assumed to be due to elastic deformation only, so that, from (16),

$$
\sigma_{kk} = 2 \frac{\partial \phi}{\partial \alpha}.
$$

(74)

The total stress is

$$
\sigma_{ij} = \frac{\partial \phi}{\partial \alpha} \delta_{ij} + \frac{1}{\det \boldsymbol{B}} \frac{\partial \phi}{\partial \beta} B_{ij}' + 2\eta D_{ij}'.
$$

(75)

When expressed on the principal stress axes, this is

$$
\sigma_1 = \frac{\partial \phi}{\partial \alpha} + \frac{1}{2} \frac{\partial \phi}{\partial \beta} (\lambda_2^{-2} - \lambda_1^{-2}) + \eta \left(\frac{\dot{\lambda}_1}{\lambda_1} - \frac{\dot{\lambda}_2}{\lambda_2} \right),
$$

$$
\sigma_2 = \frac{\partial \phi}{\partial \alpha} + \frac{1}{2} \frac{\partial \phi}{\partial \beta} (\lambda_1^{-2} - \lambda_2^{-2}) + \eta \left(\frac{\dot{\lambda}_2}{\lambda_2} - \frac{\dot{\lambda}_1}{\lambda_1} \right).
$$

(76)

A version of (76), corresponding to an isoareal Evans–Skalak model with $\phi = \mu\beta$, was used by [Mills et al. 2004] in their finite-element evaluation of cell stretching by optical tweezers.

We next determine the expressions for the components of the rate of deformation tensor, which is particularly important if the Kelvin model is combined in series with the Maxwell model to obtain a more complex or capable viscoelastic model. The deviatoric part of the rate of deformation follows from

(73) as

$$D'_{ij} = \frac{1}{2\eta} \left(\sigma'_{ij} - \frac{1}{\det \boldsymbol{B}} \frac{\partial \phi}{\partial \beta} B'_{ij} \right). \tag{77}$$

For a prescribed history of stress, it remains to determine the rate of the area change, i.e., $D_{kk} = d(\ln dA)/dt$. This can be accomplished by differentiating (74), which gives

$$\dot{\sigma}_{kk} = 2 \left(\frac{\partial^2 \phi}{\partial \alpha^2} \dot{\alpha} + \frac{\partial^2 \phi}{\partial \alpha \partial \beta} \dot{\beta} \right). \tag{78}$$

It can be readily shown that

$$\dot{\alpha} = II_B^{1/2} D_{kk}, \quad \dot{\beta} = II_B^{-1/2} B'_{ij} D_{ij}, \tag{79}$$

so that the substitution into (78) yields

$$\dot{\sigma}_{kk} = 2 \frac{\partial^2 \phi}{\partial \alpha^2} II_B^{1/2} D_{kk} + 2 \frac{\partial^2 \phi}{\partial \alpha \partial \beta} II_B^{-1/2} B'_{ij} D'_{ij}. \tag{80}$$

The trace product $B'_{ij} D'_{ij}$ can be eliminated from (80) by using (77), with the result

$$B'_{ij} D'_{ij} = \frac{1}{\eta} \left(\tfrac{1}{2} \sigma_{ij} B'_{ij} - \beta(\beta+2) \frac{\partial \phi}{\partial \beta} \right). \tag{81}$$

In the derivation, the Frobenius norm of $\boldsymbol{B}'$ is expressed as $B'_{ij} B'_{ij} = \frac{1}{2} I_B^2 - 2 II_B$. The substitution of (81) into (80) gives the desired expression for D_{kk} in terms of the current stress and its rate. This is

$$D_{kk} = \frac{II_B^{-1}}{2\partial^2 \phi / \partial \alpha^2} \left[II_B^{1/2} \dot{\sigma}_{kk} - \frac{1}{\eta} \frac{\partial^2 \phi}{\partial \alpha \partial \beta} \left(\sigma_{ij} B'_{ij} - 2\beta(\beta+2) \frac{\partial \phi}{\partial \beta} \right) \right]. \tag{82}$$

If the strain energy is given by the Evans–Skalak form $\phi = \frac{1}{2}\kappa\alpha^2 + \mu\beta$, (77) and (82) reduce to

$$D'_{ij} = \frac{1}{2\eta} \left(\sigma'_{ij} - \frac{\mu}{\det \boldsymbol{B}}, B'_{ij} \right), \quad D_{kk} = \frac{1}{(\det \boldsymbol{B})^{1/2}} \frac{\dot{\sigma}_{kk}}{2\kappa}. \tag{83}$$

On the principal stress axes, this becomes

$$D_1 - D_2 = \frac{1}{2\eta} \left(\sigma_1 - \sigma_2 + \mu(\lambda_1^{-2} - \lambda_2^{-2}) \right), \quad D_1 + D_2 = \frac{1}{2\kappa\lambda_1\lambda_2} (\dot{\sigma}_1 + \dot{\sigma}_2). \tag{84}$$

If the membrane is modeled as isoareal, with $\phi = \phi(\beta)$, then

$$D'_{ij} = \frac{1}{2\eta} \left(\sigma'_{ij} - \frac{d\phi}{d\beta} B'_{ij} \right), \quad D_{kk} = 0. \tag{85}$$

In this case, the average normal stress $\sigma_{kk}/2 = -p_0$ is unspecified by the constitutive analysis.

6. Conclusion

We have presented in this paper the rate-type constitutive analysis of erythrocyte membrane undergoing large elastic and viscoelastic deformations. The results are obtained for an arbitrary isotropic strain energy function, and for its special Evans–Skalak form, commonly used to study the mechanical response of red blood cells. The explicit representation of the fourth-order tensor of elastic moduli, which relates the objective rate of the Kirchhoff stress to the rate of the deformation tensor, is derived with respect

to the principal axes of stress, and an arbitrary set of orthogonal axes, used as the background axes in the numerical treatment of the boundary-value problems. Since the areal modulus of the erythrocyte membrane is several orders of magnitude higher than the shear modulus, the rate-type equations are also derived for an isoareal membrane. The viscoelastic constitutive equations are then derived by adopting either the Maxwell or the Kelvin viscoelastic model, i.e., by adopting the additive decomposition of the rate of deformation, or the stress tensor, into its elastic and viscous parts. More involved viscoelastic models can also be considered, such as those used in [Lubarda and Marzani 2009] in the context of small strains. The obtained constitutive equations may be readily implemented in the incremental (numerical) treatments of the path-dependent boundary-value problems, such as those associated with large deformation and shape recovery of an erythrocyte passing through and exiting from narrow capillaries [Pozrikidis 2003], or excessive aspiration of the cell membrane into micropipette, which may lead to membrane necking followed by formation of a vesicle [Peng et al. 2010]. Other potential application of derived equations is in a related study of the mechanics of nuclear membranes, bounding the nucleus of eukaryotic cells [Vaziri and Mofrad 2007].

The elastic properties of a healthy cell during its lifetime of about 120 days are essential for its main function to deliver the oxygen, while squeezing through capillaries. The degradation of elastic properties and the gradual loss of the membrane elasticity are therefore an important extension of the present work. From the mechanics point of view, this extension may proceed in the spirit of the rate-type theory of damage elasticity [Lubarda and Krajcinovic 1995], provided that the appropriate specification of the parameters which account for biochemical and physical processes of membrane damage and fatigue, and their evolution equations, are available. This, together with the description of active topological remodeling of the membrane structure during large deformations, is a challenging avenue for future research.

Appendix A: Rate-type elasticity with respect to symmetric Piola–Kirchhoff stress

If the initial configuration is used as the reference configuration in the formulation and the solution of the boundary value problem, the rate-type constitutive expressions are needed with respect to the Lagrangian strain

$$\boldsymbol{E} = \tfrac{1}{2}\,(\lambda_1^2 - 1)\,\boldsymbol{N}_1 \otimes \boldsymbol{N}_1 + \tfrac{1}{2}\,(\lambda_2^2 - 1)\,\boldsymbol{N}_2 \otimes \boldsymbol{N}_2,$$

and its conjugate symmetric Piola–Kirchhoff stress $\boldsymbol{S}$. The components of the stress $\boldsymbol{S} = S_1 \boldsymbol{N}_1 \otimes \boldsymbol{N}_1 + S_2 \boldsymbol{N}_2 \otimes \boldsymbol{N}_2$ are

$$S_1 = \frac{1}{\lambda_1}\,\frac{\partial \phi}{\partial \lambda_1}, \quad S_2 = \frac{1}{\lambda_2}\,\frac{\partial \phi}{\partial \lambda_2},$$

such that $\dot{\phi} = S_{ij}\,\dot{E}_{ij}$. Omitting details of the derivation, the corresponding rate-type constitutive expression is

$$\dot{S}_{ij} = \Lambda_{ijkl}\,\dot{E}_{kl},$$

where

$$\Lambda_{1111} = \frac{1}{\lambda_1^2}\left(\frac{\partial^2\phi}{\partial\lambda_1^2} - \frac{1}{\lambda_1}\frac{\partial\phi}{\partial\lambda_1}\right), \quad \Lambda_{2222} = \frac{1}{\lambda_2^2}\left(\frac{\partial^2\phi}{\partial\lambda_2^2} - \frac{1}{\lambda_2}\frac{\partial\phi}{\partial\lambda_2}\right),$$

$$\Lambda_{1122} = \frac{1}{\lambda_1\lambda_2}\frac{\partial^2\phi}{\partial\lambda_1\partial\lambda2}, \quad \Lambda_{1212} = \frac{1}{\lambda_1^2 - \lambda_2^2}\left(\frac{1}{\lambda_1}\frac{\partial\phi}{\partial\lambda_1} - \frac{1}{\lambda_2}\frac{\partial\phi}{\partial\lambda_2}\right),$$

with the obvious symmetries $\Lambda_{1212} = \Lambda_{2121} = \Lambda_{1221} = \Lambda_{2112}$ and $\Lambda_{1122} = \Lambda_{2211}$. The well-known relationship [Hill 1978] between the moduli $\mathcal{L}_{ijkl}$, appearing in the constitutive structure $\overset{\circ}{\tau}_{ij} = \mathcal{L}_{ijkl}D_{kl}$, and the moduli Λ_{ijkl}, appearing in the constitutive structure $\dot{S}_{ij} = \Lambda_{ijkl}\dot{E}_{kl}$, is

$$\mathcal{L}_{ijkl} = F_{im}F_{jn}\Lambda_{mnpq}F_{kp}F_{lq} + \frac{1}{2}(\tau_{ik}\delta_{jl} + \tau_{jk}\delta_{il} + \tau_{il}\delta_{jk} + \tau_{jl}\delta_{ik}).$$

If the strain energy is given by the Evans–Skalak form (6), the elastic moduli become

$$\Lambda_{1111} = \frac{1}{\lambda_1^4}\left(\kappa\lambda_1\lambda_2 + \mu\frac{3\lambda_2^2 - \lambda_1^2}{2\lambda_1\lambda_2}\right), \quad \Lambda_{1122} = \frac{1}{\lambda_1^2\lambda_2^2}\left(\kappa\lambda_1\lambda_2(2\lambda_1\lambda_2 - 1) - \mu\frac{\lambda_1^2 + \lambda_2^2}{2\lambda_1\lambda_2}\right),$$

$$\Lambda_{2222} = \frac{1}{\lambda_2^4}\left(\kappa\lambda_1\lambda_2 + \mu\frac{3\lambda_1^2 - \lambda_2^2}{2\lambda_1\lambda_2}\right), \quad \Lambda_{1212} = \frac{1}{\lambda_1^2\lambda_2^2}\left(\kappa\lambda_1\lambda_2(1 - \lambda_1\lambda_2) + \mu\frac{\lambda_1^2 + \lambda_2^2}{2\lambda_1\lambda_2}\right).$$

Appendix B: Rate-type elasticity with respect to nominal stress

The nominal stress $\boldsymbol{P} = \boldsymbol{F}^{-1}\cdot\boldsymbol{\tau}$ appears in the variational formulation of the boundary value problems, when the differential equations of motion are expressed in terms of the displacement components. The rate of work is then $\dot{\phi} = P_1\dot{\lambda}_1 + P_2\dot{\lambda}_2$, and the components of $\boldsymbol{P} = P_1\boldsymbol{N}_1 \otimes \boldsymbol{n}_1 + P_2\boldsymbol{N}_2 \otimes \boldsymbol{n}_2$ are

$$P_1 = \frac{\partial\phi}{\partial\lambda_1}, \quad P_2 = \frac{\partial\phi}{\partial\lambda_2}.$$

Again omitting details of the derivation, it follows that $\dot{P}_{ji} = \hat{\Lambda}_{jilk}\dot{F}_{kl}$, where

$$\hat{\Lambda}_{1111} = \frac{\partial^2\phi}{\partial\lambda_1^2}, \quad \hat{\Lambda}_{2222} = \frac{\partial^2\phi}{\partial\lambda_2^2}, \quad \hat{\Lambda}_{1122} = \hat{\Lambda}_{2211} = \frac{\partial^2\phi}{\partial\lambda_1\partial\lambda_2},$$

$$\hat{\Lambda}_{1212} = \hat{\Lambda}_{2121} = \frac{1}{\lambda_1^2 - \lambda_2^2}\left(\lambda_1\frac{\partial\phi}{\partial\lambda_1} - \lambda_2\frac{\partial\phi}{\partial\lambda_2}\right),$$

$$\hat{\Lambda}_{1221} = \hat{\Lambda}_{2112} = \frac{1}{\lambda_1^2 - \lambda_2^2}\left(\lambda_2\frac{\partial\phi}{\partial\lambda_1} - \lambda_1\frac{\partial\phi}{\partial\lambda_2}\right).$$

Since $\dot{\boldsymbol{P}} = \dot{\boldsymbol{S}}\cdot\boldsymbol{F}^T + \boldsymbol{P}\cdot\boldsymbol{L}^T$, the relationship between the pseudomoduli $\hat{\Lambda}_{ijkl}$ and the moduli Λ_{ijkl} from Appendix A is $\hat{\Lambda}_{jilk} = \Lambda_{jmln}F_{im}F_{kn} + S_{jl}\delta_{ik}$.

If the strain energy is given by (6), the elastic pseudomoduli become

$$\hat{\Lambda}_{1111} = \kappa \lambda_2^2 + \mu \frac{\lambda_2}{\lambda_1^3}, \quad \hat{\Lambda}_{2222} = \kappa \lambda_1^2 + \mu \frac{\lambda_1}{\lambda_2^3},$$

$$\hat{\Lambda}_{1122} = \hat{\Lambda}_{2211} = \kappa (2\lambda_1\lambda_2 - 1) - \mu \frac{\lambda_1^2 + \lambda_2^2}{2\lambda_1^2\lambda_2^2},$$

$$\hat{\Lambda}_{1212} = \hat{\Lambda}_{2121} = \mu \frac{1}{\lambda_1\lambda_2},$$

$$\hat{\Lambda}_{1221} = \hat{\Lambda}_{2112} = \kappa (1 - \lambda_1\lambda_2) + \mu \frac{\lambda_1^2 + \lambda_2^2}{2\lambda_1^2\lambda_2^2}.$$

References

[Agrawal and Steigmann 2009] A. Agrawal and D. J. Steigmann, "Boundary-value problems in the theory of lipid membranes", *Continuum Mech. Therm.* **21**:1 (2009), 57–82.

[Agre and Parker 1989] P. Agre and J. C. Parker (editors), *Red blood cell membranes: structure, function, clinical implications*, Marcel Dekker, New York, 1989.

[Boal 2002] D. H. Boal, *Mechanics of the cell*, Cambridge University Press, New York, 2002.

[Dao et al. 2003] M. Dao, C. T. Lim, and S. Suresh, "Mechanics of the human red blood cell deformed by optical tweezers", *J. Mech. Phys. Solids* **51**:11-12 (2003), 2259–2280.

[Evans and Hochmuth 1976] E. A. Evans and R. M. Hochmuth, "Membrane viscoelasticity", *Biophys J.* **16**:1 (1976), 1–11.

[Evans and Skalak 1980] E. A. Evans and R. Skalak, *Mechanics and thermodynamics of biomembranes*, CRC Press, Boca Raton, FL, 1980.

[Fung 1993] Y. C. Fung, *Biomechanics: mechanical properties of living tissues*, Springer, New York, 1993.

[Garikipati et al. 2006] K. Garikipati, J. E. Olberding, H. Narayanan, E. M. Arruda, K. Grosh, and S. Calve, "Biological remodelling: stationary energy, configurational change, internal variables and dissipation", *J. Mech. Phys. Solids* **54**:7 (2006), 1493–1515.

[Géminard et al. 2004] J.-C. Géminard, R. Bernal, and F. Melo, "Wrinkle formations in axi-symmetrically stretched membranes", *Eur. Phys. J. E* **15**:2 (2004), 117–126.

[Gov 2007] N. S. Gov, "Active elastic network: cytoskeleton of the red blood cell", *Phys. Rev. E* **75**:1 (2007), 011921.

[Haughton 2001] D. M. Haughton, "Elastic membranes", pp. 233–267 in *Nonlinear elasticity: theory and applications*, edited by Y. B. Fu and R. W. Ogden, London Mathematical Society Lecture Note Series **283**, Cambridge University Press, Cambridge, 2001.

[Haughton and Ogden 1978] D. M. Haughton and R. W. Ogden, "On the incremental equations in non-linear elasticity, I: Membrane theory", *J. Mech. Phys. Solids* **26**:2 (1978), 93–110.

[Hill 1978] R. Hill, "Aspects of invariance in solid mechanics", pp. 1–75 in *Advances in applied mechanics*, vol. 18, edited by C.-S. Yih, Academic Press, New York, 1978.

[Hochmuth 1987] R. M. Hochmuth, "Properties of red blood cells", pp. 12.1–12.17 in *Handbook of bioengineering*, edited by R. Skalak and S. Chien, McGraw-Hill, New York, 1987.

[Hochmuth and Waugh 1987] R. M. Hochmuth and R. E. Waugh, "Erythrocyte membrane elasticity and viscosity", *Annu. Rev. Physiol.* **49** (1987), 209–219.

[Li et al. 2007] J. Li, G. Lykotrafitis, M. Dao, and S. Suresh, "Cytoskeletal dynamics of human erythrocyte", *Proc. Nat. Acad. Sci. USA* **104**:12 (2007), 4937–4942.

[Lubarda 2010] V. A. Lubarda, "Constitutive analysis of thin biological membranes with application to radial stretching of a hollow circular membrane", *J. Mech. Phys. Solids* **58**:6 (2010), 860–873.

[Lubarda and Hoger 2002] V. A. Lubarda and A. Hoger, "On the mechanics of solids with a growing mass", *Int. J. Solids Struct.* **39**:18 (2002), 4627–4664.

[Lubarda and Krajcinovic 1995] V. A. Lubarda and D. Krajcinovic, "Constitutive structure of rate theory of damage in brittle elastic solids", *Appl. Math. Comput.* **67**:1-3 (1995), 81–101.

[Lubarda and Marzani 2009] V. A. Lubarda and A. Marzani, "Viscoelastic response of thin membranes with application to red blood cells", *Acta Mech.* **202**:1-4 (2009), 1–16.

[Mills et al. 2004] J. P. Mills, L. Qie, M. Dao, C. T. Lim, and S. Suresh, "Nonlinear elastic and viscoelastic deformation of the human red blood cell with optical tweezers", *Mech. Chem. Biosyst.* **1**:3 (2004), 169–180.

[Ogden 1984] R. W. Ogden, *Non-linear elastic deformations*, Ellis Horwood, Chichester, 1984.

[Park et al. 2010] Y.-K. Park, C. A. Best, T. Auth, N. S. Gov, S. A. Safran, G. Popescu, S. Suresh, and M. S. Feld, "Metabolic remodeling of the human red blood cell membrane", *Proc. Nat. Acad. Sci. USA* **107**:4 (2010), 1289–1294.

[Peng et al. 2010] Z. Peng, R. J. Asaro, and Q. Zhu, "Multiscale simulation of erythrocyte membranes", *Phys. Rev. E* **81**:3 (2010), 031904.

[Pozrikidis 2003] C. Pozrikidis (editor), *Modeling and simulation of capsules and biological cells*, Chapman & Hall/CRC Press, Boca Raton, FL, 2003.

[Taber 2009] L. A. Taber, "Towards a unified theory for morphomechanics", *Phil. Trans. R. Soc. A* **367**:1902 (2009), 3555–3583.

[Tözeren et al. 1982] A. Tözeren, R. Skalak, K.-L. P. Sung, and S. Chien, "Viscoelastic behavior of erythrocyte membrane", *Biophys. J.* **39**:1 (1982), 23–32.

[Vaziri and Mofrad 2007] A. Vaziri and M. R. K. Mofrad, "Mechanics and deformation of the nucleus in micropipette aspiration experiment", *J. Biomech.* **40**:9 (2007), 2053–2062.

[Wineman 2007] A. Wineman, "Nonlinear viscoelastic membranes", *Comput. Math. Appl.* **53**:2 (2007), 168–181.

[Zhu et al. 2007] Q. Zhu, C. Vera, R. J. Asaro, P. Sche, and L. A. Sung, "A hybrid model for erythrocyte membrane: a single unit of protein network coupled with lipid bilayer", *Biophys. J.* **93**:2 (2007), 386–400.

VLADO A. LUBARDA: vlubarda@ucsd.edu
Department of Mechanical and Aerospace Engineering, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0411, United States

and

Montenegrin Academy of Sciences and Arts, Rista Stijovića 5, 81000 Podgorica, Montenegro

∎msp

FUNCTIONALLY GRADED BARS
WITH ENHANCED DYNAMIC PERFORMANCE

KARAM Y. MAALAWI

To Charles and Marie-Louise Steele, with gratitude.

This paper presents analytical solutions for enhancing the dynamic performance of functionally graded material bars in axial motion. Optimized designs with maximized natural frequencies under mass equality constraint are given and discussed. The composition of the construction material is optimized by defining the spatial distribution of volume fractions of the material constituents using either continuous or discrete variations along the bar length. Three cases of boundary conditions have been examined: fixed-fixed, fixed-free and free-free bars. The major aim is to tailor the mass and stiffness distributions in the axial direction so as to maximize the frequencies and place them at their target values to avoid the occurrence of large amplitudes of vibration without the penalty of increasing total structural mass. The resulting optimization problem has been formulated as a nonlinear mathematical programming problem solved by invoking the Matlab optimization toolbox routines, which implement the method of feasible directions interacting with the associated eigenvalue problem routines. The proposed mathematical models have shown that the use of material grading can be promising in optimizing natural frequencies without mass penalty.

1. Introduction

Functionally graded materials (FGMs) are advanced composite materials fabricated to have graded variation of the relative fractions of the constituent materials. Commonly, these materials are made from a mixture of ceramic and metal or a combination of different metals. Ceramic provides high temperature resistance because of its low thermal conductivity, while metal secures the necessary strength and stiffness. FGMs show promise in many applications, such as spacecraft heat shields, high performance structural elements and critical engine components.

A few studies have addressed the dynamics of FGM structures. Loy et al. [1999] considered vibration analysis of FGM cylindrical shells composed of stainless steel and nickel, which were graded in thickness direction according to a volume fraction power-law distribution. Love's shell theory was implemented and the resulting eigenvalue problem was solved by Rayleigh–Ritz method. The behavior of the natural frequencies as well as the influence of the volume fractions were investigated and presented in detail. Cheng and Batra [2000] studied the steady-state vibrations of a simply supported FGM plate on an elastic foundation and subjected to uniform in-plane hydrostatic loads. They applied Reddy's third-order theory, taking the effect of rotary inertia into consideration. The Young's modulus and the Poisson's ratio were assumed to vary only in the plate thickness direction.

Closed-form solutions for calculating the natural frequencies of an axially graded beam were derived and given by [Elishakoff and Guédé 2004]. The modulus of elasticity was taken as a polynomial in the axial coordinate along the beam's length, and an inverse problem was solved to find the stiffness and mass distributions so that the chosen polynomial could lead to an exact mode shape. Three-dimensional exact solutions for free and forced vibrations of FGM plates were introduced in [Vel and Batra 2004] and [Elishakoff et al. 2005]. The material properties of composite plates were also allowed to vary in the thickness direction. Exact natural frequencies were obtained which can be used to assess the accuracy of the different plate theories. Tylikowski [2005] presented a study of parametric vibrations of functionally graded plates subjected to in-plane time-dependent forces. The material properties were graded in the thickness direction of the plate according to a volume fraction power-law distribution. Effects of power-law exponent on the stability behavior were also investigated.

Considering, next, the optimal frequency design problem, several methods have been developed and applied in solving both the frequency and mass optimization problems. They are broadly classified as mathematical programming techniques, optimality criteria methods and optimal control theory.

Mathematical programming has dealt successfully with a wide range of structural optimization problems, while control theory has been applied with limited success because of the great analytical complexity. Niordson [1965] and Brach [1968] initiated the formulation of the extremal problem of finding the maximum fundamental frequencies of vibrating beams under mass constraint. They applied variational methods and investigated the effect of beam taper on the resulting optimum solutions. However, their attained designs were characterized by zero cross-sectional area at the end supports, which results in infinite stresses near the boundaries of the beam. Following their work, several other researchers applied optimization techniques for solving one-dimensional structural members. Warner and Vavrick [1975] utilized a variational continuous optimization procedure to find the minimal mass of bars in axial motion under several frequency constraints. They indicated that the piece-wise uniform design can be economically efficient as compared to the continuous design. In [Maalawi and Warner 1984] we obtained global optimality solutions for practical rod constructions made of piecewise uniform sections. Their objective function was measured by maximization of the fundamental frequency under mass constraint, where conspicuous trends were deduced under specified boundary conditions. Other optimum shapes of one-dimensional structural models for maximum fundamental frequency were studied in [Masad 1997], who presented an efficient numerical approach for computing the rate of change of the associated eigenvalues. In [Maalawi 1999] we presented an exact method for the problem of frequency maximization of rods in axial motion under mass constraint. The dual problem of mass minimization under frequency constraint was also addressed, and results were presented and discussed for both continuous and discrete multi-element structural models.

In the field of the optimum design of FGM-structures, Qian and Batra [2005] considered frequency optimization of a cantilevered plate with variable volume fraction according to simple power-laws. They implemented genetic algorithms to find the optimum values of the power exponents, which maximize the natural frequencies, and concluded that the volume fraction needs to be varied in the longitudinal direction of the plate rather than in the thickness direction. Goupee and Vel [2006] proposed a methodology to optimize the natural frequencies of functionally graded beam with variable volume fraction of the constituent materials in the beam's length and height directions. They used a piecewise bi-cubic interpolation of volume fraction values specified at a finite number of grid points, and applied a genetic algorithm code

to find the needed optimum designs. Considering aeroelastic optimization FGM structures, Librescu and Maalawi [2007] introduced the underlying concepts of using material grading in optimizing subsonic wings against torsional instability. They developed exact mathematical models allowing the material physical and mechanical properties to change in the wing spanwise direction, where both continuous and piecewise structural models were successfully implemented.

Little is found in the literature that deals with frequency optimization of FGM bars subject to a *mass equality constraint*. The aim of the present study is, therefore, to incorporate the effect of changing the volume fractions of the constituent materials for enhancing the dynamic performance of an axially vibrating bar. Following [Librescu and Maalawi 2007; Warner 2001; Maalawi and Warner 1984], a useful optimization tool has been built for designing efficient patterns of FGM bars having maximized frequencies while maintaining the total structural mass at constraint value. This allows the search for the required optimal volume fractions which satisfy both design criteria. The corresponding increases in the frequencies calculated with respect to a baseline design have been evaluated for several configurations, including cases of fixed-fixed, fixed-free and free-free bars fabricated from certain types of composite materials. Actually, substantial improvement in the overall dynamic performance has been attained showing the usefulness of the proposed model in using the concept of material grading to arrive at the needed optimum designs for one dimensional slender bars having arbitrary cross-sectional type.

2. Structural dynamic analysis

The partial differential equation that governs the free axial motion of a vibrating bar is [Meirovitch 1997]

$$\frac{\partial}{\partial x}\left[EA(x)\frac{\partial u(x,t)}{\partial x}\right] = \rho A(x)\frac{\partial^2 u(x,t)}{\partial t^2}, \tag{1}$$

which must be satisfied over the entire bar's length L. In (1), x denotes the axial coordinate, $u(x,t)$ longitudinal displacement at any position x and time t, $E(x)$ modulus of elasticity, $A(x)$ cross-sectional area and $\rho(x)$ mass density. Three types of boundary conditions are described:

(I) fixed-fixed: $u(0,t) = 0$ and $u(L,t) = 0$

(II) fixed-free: $u(0,t) = 0$ and $EA\frac{\partial u}{\partial x}\Big|_{x=L} = 0$

(III) free-free: $EA\frac{\partial u}{\partial x}\Big|_{x=0} = 0$ and $EA\frac{\partial u}{\partial x}\Big|_{x=L} = 0$

The longitudinal displacement $u(x,t)$ is assumed to be separable in space and time, $u(x,t) = U(x) \cdot q(t)$, where the time dependence $q(t)$ is harmonic with circular frequency ω. Substituting for $d^2q/dt^2 = -\omega^2 q$, the associated eigenvalue problem can be written directly in the form

$$\frac{d}{dx}\left[EA(x)\frac{dU}{dx}\right] + \rho A(x)\omega^2 U(x) = 0 \tag{2}$$

which must be satisfied on the interval $0 < x < L$.

2.1. *Definition of the baseline design.* An essential phase in formulating an optimization problem is to define an appropriate baseline design to which the resulting optimal designs can be compared. The

baseline design has been selected to be a composite bar made of two different materials, denoted by A and B, and has uniform mass and stiffness distributions along its length, with equal volume fractions (V) of the constituent materials, i.e., $V_A = V_B = 50\%$. It is assumed that the optimized designs shall have the same total mass, length, cross sectional area and shape, type of construction material and type of boundary conditions of those of the known baseline design. Therefore, the design variables, which are subject to change in the optimization process, shall be the distribution of the volume fractions over the bar length.

2.2. *Material grading.* The physical and mechanical properties are allowed to vary lengthwise, that is, the material is graded in the axial direction. The distributions of the mass density ρ and modulus of elasticity E are taken to obey the semi-empirical formulas of [Halpin and Tsai 1969]. Assuming no voids are present, we have

$$V_A(x) + V_B(x) = 1,$$
$$\rho(x) = V_A(x)\rho_A + V_B(x)\rho_B,$$
$$E(x) = V_A(x)E_A + V_B(x)E_B.$$
(3)

Normalizing with respect to the baseline design by dividing (2) by $E_0 A / L$ and differentiating once more with respect to x, we get

$$\hat{E}\frac{d^2\hat{U}}{d\hat{x}^2} + \frac{d\hat{E}}{d\hat{x}} \cdot \frac{d\hat{U}}{d\hat{x}} + \hat{\rho}\hat{\omega}^2\hat{U} = 0$$
(4)

which must be satisfied on the interval $0 < \hat{x} < 1$. Dimensionless quantities are defined as follows:

Axial coordinate	$\hat{x} = x/L$
Longitudinal displacement	$\hat{U} = U/L$
Modulus of elasticity	$\hat{E} = E/E_0$
Mass density	$\hat{\rho} = \rho/\rho_0$
Circular frequency	$\hat{\omega} = \omega L\sqrt{\rho_0/E_0}$

where the baseline design parameters $E_0 = \frac{1}{2}(E_A + E_B)$ and $\rho_0 = \frac{1}{2}(\rho_A + \rho_B)$. The cross-sectional area A is assumed to be constant. The total structural mass is kept equal to that of the baseline:

$$\hat{M}_s = \int_0^1 \hat{\rho}\, d\hat{x} = 1,$$
(5)

implying the dimensionless mass equality constraint

$$\int_0^1 V_A(\hat{x})\, d\hat{x} = \frac{1}{2}.$$
(6)

3. Solution procedures

3.1. *Continuous model: an exact power series solution.* Kumar and Sujith [1997] presented exact analytical solutions for the longitudinal vibration of nonuniform rods. The equation of motion was reduced to a standard differential equation solved in terms of Bessel and Neumann functions. However, the solutions were restricted to rods having polynomial area variation and made of an isotropic material with

uniform properties. A more general solution of (4), where the modulus of elasticity and mass density vary spatially, can be expressed by the power series [Edwards and Penney 2004]

$$\hat{U}(\hat{x}) = \sum_{m=1}^{2} C_m \lambda_m(\hat{x}), \tag{7}$$

where the C_m are the constants of integration and the λ_m are two linearly independent solutions that have the form:

$$\lambda_m(\hat{x}) = \sum_{n=m}^{\infty} a_{m,n} \hat{x}^{n-1}, \quad n \geq m. \tag{8}$$

The unknown coefficients $a_{m,n}$ can be determined by substitution into the differential equation (4) and equating coefficients of like powers of $\hat{x}$. The variation of the volume fractions in FGM structures is usually described by power-law distributions [Elishakoff and Guédé 2004; Qian and Batra 2005; Tylikowski 2005]. In the present study, linear and parabolic model types are considered with their corresponding derived recurrence formulas:

Linear distribution.

- Type (1): $V_A(\hat{x}) = V_A(0) - \Delta V \hat{x}, \ \ \Delta V = V_A(0) - V_A(1)$

$$a_{m,n} = \frac{\gamma(n-2)^2 a_{m,n-1} - \theta \hat{\omega}^2 a_{m,n-2} + \zeta \hat{\omega}^2 a_{m,n-3}}{\beta(n-1)(n-2)} \tag{9}$$

- Type (2) – Symmetric about $\hat{x} = 1/2$: $V_A(\hat{x}) = V_A(0) - 2\Delta V \hat{x}, \ \ \Delta V = V_A(0) - V_A(1/2)$

$$a_{m,n} = \frac{2\gamma(n-2)^2 a_{m,n-1} - \theta \hat{\omega}^2 a_{m,n-2} + 2\xi \hat{\omega}^2 a_{m,n-3}}{\beta(n-1)(n-2)} \tag{10}$$

Parabolic distribution.

- Type (3) – Concave: $V_A(\hat{x}) = V_A(0) + \Delta V(\hat{x}^2 - 2\hat{x}), \ \ \Delta V = V_A(0) - V_A(1)$

$$a_{m,n} = \frac{2\gamma(n-2)^2 a_{m,n-1} - [\gamma(n-2)(n-3) + \theta \hat{\omega}^2]a_{m,n-2} + \xi \hat{\omega}^2[2a_{m,n-3} - a_{m,n-4}]}{\beta(n-1)(n-2)} \tag{11}$$

- Type (4) – Convex: $V_A(\hat{x}) = V_A(0) - \Delta V \hat{x}^2, \ \ \Delta V = V_A(0) - V_A(1)$

$$a_{m,n} = \frac{[\gamma(n-2)(n-3) - \theta \hat{\omega}^2]a_{m,n-2} + \xi \hat{\omega}^2 a_{m,n-4}}{\beta(n-1)(n-2)} \tag{12}$$

- Type (5) – Symmetric about $\hat{x} = 1/2$: $V_A(\hat{x}) = V_A(0) + 4\Delta V(\hat{x}^2 - \hat{x}), \ \ \Delta V = V_A(0) - V_A(1/2)$

$$a_{m,n} = \frac{4\gamma(n-2)^2 a_{m,n-1} - [4\gamma(n-2)(n-3) + \theta \hat{\omega}^2]a_{m,n-2} + 4\xi \hat{\omega}^2(a_{m,n-3} - a_{m,n-4})}{\beta(n-1)(n-2)} \tag{13}$$

The Greek symbols in (9)–(13) are defined as follows, with $\Delta \hat{E} = \hat{E}_A - \hat{E}_B$ and $\Delta \hat{\rho} = \hat{\rho}_A - \hat{\rho}_B$:

$$\beta = \hat{E}_B + \Delta \hat{E} V_A(0), \quad \gamma = \Delta \hat{E} \Delta V, \quad \theta = \hat{\rho}_B + \Delta \hat{\rho} V_A(0), \quad \xi = \Delta \hat{\rho} \Delta V.$$

boundary conditions		frequency equation	$\hat{\omega}_{0,i}$
fixed-fixed, symmetric modes	$\hat{U}(0) = \hat{U}'\left(\frac{1}{2}\right) = 0$	$\sum_{n=3}^{\infty} a_{2n}(n-1)/2^{n-2} = -1$	$(\pi, 3\pi, 5\pi)$
fixed-fixed, asymmetric modes	$\hat{U}(0) = \hat{U}\left(\frac{1}{2}\right) = 0$	$\sum_{n=2}^{\infty} a_{2n}/2^{n-1} = 0$	$(2\pi, 4\pi, 6\pi)$
fixed-free	$\hat{U}(0) = \hat{U}'(1) = 0$	$\sum_{n=3}^{\infty} a_{2n}(n-1) = -1$	$(\pi, 3\pi, 5\pi)/2$
free-free, symmetric modes	$\hat{U}'(0) = \hat{U}'\left(\frac{1}{2}\right) = 0$	$\sum_{n=3}^{\infty} a_{1n}(n-1)/2^{n-2} = 0$	$(2\pi, 4\pi, 6\pi)$
free-free, asymmetric modes	$\hat{U}'(0) = \hat{U}\left(\frac{1}{2}\right) = 0$	$\sum_{n=3}^{\infty} a_{1n}/2^{n-1} = -1$	$(\pi, 3\pi, 5\pi)$

Table 1. Frequency equations for different types of boundary conditions. The column $\hat{\omega}_{0,i}$ gives the dimensionless natural frequencies of the baseline design, with $\hat{\omega}_0 = 0$ corresponding to the first (rigid-body) mode of a free-free bar. Differentiation is with respect to $\hat{x}$.

A coefficient $a_{m,n}$ is set equal to zero whenever n is less than m, and the leading coefficients $a_{m,m}$ in each series are arbitrary and can be set equal to one. This method was successfully applied by [Librescu and Maalawi 2007] to determine the exact critical flight speed of a FGM subsonic wing. Table 1 summarizes the appropriate mathematical expressions of the frequency equation for any desired case, which can be obtained by application of the associated boundary conditions and consideration of nontrivial solutions.

3.2. *Piecewise model: The transfer matrix method.* A piecewise model concept was introduced in [Maalawi and Warner 1984] to obtain global optimal frequency designs of isotropic bars with piecewise uniform sections. They showed that the use of piecewise models in structural optimization gives excellent results and can be promising for similar applications. Li [2000] presented solutions of the differential equation of longitudinal vibration of bars with stiffness and mass distributions, which were described by specific power functions. He applied the transfer matrix method of [Pestel and Leckie 1963] to derive the frequency equation of multi-step bars with several boundary conditions, where analytical solutions were verified with full scale measured data of a multi-story tall building. Figure 1 shows an elastic, slender bar
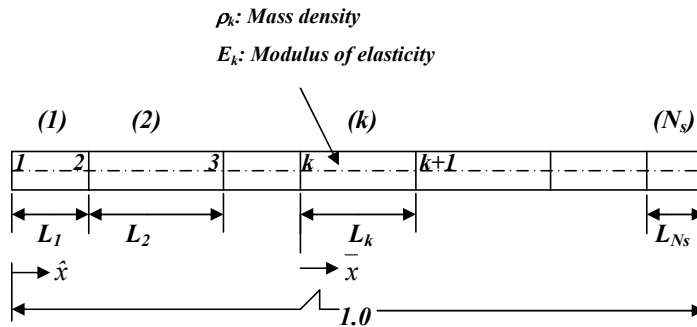
Figure 1. General configuration of a piecewise axially graded bar.

with total dimensionless length of unity constructed from any arbitrary number of uniform segments (N_s), each of which has the same cross-sectional area but different properties of the construction material. Such a configuration results in a piecewise axial grading of the material in the direction of the bar axis. Before determining the exact natural frequencies and performing the necessary mathematics, it is important to bear in mind that design optimization is only as meaningful as its core structural analysis model. Any deficiencies therein will certainly be reflected in the optimization process.

For the k-th segment, (4) reduces to

$$\frac{d^2\hat{U}}{d\bar{x}^2} + \alpha_k^2 \hat{U} = 0, \quad \alpha_k = \hat{\omega}\sqrt{\hat{\rho}_k/\hat{E}_k}, \tag{14}$$

where $\hat{E}_k = E_k/E_0$ and $\hat{\rho}_k = \rho_k/\rho_0$ are the dimensionless modulus of elasticity and mass density of the k-th segment. Equation (14) must be satisfied in the interval $0 \le \bar{x} \le \hat{L}_k$, where $\bar{x} = \hat{x} - \hat{x}_k$ and $\hat{L}_k = L_k/L$. Its general solution is:

$$\hat{U}(\bar{x}) = C_1 \sin(\alpha_k \bar{x}) + C_2 \cos(\alpha_k \bar{x}). \tag{15}$$

Expressing the constants of integration C_1 and C_2 in terms of the state variables $[\hat{U}, \hat{F} = \hat{E}\, d\hat{U}/d\bar{x}]$ at both ends of the k-th segment, we get

$$\begin{Bmatrix} \hat{U}_{k+1} \\ \hat{F}_{k+1} \end{Bmatrix} = [T^{(k)}] \begin{Bmatrix} \hat{U}_k \\ \hat{F}_k \end{Bmatrix},$$

where

$$[T^{(k)}] = \begin{bmatrix} \cos\alpha_k\hat{L}_k & \sin\alpha_k\hat{L}_k/(\alpha_k\hat{E}k) \\ -\alpha_k\hat{E}_k\sin\alpha_k\hat{L}_k & \cos\alpha_k\hat{L}_k \end{bmatrix} \tag{16}$$

is the transfer matrix of the k-th segment. It is now possible to compute the state variables progressively along the bar length by applying continuity requirements among the interconnecting boundaries of the various bar segments. Therefore, the state variables at both ends of the bar can be related to each other through an overall transfer matrix denoted by $[T]$:

$$[T] = [T^{(N_s)}][T^{(N_s-1)}]\dots[T^{(2)}][T^{(1)}]. \tag{17}$$

The required frequency equation for determining the natural frequencies can then be obtained by applying the associated boundary conditions (refer to Table 1) and considering only the nontrivial solution of the resulting matrix equation.

4. Optimization problem formulation

4.1. *General.* Attractive goals of designing efficient structural members include minimization of structural weight, maximization of the fundamental frequencies [Masad 1997], and minimization of total cost of production. Global optimization models of elastic beam-type structures were presented in [Maalawi and El-Chazly 2002], where both stability and dynamic optimization problems were formulated in a standard mathematical programming coupled with finite element analysis routines. Another important consideration is the reduction or control of the vibration level [Maalawi and Negm 2002]. Vibration can greatly influence the performance of several structural systems because of its adverse effects on stability,

fatigue life and noise. The reduction of vibration may be attained either by a direct maximization of the natural frequencies or by separating the natural frequencies of the structure from the harmonics of the exciting applied loads. This would avoid resonance and large amplitudes of vibration, which may cause severe damage of the structure. Direct maximization of the natural frequencies can ensure a simultaneous balanced improvement in both of the overall stiffness and structural mass distributions of the vibrating bar. The associated optimization problems are usually cast in nonlinear mathematical programming form [Vanderplaats 1999] where the objective is to minimize a function $F(\underline{X})$ of a vector $\underline{X}$ of design variables, subject to certain number of constraints $G_j(\underline{X}) \leq 0$, $j = 1, 2, \ldots, m$. Iterative techniques are usually used for solving such optimization problems in which a series of directed design changes are made between successive points in the design space. The method of feasible directions is one of the most powerful methods in finding the required constrained optimum solution. The search direction $\underline{S}_j$ must satisfy the conditions $\underline{S}_j \underline{\nabla} F < 0$ and $\underline{S}_j \underline{\nabla} G_j < 0$, where $\underline{\nabla} F$ and $\underline{\nabla} G_j$ are the gradient vectors of the objective and constraint functions, respectively. For checking the constrained minima, the Kuhn–Tucker test [Vanderplaats 1999] is applied at the design point $\underline{X}_D$, which lies on one or more set of active constraints. The Kuhn–Tucker equations are necessary conditions for optimality for a constrained optimization problem and their solution forms the basis to the method of feasible directions.

4.2. *Optimal frequency model.* In the present optimization problem, two alternatives of the objective function form have been implemented and examined. The first one is represented by a direct maximization of a weighted sum of the natural frequencies, which is expressed mathematically as follows:

$$\text{Minimize } F(\underline{X}) = -\sum_i \alpha_i \hat{\omega}_i, \qquad \sum_i \alpha_i = 1, \quad 0 \leq \alpha_i \leq 1, \tag{18}$$

where $\hat{\omega}_i$ are the normalized frequencies, α_i weighting factors measuring the relative importance of each frequency [Kasprzak and Lewis 2001] and $\underline{X}$ is the chosen design variable vector. For continuous models with known power-law distribution, $\underline{X}$ represents the volume fractions at the ends of the bar, while for discrete model $\underline{X} = (V_{A,k}, \hat{L}_k)_{k,1,2,\ldots N_s}$. The second alternative is to minimize a weighted sum of the squares of the difference between each frequency $\hat{\omega}_i$ and its target or desired value $\hat{\omega}_i^*$:

$$\text{Minimize } F(\underline{X}) = \sum_i \alpha_i (\hat{\omega}_i - \hat{\omega}_i^*)^2. \tag{19}$$

Both objectives are subject to the constraints

$$\text{Mass constraint:} \quad \hat{M}_s = 1, \tag{20}$$

$$\text{Side constraints:} \quad \underline{X}_L \leq \underline{X} \leq \underline{X}_U, \tag{21}$$

where $\underline{X}_L$ and $\underline{X}_U$ are the lower and upper limiting values imposed on the design variables vector $\underline{X}$ in order not to obtain unrealistic odd-shaped designs in the final optimum solutions. Approximate values of the target frequencies are usually chosen to be within close ranges; sometimes called frequency-windows; of those corresponding to an initial baseline design, which are adjusted to be far away from the critical exciting frequencies. The proper choice of the best form of the objective function has to wait for actual computer experimentation. Several computer program packages are available now for solving the above design optimization model, which can be coded to interact with structural and eigenvalue

	Material A (carbon fibers)	Material B (epoxy matrix)
Mass density (g/cm^3)	$\rho_f = 1.81$	$\rho_m = 1.27$
Young's modulus (GPa)	$E_{1f} = 235$	$E_m = 4.3$
Shear modulus (GPa)	$G_{12f} = 27$	$G_m = 1.6$
Poisson's ratio	$\nu_{12f} = 0.2$	$\nu_m = 0.35$

Table 2. Material properties of carbon-AS4 / epoxy-3501-6 composite.

analysis software. The Matlab optimization toolbox is a powerful tool that includes many routines for different types of optimization encompassing both unconstrained and constrained minimization algorithms [Venkataraman 2002]. One of its useful routines, fmincon, implements the method of feasible directions in finding the constrained minimum of an objective function of several variables.

5. Optimization results and discussions

The mathematical models developed above have been applied to obtain the required optimal solutions of FGM bars with different boundary conditions (see Table 1). The selected construction material is composed of carbon-AS4 (material A) and epoxy3501-6 (material B), which has favorable characteristics and is highly desirable in several mechanical, civil and aerospace engineering applications [Daniel and Ishai 2006]; see Table 2.

Two different forms of the objective function measuring frequency optimization have been defined in (18) and (19). Whatever is the approach taken, it can be useful to investigate first the behavior of the natural frequencies for the different boundary conditions and see how they are changed with the selected design variables while maintaining the total structural mass constant. Computer experimentation for optimizing fixed-fixed and free-free bars with either discrete or continuous material grading have indicated that good patterns with improved dynamic characteristics would have symmetric grading about the mid-span point of the bar ($\hat{x} = 0.5$). Therefore, half of the bar can only be analyzed with the benefit of reducing the total number of the design variables, and consequently, the computational time.

5.1. *Piecewise models.* Considering first the case of fixed-fixed bar constructed from four segments with symmetry about the midspan, half of the bar can be analyzed using only four variables denoted by $(V_A, \hat{L})_{k=1,2}$. Furthermore, one of the segment lengths can be eliminated because of the equality constraint imposed on the total length. Another variable can also be discarded by applying the mass equality constraint (6), which reduces the number of design variables to only two, say $(V_A, \hat{L})_1$.

Figure 2 depicts the functional behavior of the dimensionless first and second frequencies augmented with the mass equality constraint. It is seen that the functions are well behaved and continuous everywhere in the design space $(V_A - \hat{L})_1$, except in the empty regions located at the upper right of the whole domain, where the mass equality constraint is violated. Such empty regions are bounded from below by the curve $(V_A \hat{L})_1 = 0.5$; that is the volume fraction of material (A) is equal to zero in the second segment: $V_{A,2} = 0$.

The level curves of the lowest fundamental frequency associated with the first symmetric mode are shown in on the left in Figure 2. The feasible domain is seen to be split by the baseline contours ($\hat{\omega}_1 = \pi$)
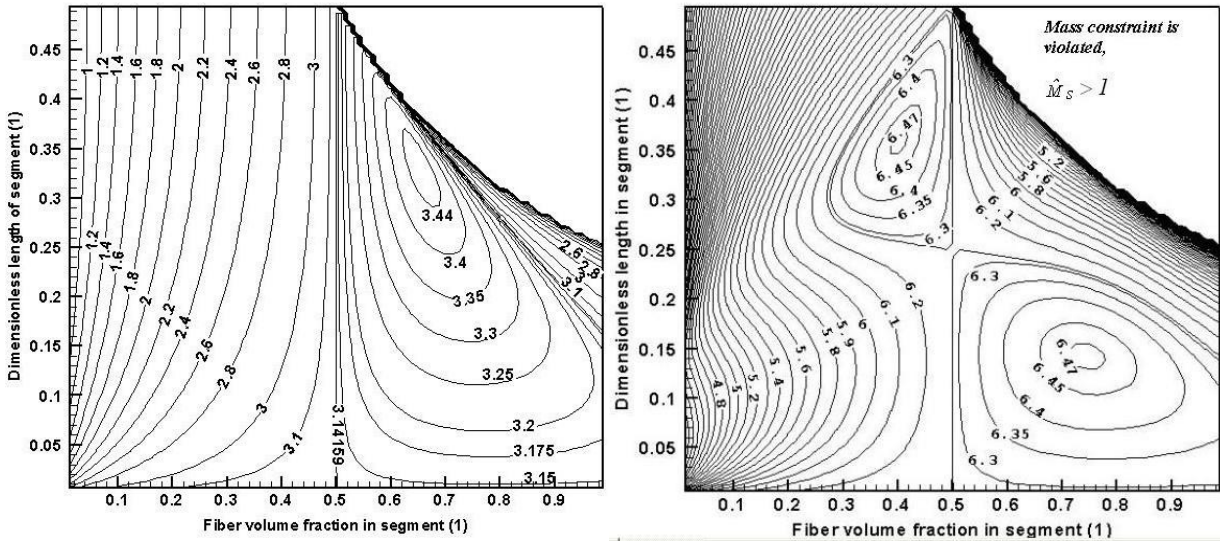
Figure 2. Isomerits of the dimensionless frequency under mass equality constraint of a 4-segment fixed-fixed bar, symmetric about midspan. Left: fundamental frequency (symmetric mode). Right: second frequency (asymmetric mode). Construction material is carbon fibers-AS4 / epoxy-3501-6.

into two distinct zones. The one to the right encompasses the constrained global maxima, which is calculated to be $\hat{\omega}_{1,max} = 3.45406$ at the optimal design point $(V_A, \hat{L})_{k=1,2} = (0.650, 0.3375), (0.1885, 0.1625)$. Actually, each design point inside the feasible domain corresponds to different material properties as well as different stiffness and mass distributions, while maintaining the total structural mass of the vibrating rod constant. Figure 2, right, shows the developed isomerits of the second frequency associated with the first asymmetric mode. Two global maxima can be observed having the calculated values $\hat{\omega}_{2,max} = 6.4732$ and 6.4711, which correspond, respectively, to the design points $(V_A, \hat{L})_{k=1,2} = (0.405, 0.365), (0.757, 0.135)$ located at the upper left region, and $(0.750, 0.145), (0.3975, 0.355)$ at the lower right region of the feasible domain. The developed isomerits for other cases of fixed-free and free-free bars made of carbon/epoxy composites are shown in Figures 3 and 4, respectively. It is seen that a bar with free-free boundary conditions behaves in the opposite trend to the case of fixed-fixed bar. The global maximum of the fundamental frequency in the former occurs at the lower region to the left of the design space (Figure 4) having the same value of $\hat{\omega}_{1,max} = 3.45406$ with altering segment locations, namely, $(V_A, \hat{L})_{k=1,2} = (0.1885, 0.1625), (0.650, 0.3375)$.

Table 3 summarizes the attained optimal solutions with increasing the number of segments for the different types of boundary conditions. It is seen that the attained optimization gain (i.e. the percentage increase in the fundamental frequency above its baseline value) increases with the number of segments. However, it should be kept in mind that the cost of manufacturing will also be increased. Therefore, a compromise has to be made between the reduction of vibration and the cost of manufacturing. In all, the bars can be economically built from a fewer number of segments having different length and material properties, since each pattern with a specified number of segments has its own exact global
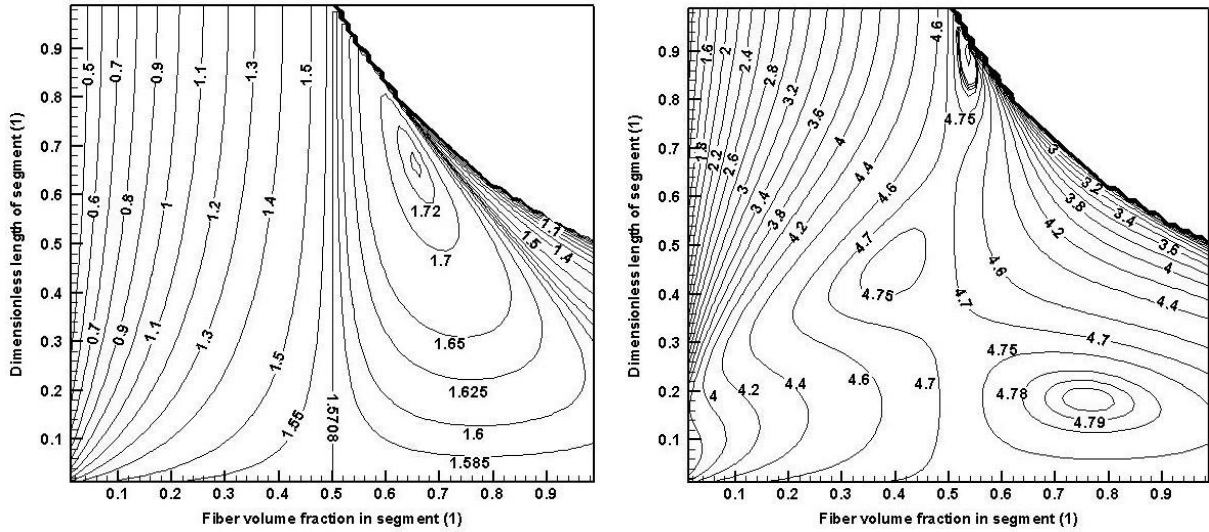
Figure 3. Frequency isomerits of a 2-segment, fixed-free bar under mass constraint ($\hat{M}_s = 1.0$): fundamental frequency (left) and second frequency (right). Material is carbon fibers-AS4 / epoxy-3501-6.
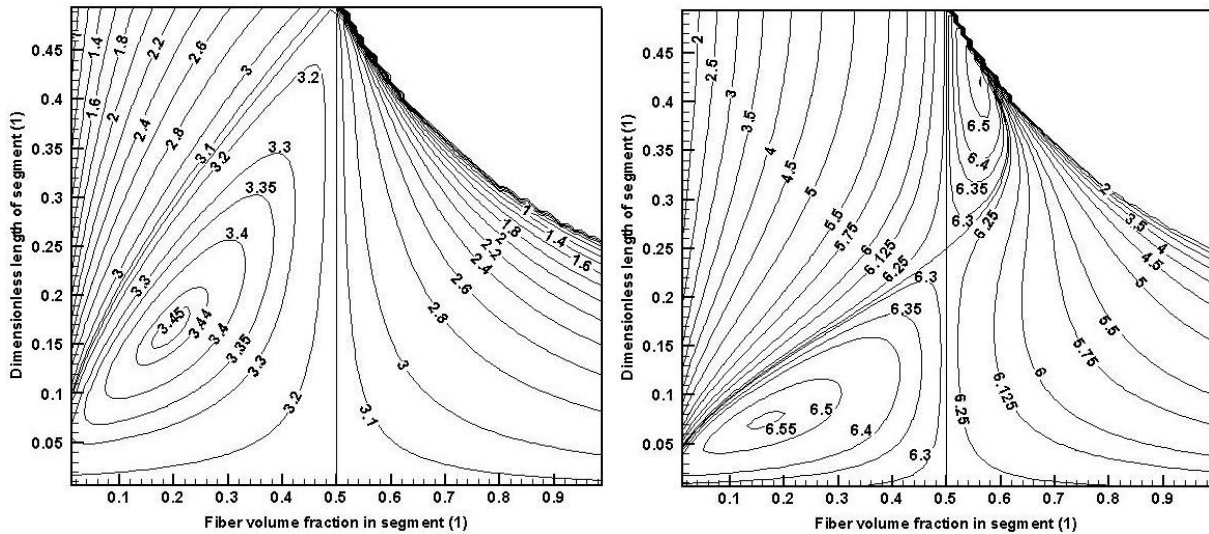
Figure 4. Frequency isomerits of a symmetric, 4-segment, free-free bar under mass constraint ($\hat{M}_s = 1.0$). Left: fundamental frequency (asymmetric mode). Right: second frequency (symmetric mode). Material is carbon fibers-AS4 / epoxy-3501-6.

optimal solution. One does not have to use more segments in order to increase the accuracy of the resulting solutions. This is one of the major outcomes of the present model formulation, which result in exact solutions no matter the number of segments is. Another remarkable observation is that the attained optimal designs based on maximization of a single frequency do not guarantee maximization of other

$N_s =$	1	2	3	4
fixed-fixed (half-span)	$(0.50, 0.50)_1$ (baseline)	$(0.650, 0.338)_1$ $(0.189, 0.163)_2$	$(0.719, 0.244)_1$ $(0.413, 0.144)_2$ $(0.138, 0.112)_3$	$(0.750, 0.23125)_1$ $(0.425, 0.13125)_2$ $(0.225, 0.07500)_3$ $(0.0625, 0.0625)_4$
$(\hat{\omega}_1, \hat{\omega}_2, \hat{\omega}_3)$	$(3.142, 6.283, 9.425)$	$(3.454, 5.537, 8.081)$	$(3.531, 5.487, 8.287)$	$(3.562, 5.133, 8.231)$
increase in $\hat{\omega}_1$	0%	9.93%	12.38%	13.37%
fixed-free	$(0.50, 1.00)_1$ (baseline)	$(0.650, 0.675)_1$ $(0.189, 0.325)_2$	$(0.725, 0.500)_1$ $(0.381, 0.306)_2$ $(0.107, 0.194)_3$	$(0.750, 0.4625)_1$ $(0.425, 0.2625)_2$ $(0.225, 0.1500)_3$ $(0.0625, 0.125)_4$
$(\hat{\omega}_1, \hat{\omega}_2, \hat{\omega}_3)$	$(1.571, 4.712, 7.854)$	$(1.727, 4.044, 7.354)$	$(1.766, 4.105, 6.552)$	$(1.781, 4.116, 6.454)$
increase in $\hat{\omega}_1$	0%	9.93%	12.41%	13.37%
free-free (half-span)	$(0.50, 0.50)_1$ (baseline)	$(0.200, 0.165)_1$ $(0.648, 0.335)_2$	$(0.138, 0.112)_1$ $(0.413, 0.144)_2$ $(0.719, 0.244)_3$	$(0.0625, 0.0625)_1$ $(0.2250, 0.0750)_2$ $(0.425, 0.13125)_3$ $(0.750, 0.23125)_4$
$(\hat{\omega}_1, \hat{\omega}_2, \hat{\omega}_3)$	$(3.142, 6.283, 9.425)$	$(3.454, 5.990, 8.170)$	$(3.531, 5.918, 8.287)$	$(3.562, 5.874, 8.231)$
increase in $\hat{\omega}_1$	0%	9.93%	12.38%	13.37%

Table 3. Optimal patterns $(V_A, \hat{L})_{k=1,2,\ldots,N_s}$ of multisegment FGM bars with maximized fundamental frequency $\hat{\omega}_1$ under mass equality constraint. Patterns for fixed-fixed and free-free bars are symmetric about the midpoint of the bar.

frequencies. For example, in the case of fixed-fixed bar, the design based upon maximization of $\hat{\omega}_2$ alone (see Figure 2) can result in a degraded value of the fundamental frequency ($\hat{\omega}_1 = 2.820$), located in the region to the left of the design space, which is lower than that of the baseline design by about 10.24%. Therefore, if one really seeks to maximize the overall stiffness-to-mass ratio of the vibrating rod, a multi-objective design optimization ought to be implemented instead [Kasprzak and Lewis 2001]. In this regard, the proper determination of the values of the weighting factors α_i ought to be based on the fact that each frequency shall be maximized from its initial value corresponding to the uniform baseline design. The author suggests the following procedure:

- Initialize the values of the weighting factors by the reciprocal of the initial reference values; $\alpha_i = 1/\hat{\omega}_{0,i}$.
- Normalize the resulting values by dividing each one by their sum in order to make the final sum equal to 1.0, i.e. $\alpha_i \leftarrow \alpha_i / \sum_i \alpha_i$.

Considering only the first three frequencies, Table 4 gives the appropriate values of the weighting factors corresponding to the different types of boundary conditions. The attained optimal solutions for different optimization strategies are given in Table 5 for bars constructed from two segments.

	α_1	α_2	α_3
fixed-fixed	54.5%	27.3%	18.2%
fixed-free	65.2%	21.7%	13.1%
free-free	54.5%	27.3%	18.2%

Table 4. Appropriate values of the weighting factors for different boundary conditions.

	max $\hat{\omega}_1$	max $\hat{\omega}_2$	max $\sum_{i=1}^{3} \alpha_i \hat{\omega}_i$
fixed-fixed	$(0.6500, 0.3375)_1$	$(0.405, 0.365)_1$	$(0.750, 0.145)_1$
(half length)	$(0.1885, 0.1625)_2$	$(0.757, 0.135)_2$	$(0.398, 0.355)_2$
$(\hat{\omega}_1, \hat{\omega}_2, \hat{\omega}_3)$	$(3.454, 5.537, 8.081)$	$(2.820, 6.473, 8.914)$	$(3.289, 6.471, 9.468)$
fixed-free	$(0.6500, 0.675)_1$	$(0.538, 0.90)_1$	$(0.5315, 0.9125)_1$
	$(0.1885, 0.325)_2$	$(0.163, 0.10)_2$	$(0.1715, 0.0875)_2$
$(\hat{\omega}_1, \hat{\omega}_2, \hat{\omega}_3)$	$(1.727, 4.041, 7.352)$	$(1.636, 4.845, 7.834)$	$(1.625, 4.839, 7.922)$
free-free	$(0.200, 0.165)_1$	$(0.560, 0.425)_1$	$(0.163, 0.075)_1$
(half length)	$(0.648, 0.335)_2$	$(0.163, 0.075)_2$	$(0.560, 0.425)_2$
$(\hat{\omega}_1, \hat{\omega}_2, \hat{\omega}_3)$	$(3.454, 5.990, 8.170)$	$(2.537, 6.557, 8.502)$	$(3.335, 6.557, 9.526)$

Table 5. Optimal patterns $(V_A, \hat{L})_{k=1,2}$ of a two-segment FGM bar for different optimization strategies. Patterns for fixed-fixed and free-free bars are symmetric about the middle of bar length.

More calculations by implementation of the models described by (18) and (19) have revealed the fact that maximization of the natural frequencies is a much better design criterion than the frequency placement criterion. The latter resulted in a slightly time consuming optimization process and a slow rate of convergence towards the optimum solution. Such a quadratic formulation complicates matters by increasing nonlinearities in the design space and, therefore, difficult to minimize globally. Moreover, the determination of the target frequencies in most cases is not easy and the resulting solutions are greatly dependent on the choice of their values. Direct maximization of the natural frequencies without regard to frequency-windows was found to be more significant and best representative to the overall dynamic performance. If it happened that the maximum frequency violates frequency windows, which was found to be a rare situation, another value of the frequency can be chosen near the global optima, and the frequency equation can be solved via an inverse approach for any one of the unknown design variables instead.

5.2. *Continuous models.* Linear and parabolic models for material grading along the bar span (Figure 5) have been examined. The attained optimal solutions for both cases are given in Table 6 for the different types of boundary conditions. Two consecutive solutions are given for each case; the first solution is based on maximization of the fundamental frequency alone, while the second is based on maximization of a weighted sum of the first three frequencies. Upper and lower limits were imposed on the volume fraction of material $(A)(0.0 \le V_A \le 1.0)$ in order to avoid having unrealistic configurations in the final
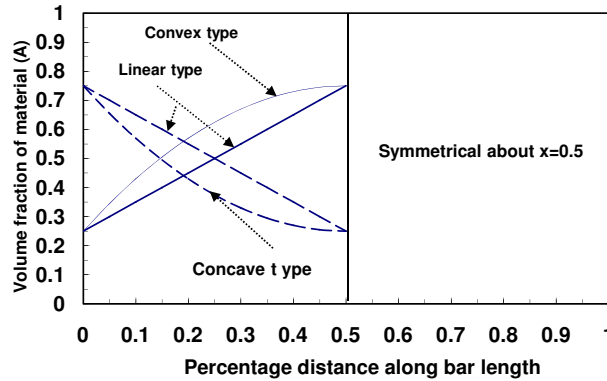
Figure 5. Symmetric shape models of volume fraction distribution along bar length.

Volume fraction distribution		$(\hat{\omega}_1, \hat{\omega}_2, \hat{\omega}_3)$
fixed-fixed ($0 < \hat{x} < 0.5$)		
Linear:	$V_A(\hat{x}) = 1.0 - 2\hat{x}$	(3.5922, 4.9481, 8.3353)
	$V_A(\hat{x}) = 0.75 - \hat{x}$	(3.4416, 6.0931, 9.3144)
Parabolic (concave):	$V_A(\hat{x}) = 1.0 + 3(\hat{x}^2 - \hat{x})$	(3.4706, 6.1781, 9.0588)
	$V_A(\hat{x}) = 0.75 + 1.5(\hat{x}^2 - \hat{x})$	(3.3566, 6.3231, 9.3781)
fixed-free ($0 < \hat{x} < 1.0$)		
Linear:	$V_A(\hat{x}) = 1.0 - \hat{x}$	(1.7961, 4.1677, 6.6348)
	$V_A(\hat{x}) = 0.75 - 0.5\hat{x}$	(1.7208, 4.6572, 7.6833)
Parabolic (convex):	$V_A(\hat{x}) = 0.75(1 - \hat{x}^2)$	(1.7928, 4.4189, 7.0611)
	$V_A(\hat{x}) = 0.625 - 0.375\hat{x}^2$	(1.6958, 4.6936, 7.7527)
Parabolic (concave):	$V_A(\hat{x}) = 1.0 + 0.75(\hat{x}^2 - 2\hat{x})$	(1.7353, 4.5294, 7.4584)
	$V_A(\hat{x}) = 0.75 + 0.375(\hat{x}^2 - 2\hat{x})$	(1.6781, 4.6891, 7.7692)
free-free ($0 < \hat{x} < 0.5$)		
Linear:	$V_A(\hat{x}) = 2\hat{x}$	(3.5922, 5.7091, 8.3353)
	$V_A(\hat{x}) = 0.25 + \hat{x}$	(3.4416, 6.1841, 9.3144)
Parabolic (convex):	$V_A(\hat{x}) = 3(\hat{x} - \hat{x}^2)$	(3.5856, 6.2167, 8.8378)
	$V_A(\hat{x}) = 0.25 - 1.5(\hat{x}^2 - \hat{x})$	(3.3916, 6.3588, 9.3872)

Table 6. Optimal constant-mass patterns with linear and parabolic distributions of $V_A(\hat{x})$ for different types of boundary conditions. Patterns with fixed-fixed or free-free boundary conditions are symmetric about the middle of bar length.

attained solutions. In all cases the total structural mass was maintained constant by imposing the equality constraint described by (6). As a general observation, for fixed-fixed and fixed-free boundary conditions, patterns with higher fiber volume fraction near the fixed ends are always favorable. The opposite trend is true for cases of free-free bars. Maximization of the fundamental frequency alone produces an optimization gain of about 14.33% for the linear model with 0% and 100% volume fractions at the ends of the optimized bars with different boundary conditions. However, a drastic reduction in the second and third

frequencies can be observed. Better solutions have been achieved by maximizing a weighted sum of the first three frequencies, where the parabolic model was found to excel the linear one in producing balanced improvements in all frequencies. Results have also indicated that the fixed-fixed bars are recommended to have concave distribution rather than convex one. The latter produce poor patterns with degraded stiffness-to-mass ratio levels. The opposite trend was observed for the free-free bars, where the convex type is much more favorable than the concave type. Both concave and convex shapes can be accepted for a cantilevered bar. Finally, it should be pointed out here that, for certain ranges of the design variables many terms ought to be taken in the assumed power series (see Section 3.1) in order to achieve convergence. The proper number of terms to be taken in the series needs to be investigated.

6. Conclusions

In view of the importance of enhancing the dynamic performance and raising the overall (stiffness/mass) level of a FGM bar in axial motion, appropriate optimization models have been formulated for both continuous and discrete distributions of the volume fractions of the selected composite material. Several forms of the objective function have been examined by maximizing the natural frequencies while preserving the total structural mass at a constant value equals to that of a known baseline design. It was found that maximization of the fundamental frequency alone for any type of boundary conditions does not guarantee maximization of other higher frequencies. A local maximum for any one of the frequencies can correspond to a local minimum for another frequency. A weighted-sum formulation must, therefore, be considered in order to achieve a simultaneous balanced improvement among all frequencies. The weighting factors are best determined from the fact that each frequency ought to be maximized from its initial reference value of the uniform baseline design. Moreover, computer experiments have revealed that maximization of the natural frequencies is a much better design criterion than the frequency placement criterion. The attained optimal solutions in the latter were found to be too sensitive to the selected target frequencies, which are not easy to determine. If it happened that any one of the maximized frequencies violates frequency windows, which was found to be a rare situation, another value of the frequency can be chosen near the global optima, and an inverse approach can be utilized by solving the associated frequency equation for any one of the unknown design variables instead. Optimization of multisegment fixed-fixed or free-free bars has indicated that good patterns should be symmetric about the midspan of the bar. The given exact structural analysis leads to the exact frequencies no matter the number of segments is. The optimized bars can be fabricated economically from any arbitrary number of uniform segments with material grading in the axial direction. The increase in the number of segments would naturally result in an increase in the attained maximized frequencies, but care ought to be taken regarding the penalty of increasing the production and manufacturing costs. It has been confirmed that the segment length is most significant design variable in the whole optimization process. Some investigators who apply finite elements have not recognized that the length of each element can be taken as a main design variable in the whole set of optimization variables. It has also been shown that normalization of all terms results in a naturally scaled objective function, constraints and design variables, which is recommended when applying different optimization techniques. The results from the present approach reveals that piecewise grading of the material can be promising producing truly efficient bar designs with enhanced dynamic performance. In conclusion, a powerful design tool has been obtained by formulating an appropriate

objective function and applying mathematical programming techniques to the resulting optimization problem. It is the author's wish that the results presented in this paper will be compared and validated through other optimization techniques such as genetic algorithms or any appropriate global optimization algorithm. Future potential areas such as optimization of FGM beams in bending and torsion shall be formulated and identified. Other secondary effects such as material and geometric nonlinearities due to large deformation shall be investigated in future studies.

A personal note

Charles and Marie-Louise Steele hold a special place in my heart for their support and encouragement of my research in structural mechanics. In 2000, as editor of the *International Journal of Solids and Structures*, Marie-Louise encouraged me to submit there my "Buckling optimization of flexible columns". I was initially discouraged, but she pushed me to improve the paper, which was eventually accepted after extensive revision and published in 2002, outlining a new approach to the problem under consideration. She was likewise supportive after she and Charles founded JoMMS, where I published papers in 2007 and 2008. Early in 2009, she encouraged me to apply for the Fulbright research program at Stanford; we had been planning to collaborate in the academic year 2009/2010 on a research project about wind turbine blades made of functionally graded material.

Marie-Louise was always helping people to an uncommon degree. I believe that her passing away has been a great loss not only to her family and friends but to the research community. God rest her soul.

References

[Brach 1968] R. M. Brach, "On the extremal fundamental frequencies of vibrating beams", *Int. J. Solids Struct.* **4**:7 (1968), 667–674.

[Cheng and Batra 2000] Z.-Q. Cheng and R. C. Batra, "Exact correspondence between eigenvalues of membranes and functionally graded simply supported polygonal plates", *J. Sound Vib.* **229**:4 (2000), 879–895.

[Daniel and Ishai 2006] I. M. Daniel and O. Ishai, *Engineering mechanics of composite materials*, 2nd ed., Oxford University Press, New York, 2006.

[Edwards and Penney 2004] C. H. Edwards and D. E. Penney, *Differential equations and boundary value problems: computing and modeling*, 3rd ed., Pearson Education, Upper Saddle River, NJ, 2004.

[Elishakoff and Guédé 2004] I. Elishakoff and Z. Guédé, "Analytical polynomial solutions for vibrating axially graded beams", *Mech. Adv. Mater. Struct.* **11**:6 (2004), 517–533.

[Elishakoff et al. 2005] I. Elishakoff, C. Gentilini, and E. Viola, "Three-dimensional analysis of an all-round clamped plate made of functionally graded materials", *Acta Mech.* **180** (2005), 21–36.

[Goupee and Vel 2006] A. J. Goupee and S. S. Vel, "Optimization of natural frequencies of bidirectional functionally graded beams", *Struct. Multidiscip. Optim.* **32**:6 (2006), 473–484.

[Halpin and Tsai 1969] J. C. Halpin and S. W. Tsai, "Effects of environmental factors on composite materials", Technical report AFML-TR-67-423, Air Force Materials Laboratory, Wright-Patterson Air Force Base, Dayton, OH, 1969, Available at http://handle.dtic.mil/100.2/ADA306357.

[Kasprzak and Lewis 2001] E. M. Kasprzak and K. E. Lewis, "Pareto analysis in multiobjective optimization using the collinearity theorem and scaling method", *Struct. Multidiscip. Optim.* **22**:3 (2001), 208–218.

[Kumar and Sujith 1997] B. M. Kumar and R. I. Sujith, "Exact solutions for the longitudinal vibration of non-uniform rods", *J. Sound Vib.* **207**:5 (1997), 721–729.

[Li 2000] Q. S. Li, "Free longitudinal vibration analysis of multi-step non-uniform bars based on piecewise analytical solutions", *Eng. Struct.* **22**:9 (2000), 1205–1215.

[Librescu and Maalawi 2007] L. Librescu and K. Y. Maalawi, "Material grading for improved aeroelastic stability in composite wings", *J. Mech. Mater. Struct.* **2**:7 (2007), 1381–1394.

[Loy et al. 1999] C. T. Loy, K. Y. Lam, and J. N. Reddy, "Vibration of functionally graded cylindrical shells", *Int. J. Mech. Sci.* **41**:3 (1999), 309–324.

[Maalawi 1999] K. Y. Maalawi, "Optimal mass and frequency design of rods in longitudinal vibrations", pp. 259–273 in *6th Conference on Theoretical and Applied Mechanics* (Cairo, 1999), Academy of Scientific Research and Technology, Cairo, 1999.

[Maalawi and El-Chazly 2002] K. Y. Maalawi and N. M. El-Chazly, "Global optimization of multi-element beam-type structures", in *Proceedings of the 2nd International Conference on Advances in Structural Engineering and Mechanics* (*ASEM'02*) (Busan, 2002), edited by C.-K. Choi and W. C. Schnobrich, Techno-Press, Daejeon, 2002.

[Maalawi and Negm 2002] K. Y. Maalawi and H. M. Negm, "Optimal frequency design of wind turbine blades", *J. Wind Eng. Ind. Aerod.* **90**:8 (2002), 961–986.

[Maalawi and Warner 1984] K. Y. Maalawi and W. H. Warner, "Global optimization of multi-segment rods in axial motion", Technical report AEM 8880, Department of Aerospace Engineering and Mechanics, University of Minnesota, 1984. Plan B Project.

[Masad 1997] J. A. Masad, "Optimization methods with structural dynamics applications", *Comput. Struct.* **62**:3 (1997), 521–525.

[Meirovitch 1997] L. Meirovitch, *Principles and techniques of vibrations*, Prentice Hall, Upper Saddle River, NJ, 1997.

[Niordson 1965] F. I. Niordson, "On the optimal design of a vibrating beam", *Quart. Appl. Math.* **23**:1 (1965), 47–53.

[Pestel and Leckie 1963] E. C. Pestel and F. A. Leckie, *Matrix methods in elastomechanics*, McGraw-Hill, New York, 1963.

[Qian and Batra 2005] L. F. Qian and R. C. Batra, "Design of bidirectional functionally graded plate for optimal natural frequencies", *J. Sound Vib.* **280**:1–2 (2005), 415–424.

[Tylikowski 2005] A. Tylikowski, "Dynamic stability of functionally graded plate under in-plane compression", *Math. Probl. Eng.* **2005**:4 (2005), 411–424.

[Vanderplaats 1999] G. Vanderplaats, *Numerical optimization techniques for engineering design: with applications*, McGraw-Hill, New York, 1999.

[Vel and Batra 2004] S. S. Vel and R. C. Batra, "Three-dimensional exact solution for the vibration of functionally graded rectangular plates", *J. Sound Vib.* **272**:3–5 (2004), 703–730.

[Venkataraman 2002] P. Venkataraman, *Applied optimization with MATLAB programming*, Wiley, Hoboken, NJ, 2002.

[Warner 2001] W. H. Warner, "Optimal design of elastic rods: extension of a minimum energy solution", *Int. J. Solids Struct.* **38**:16 (2001), 2879–2891.

[Warner and Vavrick 1975] W. H. Warner and D. J. Vavrick, "Optimal design in axial motion for several frequency constraints", *J. Optim. Theory Appl.* **15**:1 (1975), 157–166.

KARAM Y. MAALAWI: nrc.aero@gmail.com
Department of Mechanical Engineering, National Research Center, 33 El Behous Street, Dokki, Cairo 12622, Egypt
http://www.nrc.sci.eg

■
■msp

BASIC STRAIN GRADIENT PLASTICITY THEORIES
WITH APPLICATION TO CONSTRAINED FILM DEFORMATION

Christian F. Niordson and John W. Hutchinson

A family of basic rate-independent strain gradient plasticity theories is considered that generalize conventional J_2 deformation and flow theories of plasticity to include a dependence on strain gradients in a simple way. The theory builds on three recent developments: the work of Gudmundson (*J. Mech. Phys. Solids* **52** (2004), 1379–1406) and Gurtin and Anand (*J. Mech. Phys. Solids* **57** (2009), 405–421), proposing constitutive relations for flow theories consistent with requirements of positive plastic dissipation; the work of Fleck and Willis (*J. Mech. Phys. Solids* **57** (2009), 161–177 and 1045–1057), who clarified the structure of the new flow theories and presented the underlying variational formulation; and observations of Evans and Hutchinson (*Acta Mater.* **57** (2009), 1675–1688) related to preferences for specific functional compositions of strains and strain gradients. The starting point in this paper is the deformation theory formulation of Fleck and Hutchinson (*J. Mech. Phys. Solids* **49** (2001), 2245–2271) which provides the clearest insights into the role of strain gradients and serves as a template for the flow (incremental) theory. The flow theory is constructed such that it coincides with the deformation theory under proportional straining, analogous to the corresponding coincidence in the conventional J_2 theories. The generality of proportional straining is demonstrated for pure power-law materials, and the utility of power-law solutions is illustrated for the constrained deformation of thin films: the compression or extension of a finite layer joining rigid platens. Full elastic-plastic solutions are obtained for the same problem based on a finite element method devised for the new class of flow theories. Potential difficulties and open issues associated with the new class of flow theories are identified and discussed.

1. Introduction

Both [Gudmundson 2004] and [Gurtin and Anand 2009] have called attention to the fact that the phenomenological flow (incremental) strain gradient plasticity theory of [Fleck and Hutchinson 2001] does not guarantee that plastic dissipation will be positive for all straining histories. The advantage of the Fleck–Hutchinson theory is that it generalizes the most widely used conventional plasticity theories, J_2 deformation and flow theories, in a straightforward way that can be readily implemented numerically. In its simplest form, the theory introduces only a single new material length parameter with no other inputs than those of conventional J_2 theory. By invoking viscous plasticity, [Gudmundson 2004] and [Gurtin and Anand 2009] suggest ways that the flow can be modified such that the thermodynamic requirement that

plastic dissipation be nonnegative is met. Following up on these works, Fleck and Willis [2009a; 2009b] have proposed specific forms of the constitutive law for both viscous and rate-independent materials. The simplest version proposed by Fleck and Willis coincides with the original theory of [Fleck and Hutchinson 2001] under proportional straining, and again introduces only a single length parameter. In addition, Fleck and Willis call attention to the fact that the new form of the constitutive relation for the rate-independent flow theory has close parallels to aspects of conventional rigid plasticity in the sense that some of the higher-order stress quantities are not determined by prior history but depend on the current strain rates. Most importantly, Fleck and Willis establish variational principles for the new flow theory formulations.

In this paper, a unified family of small-strain, rate-independent, gradient plasticity theories will be adopted from [Fleck and Willis 2009a; 2009b] which meet the thermodynamic requirements detailed in [Gurtin and Anand 2009]. Consideration of the extended family is motivated by a recent assessment of basic aspects of strain gradient theories [Evans and Hutchinson 2009] which examined the manner in which plastic strain gradients are incorporated into the theories in light of experimental trends. Specifically, with ε_P as a measure of the amplitude of the plastic strain and ε_P^* as a positive measure of the plastic strain gradients, Evans and Hutchinson assessed the family of generalized effective plastic strains defined by

$$E_P = \left((\varepsilon_P)^\mu + (\ell \varepsilon_P^*)^\mu \right)^{1/\mu}, \tag{1-1}$$

as introduced originally in [Fleck and Hutchinson 1997]. The material length parameter, ℓ, arises due to dimensional consistency.

The composition with $\mu = 2$,

$$E_P = \sqrt{(\varepsilon_P)^2 + (\ell \varepsilon_P^*)^2}, \tag{1-2}$$

has generally been preferred, primarily for mathematical reasons, as in [Fleck and Hutchinson 2001] and also in the new formulation of [Fleck and Willis 2009a]. A quadratic dependence on ℓ is also implicit in the various formulations of [Gudmundson 2004; Fredriksson and Gudmundson 2005; Gurtin and Anand 2009]. The argument for the linear composition ($\mu = 1$),

$$E_P = \varepsilon_P + \ell \varepsilon_P^*, \tag{1-3}$$

rests on the fact that it leads to gradient effects proportional to ℓ/h for a sequence of objects of increasing size h as the conventional limit is approached. This trend in experimental indentation hardness data has been noted for many metals, and has been invoked in the constitutive equation of [Nix and Gao 1998]. By contrast, compositions depending quadratically on $\ell \varepsilon_P^*$, as in (1-2), predict gradient effects proportional to $(\ell/h)^2$ as the conventional limit is approached for many problems such as indentation, and would appear unable to capture experimental trends as well as the linear composition as the conventional limit is approached.

This paper takes as its starting point the postulation of the deformation theory of plasticity (that is, effectively, a strain gradient theory of small-strain nonlinear elasticity) as proposed in [Fleck and Hutchinson 2001]. Then, following that paper, the flow theory is constructed to coincide with the deformation theory for proportional straining. This approach unifies the theories, as in the conventional J_2 theories, and illuminates the role gradients play in the clearest possible manner. In addition, it provides a simpler framework for solving problems when use of deformation theory is justified, and underpins theoretical

developments such as the J-integral of crack mechanics which rely on the existence of a deformation theory. The objective here continues to be a theory generalizing conventional rate-independent J_2 theory in the simplest meaningful way, building, in particular, on [Fleck and Hutchinson 2001] and the more recent developments in [Fleck and Willis 2009a; 2009b]. As in [Idiart et al. 2009], the homogeneous composition of the generalized effective plastic strain in (1-1) will be considered with μ regarded as having a fixed value based on mechanistic considerations such as those discussed in [Evans and Hutchinson 2009]. A parallel approach developed for an alternative effective strain composition favored by Nix and Gao [1998] is given in the Appendix.

Following introduction of the underlying theory, application to pure power-law materials is discussed. Deformation theory and flow theory solutions for these materials coincide because proportional straining is precisely satisfied. Compression or separation of a finite-length thin film bonding rigid platens is analyzed for the insights it provides into strain gradient effects on thin metallic bonding layers and as a possible test configuration for obtaining material data. Deformation and flow theory solutions will be generated, illustrating their correspondence and revealing the role of the material length parameter and higher-order boundary conditions. In solving specific problems, some seemingly anomalous predictions have been found for flow theory. These issues are addressed at the end of the paper.

2. Preliminaries

The paper lays out a small-strain, isotropic, rate-independent phenomenological theory. The notation and theoretical framework is similar to that in [Fleck and Hutchinson 2001]. The conventional Cauchy stress is denoted by σ_{ij}, its deviator stress by s_{ij}, the total strain by $\varepsilon_{ij} = \frac{1}{2}(u_{i,j} + u_{j,i})$ with u_i as the displacement, and the elastic strain by $\varepsilon_{ij}^e = \varepsilon_{ij} - \varepsilon_{ij}^P$ with ε_{ij}^P as the plastic strain. The conventional effective stress is defined by $\sigma_e = \sqrt{3 s_{ij} s_{ij}/2}$, and a dimensionless deviator tensor codirectional to the deviator stress is defined by $m_{ij} = (3/2)s_{ij}/\sigma_e$. The stress is given by $\sigma_{ij} = \mathcal{L}_{ijkl}\varepsilon_{kl}^e$, with isotropic elastic moduli determined by the Young's modulus, E, and Poisson's ratio, ν. The other input to conventional J_2 theory is the uniaxial tensile stress-strain curve. Denote the relation between the stress and the plastic strain in uniaxial tension by $\sigma(\varepsilon_P)$. Furthermore, denote the plastic work density to deform a material element in uniaxial tension to plastic strain ε_P by

$$U_P(\varepsilon_P) = \int_0^{\varepsilon_P} \sigma(\varepsilon_P)\, d\varepsilon_P. \tag{2-1}$$

2A. *Measures of plastic strain and plastic strain gradient in the deformation theories.* With

$$\varepsilon_{ij}^P = \varepsilon_{ij} - \varepsilon_{ij}^e \qquad \text{and} \qquad \varepsilon_P = \sqrt{2\varepsilon_{ij}^P \varepsilon_{ij}^P/3}$$

as the plastic strain amplitude, the plastic strain is required to be codirectional with the stress deviator such that $\varepsilon_{ij}^P = \varepsilon_P m_{ij}$. The *one-parameter* measure of plastic strain gradient,

$$\varepsilon_P^* = \sqrt{\varepsilon_{P,i}\varepsilon_{P,i}}, \tag{2-2}$$

is the same as that employed in [Aifantis 1984; Mühlhaus and Aifantis 1991]. More general measures based on the three quadratic invariants of the plastic strain gradient, $\varepsilon_{ij,k}^P = \varepsilon_{P,k} m_{ij} + \varepsilon_P m_{ij,k}$, and involving three length parameters, ℓ_i ($i = 1, 2, 3$) have been considered in [Fleck and Hutchinson 2001;

Fleck and Willis 2009a], but here attention is confined to the one-parameter measure (2-2) which will be referred to as the *scalar measure* following the terminology of Fleck and Willis. A one-parameter *tensor measure*, in the Fleck–Willis terminology, is

$$\varepsilon_P^* = \sqrt{2\varepsilon_{ij,k}^P \varepsilon_{ij,k}^P / 3}. \tag{2-3}$$

The definition in (2-3) is such that it coincides with (2-2) for a shearing with only one nonzero gradient, for example, $\varepsilon_{12,2}^P = \varepsilon_{21,2}^P$. For other strain gradients the two measures will generally not coincide. For both (2-2) and (2-3), the generalized effective plastic strain, E_P, is defined by (1-1).

The length parameter must be determined by experiment, as for the other material properties. The value of the parameter depends on the specific theory, as discussed in some detail in [Evans and Hutchinson 2009], and possibly on the specific experiment. The 3-parameter versions of strain gradient plasticity have greater flexibility in accurately encompassing a wide array of problems, but that is not the main concern here.

2B. *Measures of plastic strain and plastic strain gradient in the flow theories.* The fundamental measures for flow theory are defined for increments (rates), otherwise, they are similar to those above. Specifically, with $\dot{\varepsilon}_{ij}^P = \dot{\varepsilon}_{ij} - \dot{\varepsilon}_{ij}^e$, $\dot{\varepsilon}_P = \sqrt{2\dot{\varepsilon}_{ij}^P \dot{\varepsilon}_{ij}^P / 3}$, and with a plastic strain rate codirectional with the stress deviator, $\dot{\varepsilon}_{ij}^P = \dot{\varepsilon}_P m_{ij}$. For the one-parameter scalar version, the gradient rate measure is

$$\dot{\varepsilon}_P^* = \sqrt{\dot{\varepsilon}_{P,i} \dot{\varepsilon}_{P,i}} \tag{2-4}$$

and for the one-parameter tensor version it is

$$\dot{\varepsilon}_P^* = \sqrt{\tfrac{2}{3}\dot{\varepsilon}_{ij,k}^P \dot{\varepsilon}_{ij,k}^P}. \tag{2-5}$$

The generalized effective plastic strain rate is taken as

$$\dot{E}_P = \left((\dot{\varepsilon}_P)^\mu + (\ell\dot{\varepsilon}_P^*)^\mu\right)^{1/\mu}. \tag{2-6}$$

Current values of the plastic strain measures are integrals over the history:

$$\varepsilon_P = \int \dot{\varepsilon}_P \quad \text{and} \quad E_P = \int \dot{E}_P. \tag{2-7}$$

Throughout this paper, the notion of proportional straining requires that the plastic strain, ε_{ij}^P, and its gradient, $\varepsilon_{ij,k}^P$, change monotonically and proportionally. For proportional straining the definitions for the flow theories coincide with those for the respective deformation theories.

It is important to be cognizant of restrictions inherent to the effective strain measures ε_P and E_P in the flow theory. Both measures are nondecreasing since their rates are intrinsically positive. The theories proposed below are similar to conventional J_2 theory in that they take hardening and the yield surface to depend on the total plastic strain. The theory invokes an isotropic expansion of the yield surface which should not be expected to correctly reproduce histories involving significant stress and strain reversal. As with conventional J_2 theory, the strain gradient version proposed here is not intended for application to problems involving significant stress reversal without due consideration to yield surface shape changes.

2C. *The principle of virtual work and the equilibrium equations.* The principle of virtual work for the scalar theory is [Fleck and Hutchinson 2001]

$$\int_V (\sigma_{ij}(\delta\varepsilon_{ij} - \delta\varepsilon_P m_{ij}) + Q\delta\varepsilon_P + \tau_i\delta\varepsilon_{P,i})dV = \int_S (T_i\delta u_i + t\delta\varepsilon_P)dS \qquad (2\text{-}8)$$

for all admissible δu_i and $\delta\varepsilon_P$. The associated equilibrium equations are

$$\sigma_{ij,j} = 0, \qquad \tau_{i,i} - Q + \sigma_e = 0, \qquad (2\text{-}9)$$

and the boundary data pairs are (with n_i as the surface normal)

$$(T_i = \sigma_{ij}n_j, u_i) \qquad \text{and} \qquad (t = \tau_i n_i, \varepsilon_P). \qquad (2\text{-}10)$$

In conventional theory, $Q = \sigma_e$, but in the higher-order theory Q and τ_i are new stress variables that are work conjugate to ε_P and $\varepsilon_{P,i}$, respectively. With stress quantities replaced by their increments, (2-8)–(2-10) apply for incremental equilibrium.

The corresponding principle of virtual work for the tensor formulation is

$$\int_V \left(\sigma_{ij}(\delta\varepsilon_{ij} - \delta\varepsilon_{ij}^P) + Q_{ij}\delta\varepsilon_{ij}^P + \tau_{ijk}\delta\varepsilon_{ij,k}^P\right)dV = \int_S (T_i\delta u_i + t_{ij}\delta\varepsilon_{ij}^P)dS, \qquad (2\text{-}11)$$

with equilibrium equations

$$\sigma_{ij,j} = 0, \qquad \tau_{ijk,k} - Q_{ij} + s_{ij} = 0, \qquad (2\text{-}12)$$

and boundary pairs

$$(T_i = \sigma_{ij}n_j, u_i), \qquad (t_{ij} = \tau_{ijk}n_k, \varepsilon_{ij}^P). \qquad (2\text{-}13)$$

The reader is referred to [Fleck and Willis 2009b] for complete details of the tensor version which parallel those for the scalar version. In what follows details will only be presented for the scalar version, but selected results for the tensor version will be given.

2D. *The role of the gradients of plastic strain in the theory.* The physical interpretation underlying E_P in (1-1) is that it represents the collective sum of the movement of statistically stored dislocations tied to ε_P and of the geometrically necessary dislocations associated with ε_P^*. The specific composition of E_P is phenomenological, but it is intended to measure the history of all the dislocation motion. It reduces to the conventional measure when the gradients are small. In conventional J_2 deformation theory the energy density of a material element following proportional straining is the sum of elastic and "plastic" parts according to $\frac{1}{2}\mathscr{L}_{ijkl}\varepsilon_{ij}^e\varepsilon_{kl}^e + U_P(\varepsilon_P)$ with U_P given by (2-1). Following [Fleck and Hutchinson 2001], the energy density for strain gradient deformation plasticity is taken as $\frac{1}{2}\mathscr{L}_{ijkl}\varepsilon_{ij}^e\varepsilon_{kl}^e + U_P(E_P)$. The replacement of ε_P by E_P in $U_P(\varepsilon_P)$ reveals the essence of the plastic strain gradient in this family of theories. In words, the plastic work needed to deform the material element in the presence of strain gradients under proportional straining as measured by E_P in (1-1) is taken equal to that at the same strain, $\varepsilon_P = E_P$, in the absence of gradients, consistent with the notion that E_P and ε_P measure the dislocation motion under the two circumstances.

3. The scalar theory

3A. *Deformation theory.* The potential energy functional for the deformation theory is

$$\Phi(\boldsymbol{u}, \varepsilon_P) = \int_V \left(\tfrac{1}{2}\mathscr{L}_{ijkl}\varepsilon_{ij}^e\varepsilon_{kl}^e + U_P(E_P)\right)dV - \int_{S_T}(T_i u_i + t\varepsilon_P)dS, \qquad (3\text{-}1)$$

with $\varepsilon_{ij}^e = \varepsilon_{ij} - \varepsilon_P m_{ij}$ and where T_i and t are prescribed on S_T. A solution minimizes the functional with respect to admissible $\boldsymbol{u}$ and ε_P assuming the tensile stress-strain curve, $\sigma(\varepsilon_P)$, defining U_P in (2-1) is monotone. The higher-order stresses are

$$Q = \frac{\partial U_P}{\partial \varepsilon_P} = \sigma(E_P)\frac{(\varepsilon_P)^{\mu-1}}{(E_P)^{\mu-1}} \qquad \text{and} \qquad \tau_i = \frac{\partial U_P}{\partial \varepsilon_{P,i}} = \sigma(E_P)\frac{\ell^2(\ell\varepsilon_P^*)^{\mu-2}\varepsilon_{P,i}}{(E_P)^{\mu-1}}. \qquad (3\text{-}2)$$

One can easily show that the following generalized effective stress satisfies

$$\Sigma \equiv \left(Q^{\mu/(\mu-1)} + (\tau/\ell)^{\mu/(\mu-1)}\right)^{(\mu-1)/\mu} = \sigma(E_P), \qquad (3\text{-}3)$$

with $\tau = \sqrt{\tau_i \tau_i}$.

3B. *Flow theory.* With rates and integrated quantities as defined in Section 2B, the stress quantities Q and τ_i are chosen such that they coincide with those of deformation theory in (3-2) when the straining is proportional, that is,

$$Q = \sigma(E_P)\frac{(\dot{\varepsilon}_P)^{\mu-1}}{(\dot{E}_P)^{\mu-1}} \qquad \text{and} \qquad \tau_i = \sigma(E_P)\frac{\ell^2(\ell\dot{\varepsilon}_P^*)^{\mu-2}\dot{\varepsilon}_{P,i}}{(\dot{E}_P)^{\mu-1}}. \qquad (3\text{-}4)$$

This choice guarantees a positive plastic dissipation rate because, as shown by direct calculation,

$$Q\dot{\varepsilon}_P + \tau_{,i}\dot{\varepsilon}_{P,i} = \sigma(E_P)\dot{E}_P. \qquad (3\text{-}5)$$

The choice (3-4) is a special case of the wide range of constitutive possibilities outlined in [Gudmundson 2004; Gurtin and Anand 2009] and is identical to that of [Fleck and Willis 2009a] for $\mu = 2$ and that of [Idiart et al. 2009] for arbitrary μ. As [Fleck and Willis 2009a] emphasized, (3-4) specifies the higher-order stresses, Q and τ_i, in terms of the rate of the plastic strain and its gradient. In this respect it is akin to conventional rigid plasticity. One important consequence of (3-4) is that Q and τ_i are not determined by the prior stress history; rather, they depend on the strain rates and thus on the boundary conditions of the incremental problem.

The fact that Q and τ_i are not predetermined in the current state is seen more clearly when the higher-order equilibrium equation in (2-9) is expressed in terms of the plastic strain rate using (3-4):

$$\left(\sigma(E_P)\frac{\ell(\ell\dot{\varepsilon}_P^*)^{\mu-2}\ell\dot{\varepsilon}_{P,i}}{(\dot{E}_P)^{\mu-1}}\right)_{,i} - \sigma(E_P)\frac{(\dot{\varepsilon}_P)^{\mu-1}}{(\dot{E}_P)^{\mu-1}} = -\sigma_e. \qquad (3\text{-}6)$$

The distributions of conventional effective stress, σ_e, and E_P *are* predetermined by the prior history. Equation (3-6) is a nonlinear, second-order partial differential equation homogeneous of degree zero in $\dot{\varepsilon}_P$. The solution $\dot{\varepsilon}_P$ depends on the boundary conditions with prescribed $t = \tau_i n_i = t^0$ on S_T and $\dot{\varepsilon}_P$ on S_u. Neither these boundary conditions, nor even S_T and S_u, need depend on prior history. Consequently, Q and τ_i from (3-4) can undergo discontinuous changes when incremental boundary conditions are

abruptly altered. However, when the boundary conditions for a sequence of incremental problems change continuously, as is the case for many problems of interest, Q and τ_i may also vary continuously.

It is readily verified that (3-6) is the Euler equation associated with stationary

$$H(\dot{\varepsilon}^P) = \int_V \left(\sigma(E_P)\dot{E}_P - \sigma_e \dot{\varepsilon}_P \right) dV - \int_{S_T} t^0 \dot{\varepsilon}_P \, dS \qquad (3\text{-}7)$$

with respect to all positive $\dot{\varepsilon}_P$ satisfying prescribed conditions on S_u. This is Minimum Principle I of [Fleck and Willis 2009a]. In each zone in which $\dot{\varepsilon}_P$ is nonzero, $\dot{\varepsilon}_P$ is fully determined by (3-6) if nonzero $\dot{\varepsilon}_P$ is prescribed on any portion of S. Otherwise, $\dot{\varepsilon}_P$ is determined to within a multiplicative constant within each nonoverlapping yielded zone.

From (3-5) it follows that

$$\dot{Q}\dot{\varepsilon}_P + \dot{\tau}_{,i}\dot{\varepsilon}_{P,i} + Q\ddot{\varepsilon}_P + \tau_{,i}\ddot{\varepsilon}_{P,i} = h(E_P)\dot{E}_P^2 + \sigma(E_P)\ddot{E}_P, \qquad (3\text{-}8)$$

where $h(E_P) = d\sigma(E_P)/dE_P$. This equation holds even when Q and τ_i change discontinuously, as described above. Based on (3-4) and on the derivatives of the plastic strain-rate quantities, one can show: $Q\ddot{\varepsilon}_P + \tau_{,i}\ddot{\varepsilon}_{P,i} = \sigma(E_P)\ddot{E}_P$. Consequently, whether Q and τ_i change continuously or discontinuously,

$$\dot{Q}\dot{\varepsilon}_P + \dot{\tau}_{,i}\dot{\varepsilon}_{P,i} = h(E_P)\dot{E}_P^2. \qquad (3\text{-}9)$$

The quadratic work terms for the incremental problem are therefore

$$\tfrac{1}{2}\{\dot{\sigma}_{ij}\dot{\varepsilon}_{ij}^e + \dot{Q}\dot{\varepsilon}_P + \dot{\tau}_{,i}\dot{\varepsilon}_{P,i}\} = \tfrac{1}{2}\{\mathcal{L}_{ijkl}(\dot{\varepsilon}_{ij} - \dot{\varepsilon}_P m_{ij})(\dot{\varepsilon}_{kl} - \dot{\varepsilon}_P m_{kl}) + h(E_P)\dot{E}_P^2\}.$$

The quadratic work terms are combined with the prescribed incremental surface tractions to form the functional whose minimum with respect to admissible distributions $(\dot{u}_i, \dot{\varepsilon}_P)$ provides the solution:

$$J(\dot{u}_i, \dot{\varepsilon}_P) = \frac{1}{2}\int_V \{\mathcal{L}_{ijkl}(\dot{\varepsilon}_{ij} - \dot{\varepsilon}_P m_{ij})(\dot{\varepsilon}_{kl} - \dot{\varepsilon}_P m_{kl}) + h(E_P)\dot{E}_P^2\}dV - \int_{S_T}(\dot{T}_i\dot{u}_i + \dot{t}\dot{\varepsilon}_P)dS, \qquad (3\text{-}10)$$

with $\dot{T}_i$ and $\dot{t}$ prescribed on S_T and with $\dot{u}_i$ and $\dot{\varepsilon}_P$ prescribed on S_u. If the complete distributions of $(\dot{u}_i, \dot{\varepsilon}_P)$ were unknown, as in the original version of [Fleck and Hutchinson 2001], then the minimum of J with respect to all admissible $(\dot{u}_i, \dot{\varepsilon}_P)$ produces the entire incremental solution. However, due to the special nature of the constitutive law (3-4) discussed above, $\dot{\varepsilon}_P$ is either known, or known to within a multiplicative constant, within each yielded region of the body. With $\dot{\varepsilon}_P$ constrained by the Minimum Principle I, it was proved in [Fleck and Willis 2009a] that the minimum of (3-10) delivers the unknown multiplicative constants of $\dot{\varepsilon}_P$ as well as $\dot{u}_i$ and $\dot{\sigma}_{ij}$. This is Minimum Principle II. The same reference also proves uniqueness and discusses bounding principles for the incremental problem. The incremental problem is fully characterized, apart from a condition for distinguishing loading and elastic unloading, as will be discussed shortly. The conventional stress is updated using $\sigma_{ij} \rightarrow \sigma_{ij} + \dot{\sigma}_{ij}$ in the standard manner, as is E_P based on (2-7); these and the distribution of m_{ij} are known entering the repetition of the solution process for the next incremental step.

Equation (3-4) implies the existence of a surface in terms of Q and τ_i given explicitly by

$$\Sigma \equiv \left(Q^{\mu/(\mu-1)} + (\tau/\ell)^{\mu/(\mu-1)} \right)^{(\mu-1)/\mu} = \sigma(E_P) \qquad \text{with} \qquad \tau = \sqrt{\tau_i \tau_i}. \qquad (3\text{-}11)$$
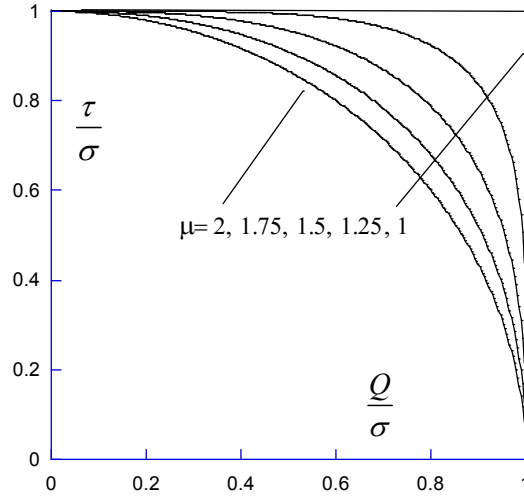
Figure 1. Yield surfaces for one-parameter family of scalar theories.

For $\mu = 2$, this same surface, or ones analogous to it, are found in the theories of [Fleck and Hutchinson 1997; Gudmundson 2004; Fleck and Willis 2009a]. The surface is plotted in Figure 1 for five values of μ, including the limit for $\mu = 1$ which has a corner. The plastic strain rate and its gradient are normal to the surface:

$$\dot{\varepsilon}_P = \dot{E}_P \frac{\partial \Sigma}{\partial Q} = \dot{E}_P \left(\frac{Q}{\sigma} \right)^{1/(\mu-1)}, \qquad \dot{\varepsilon}_{P,i} = \dot{E}_P \frac{\partial \Sigma}{\partial \tau_i} = \ell^{-1} \dot{E}_P \left(\frac{\tau}{\ell \sigma} \right)^{1/(\mu-1)} \frac{\tau_i}{\tau}. \tag{3-12}$$

These equations are equivalent to (3-4). Plastic loading requires $\dot{\varepsilon}_P \geq 0$. From (3-11) it follows that for any plastic loading increment $\dot{\Sigma} = h(E_P)\dot{E}_P \geq 0$ because $\dot{E}_P$ is intrinsically nonnegative and material softening is not considered, that is, $h(E_P) \geq 0$. In the present theory, the surface (3-11) is not a yield surface in the conventional sense because any field $\dot{\varepsilon}_P > 0$ satisfying (3-4) always produces Q and τ_i that lie on the surface.

   While this formulation guarantees positive plastic work, no attempt has been made here to identify whether the contributions of the gradients of plastic strains to the plastic work are energetic or dissipative in the terminology of [Gurtin and Anand 2009]. In this respect, the situation is analogous to conventional J_2 flow theory where the fractional proportions of the plastic work dissipated as heat and stored as the elastic energy of the dislocations are largely irrelevant to the formulation and, indeed, can differ from material to material without being reflected in the theory. It was noted earlier that the flow theories laid out here with isotropic hardening are not usually expected to be good models for histories involving reversed plastic straining due to disregard of Bauschinger effects. Efforts to associate energetic contributions with a component of kinematic hardening and dissipative contributions with the isotropic component of hardening, as some have attempted, fall outside the aims of this paper.

4. Tensor theory

The deformation and flow theory versions employing the measures of strain and strain rate based on tensor theory are analogous to those laid out above for the scalar version and can be found in [Fleck and Willis 2009b]. One major difference between the two versions concerns the manner in which plastic

flow is constrained in the scalar version contrasted with the tensor version. In the scalar version the plastic strain increment is constrained to be proportional to the deviator of the Cauchy stress, s_{ij}, which is known (and fixed) in the current state. The plastic strain increment is not directly coupled to the Cauchy stress in the tensor version. In the tensor version, the direction of the plastic strain increment, $\dot{\varepsilon}_{ij}^P$, is not constrained by any predetermined quantity in the current state. Instead, Q_{ij} is codirectional to $\dot{\varepsilon}_{ij}^P$ and τ_{ijk} is codirectional to $\dot{\varepsilon}_{ij,k}^P$, and both of the higher-order stresses are not known in the current state but are determined only by the solution to the incremental problem. Implications of this difference will emerge in the numerical solutions to the constrained film problem.

The *deformation theory* for the tensor version is still defined by the functional Φ in (3-1) with $\varepsilon_P = \sqrt{2\varepsilon_{ij}^P\varepsilon_{ij}^P/3}$ and E_P defined in (1-1). Now, however, ε_P^* is defined by (2-5). For proportional plastic straining the *flow theory* again coincides with the deformation theory.

5. Pure power-law solutions

A pure power-law material has uniaxial behavior specified by $\sigma(\varepsilon_P) = \sigma_0 \varepsilon_P^N$ with plastic work, $U_P(\varepsilon_P) = \sigma_0 \varepsilon_P^{N+1}/(N+1)$. Pure power-law solutions to the deformation theory neglect elasticity such that (3-1) becomes

$$\Phi(\boldsymbol{u}, \varepsilon_P) = \int_V U_P(E_P)\,dV - \int_{S_T} (T_i u_i + t\varepsilon_P)\,dS. \tag{5-1}$$

Any solution, ε_P, minimizes Φ subject to $\varepsilon_{ij}^P = (u_{i,j} + u_{j,i})/2$ (with $u_{j,j} = 0$) and satisfaction of conditions on S_u.

Consider *traction-prescribed* boundary value problems for the scalar theory with load parameter λ such that $(T_i, t) = \lambda(T_i^0, t^0)$ on S_T with (T_i^0, t^0) as fixed spatial distributions and $(u_i, \varepsilon_P) = 0$ on S_u. It is straightforward to show that solutions to the scalar theory have the form

$$(u_{,i}, \varepsilon_P) = \lambda^{1/N}(u_{,i}^0, \varepsilon_P^0), \qquad (s_{ij}, Q, \tau_i) = \lambda(s_{ij}^0, Q^0, \tau_i^0), \tag{5-2}$$

where the quantities with superscript "0" are functions of position but independent of λ. Similarly, for *displacement-prescribed* boundary value problems with $(u_i, \varepsilon_P) = \lambda(u_i^0, \varepsilon_P^0)$ on S_u and $(T_i, t) = 0$ on S_T, the form of the solution is

$$(u_{,i}, \varepsilon_P) = \lambda(u_{,i}^0, \varepsilon_P^0), \qquad (s_{ij}, Q, \tau_i) = \lambda^N(s_{ij}^0, Q^0, \tau_i^0). \tag{5-3}$$

For both types of boundary conditions, proportional straining occurs at every point in the body if λ is increased monotonically. Consequently, any deformation theory solution is a solution to the corresponding flow theory. This is the extension of Ilyushin's theorem to strain gradient plasticity. It holds for the tensor version as well. It is worth noting that any such solution is also a solution to the flow theory proposed in [Fleck and Hutchinson 2001]: the possibility of negative plastic dissipation is not an issue for such solutions.

6. Plane strain compression or separation of thin films: pure-power-law solutions

A slab of thin film is depicted in the insert in Figure 2 with height h and width L. The slab is bonded to rigid platens on the top and bottom; the lower platen is fixed while the vertical displacement of the upper slab is denoted by Δ. The sides ($x_1 = \pm L/2$) are traction-free. It is well known that the bonding
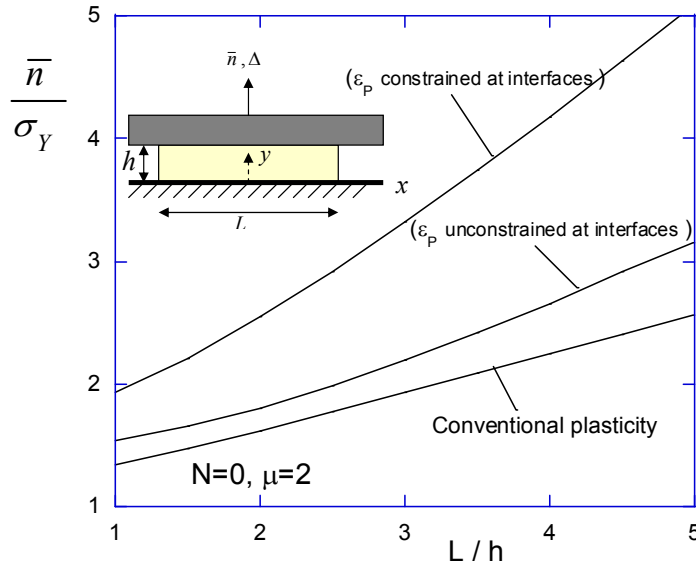
Figure 2. Effect of film aspect ratio on the average normal stress required to yield a rigid-perfectly plastic material ($N = 0$). The curves for the strain gradient material are based on the scalar model with $\mu = 2$ and dimensionless length parameter $\ell/h = \frac{1}{4}$. The upper curve assumes plastic flow is blocked at the platen/film interfaces ($\varepsilon_P = 0$) while the lower two curves are determined with no such restriction on plastic flow.

constraint of the platens gives rise to high stress triaxiality in the central region of the film accompanied by horizontal flow to accommodate the relative platen motion. The velocity of material elements paralleling the interfaces is approximately parabolic producing a linear variation of shear strain across the layer. Thus, as in film bending or wire torsion, compression or separation of the film intrinsically produces gradients of strain. Conventional theory for a rigid-perfectly plastic solid ($N = 0$) predicts that the magnitude of the average normal stress, $\bar{n}$, to yield a constrained layer in tension or compression increases with aspect ratio of the film according to [Hill 1950]

$$\frac{\bar{n}}{\sigma_Y} \cong 1 + \frac{1}{2\sqrt{3}}\frac{L}{h}, \tag{6-1}$$

where $\bar{n} = P/L$, with P as the total force/depth. The purpose of this section is to illuminate how plasticity size effects related to the film thickness further elevate the average normal stress.

Compression or separation of a slab of metal film that is well bonded to stiff platens may be a useful configuration for the experimental determination of material length scales and/or for assessing higher-order boundary conditions, especially if companion test data are available for the slab in shear. Moreover, the effective yield strength of the metal layer under normal separation exhibits a thickness effect that has direct relevance to metallic bonding. Plane strain compression or separation of a rectangular slab of film also illustrates the simplicity and utility of power-law solutions. Full elastic-plastic solutions based on the flow theory will be presented in Section 7 to bring out the interaction between elasticity and plasticity.

Two boundary conditions will be considered on the surfaces where the slab is bonded to the platens: (i) *unconstrained plastic flow* with no constraint on ε_P (or on ε_{12}^P in the tensor version), and (ii) *constrained plastic flow* with $\varepsilon_P = 0$ (or $\varepsilon_{12}^P = 0$ in the tensor version). The second boundary condition models dislocations blocked at a well-bonded interface joining a metal to a stiff elastic solid. In a continuum description, the plastic strain must vanish at the bonded surface if slip planes parallel to the surface are not available. Conditions intermediate to these limiting conditions can be defined (for example, [Fleck and Willis 2009a]) to model internal surfaces between different materials or other types of boundaries, but these will not be pursued here.

The uniaxial tensile response of the metal is taken as $\sigma(\varepsilon_P) = \sigma_0 \varepsilon_P^N$ or, equivalently, as $\sigma(\varepsilon_P) = \sigma_Y(\varepsilon_P/\varepsilon_Y)^N$ with $\sigma_0 = \sigma_Y/\varepsilon_Y^N$. The plastic work is $U_P(\varepsilon_P) = \sigma_Y \varepsilon_Y (\varepsilon_P/\varepsilon_Y)^{N+1}/(N+1)$. Let $x = x_1/h$ and $y = x_2/h$, and consider the displacement field

$$
\begin{aligned}
u_1/h &= -c_1 x\big(y(1-y)\big) - c_2 x\big(y(1-y)\big)^2 - c_3 x\big(y(1-y)\big)^3, \\
u_2/h &= c_1\big(\tfrac{1}{2}y^2 - \tfrac{1}{3}y^3\big) + c_2\big(\tfrac{1}{3}y^3 - \tfrac{1}{2}y^4 + \tfrac{1}{5}y^5\big) + c_3\big(\tfrac{1}{4}y^4 - \tfrac{3}{5}y^5 + \tfrac{1}{2}y^6 - \tfrac{1}{7}y^7\big),
\end{aligned}
\tag{6-2}
$$

with

$$
\Delta/h = u_2(h)/h = \tfrac{1}{6}c_1 + \tfrac{1}{30}c_2 + \tfrac{1}{140}c_3.
\tag{6-3}
$$

With $\{c_i\}$ as free coefficients, the field generates an admissible plastic strain field, ε_{ij}^P, that satisfies $u_{i,i} = 0$ and the boundary conditions on the top and bottom surfaces, subject to the additional condition that $c_1 = 0$ for the constrained flow condition (ii). The effective plastic strain is $\varepsilon_P = \sqrt{\tfrac{4}{3}(\varepsilon_{11}^{P2} + \varepsilon_{12}^{P2})}$ and the nonzero strain gradients from (6-2) are $\varepsilon_{12,2}^P$ and $\varepsilon_{11,2}^P = 2\varepsilon_{12,1}^P$. For the gradient measure for the scalar version, (2-2),

$$
\varepsilon_P^{*2} = \frac{4}{9\varepsilon_P^2}\big(\varepsilon_{12}^{P2}\varepsilon_{11,2}^{P2} + 4(\varepsilon_{11}^P \varepsilon_{11,2}^P + \varepsilon_{12}^P \varepsilon_{12,2}^P)^2\big),
\tag{6-4}
$$

while the measure for the tensor version, (2-3), is

$$
\varepsilon_P^{*2} = \tfrac{4}{3}(\varepsilon_{11,2}^{P2} + \varepsilon_{12,2}^{P2}).
\tag{6-5}
$$

The solution process is as follows. Using (6-4) or (6-5) together with the generalized effective strain, E_P, in (1-1), one obtains an expression for the energy functional, Φ, in (5-1) (the contribution from S_T vanishes). It depends on the free parameters, $\{c_i\}$, and on the parameters prescribing the problem according to

$$
\Phi = \sigma_Y \varepsilon_Y \frac{hL}{N+1}\Big(\frac{1}{\varepsilon_Y}\frac{\Delta}{h}\Big)^{N+1} f\Big(c_i, \frac{L}{h}, N, \frac{\ell}{h}, \mu\Big).
\tag{6-6}
$$

Numerical integration is employed to evaluate the integral in (5-1) for any $\{c_i\}$. The minimum of Φ with respect to the $\{c_i\}$ is also determined numerically, subject to (6-3), with $c_1 = 0$ for condition (ii). The results presented below have been computed using (6-2) with two free coefficients for each of the two cases, providing f with an accuracy of no less than about one percent. As an illustration, for condition (i), let the reference be results obtained using all three coefficients. Then, f is accurate to within 5% if only c_1 is used, while it is accurate, typically, to within a small fraction of a percent if c_1 and c_2 are used.

From $Pd\Delta = d\Phi$ and (6-6), the average normal stress, $\bar{n} = P/L$, for separating the platens is the power law

$$
\frac{\bar{n}}{\sigma_Y} = \Big(\frac{1}{\varepsilon_Y}\frac{\Delta}{h}\Big)^N f,
\tag{6-7}
$$

where f is the value obtained in the minimization of Φ. This expression also holds for compression with $\bar{n}$ as the compressive stress and Δ as the compression.

The results presented in Figure 2 display the most basic aspects of size-dependence related to films having a finite aspect ratio. In the limit of (6-7) for a rigid-perfectly plastic material ($N = 0$), the average normal stress, $\bar{n}/\sigma_Y = f$, required to separate (or compress) the platens is independent of Δ. In Figure 2, the dependence of $\bar{n}/\sigma_Y = f$ on L/h shown for conventional plasticity was computed using the scheme described above with $\ell = 0$ and no constraints on ε_P; it is in close agreement with the approximate formula (6-1). Included in the figure are results computed for the scalar measure (6-4) of effective plastic strain with $\mu = 2$ and $\ell/h = \frac{1}{4}$ for both unconstrained and constrained plastic flow at the film/platen interfaces. Film size effects become significant when the film thickness is reduced to several times the material length parameter. Moreover, plastic flow constraint at the film/platen interfaces further enhances the average flow strength. The trends in Figure 2 imply a significant elevation of the average flow strength stress due to strain gradient effects that superimpose on those associated with the film aspect ratio. Several aspects of these trends will be further explored in this section and the next.

The power-law relation between $\bar{n}$ and Δ from (6-7) is shown in Figure 3 for $N = 0.2$ and $L/h = 1$ for the scalar version. Due to the homogeneous composition (1-2) of the effective plastic strain, the relative elevation in strength due to the size effect for pure power materials is independent of strain, as evident from (6-7). For these examples, the effect of plastic flow constraint at the platens is roughly equivalent to a doubling of ℓ/h relative to the case of unconstrained plastic flow. For the unconstrained film the flow strength, $\bar{n}$, is elevated by a factor of two when the film thickness is reduced to about 2ℓ.

Figure 4 compares the responses predicted for the scalar and tensor versions for two dimensionless length parameters for $L/h = 3$. The normalizations of the plastic strain gradient measure, ε_P^*, for the
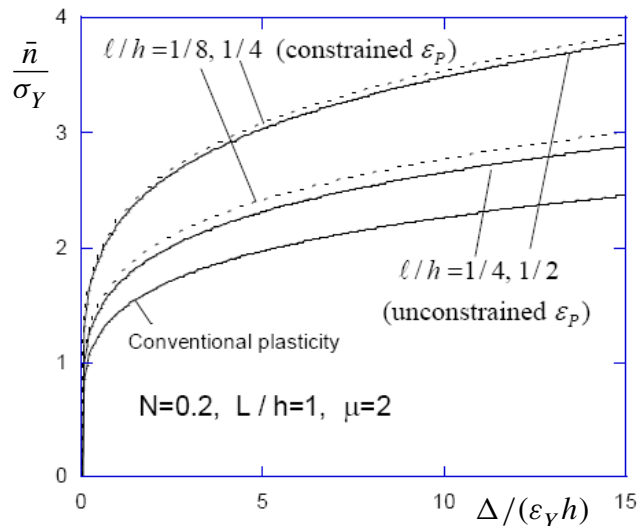
Figure 3. Pure-power relation between average normal stress and platen separation for a film slab with $L/h = 1$ based on the scalar measure of effective plastic strain. The relation applies in tension or compression with appropriate choice of signs. The solid curves are for films with no constraint on plastic strain at the film/platen interfaces while the plastic strain at the interfaces is constrained to be zero for the dashed curves.
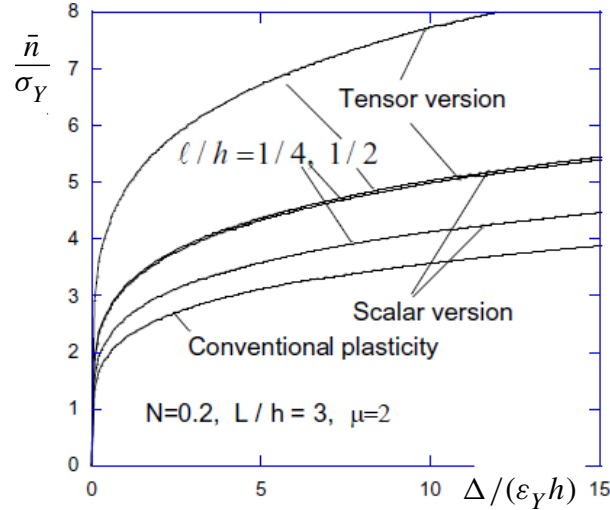
Figure 4. Comparison between pure-power law solutions based on scalar version and tensor version. Plastic flow at the film/platen interfaces is not constrained.

scalar and tensor versions defined in (2-2) and (2-3) coincide for shearing with a gradient (for example, with $\varepsilon_{12,2}^P = \varepsilon_{21,2}^P$ as the only nonzero gradient) but they do not coincide for the present problem. For this problem, the tensor measure of ε_P^* is approximately twice the scalar measure as reflected by the greater strengthening enhancement of the tensor version.[1]

The material length parameters must be calibrated against experimental data. As the discussion above illustrates, the numerical value of a length so obtained will depend on the choice of plasticity theory, possibly with extreme differences [Evans and Hutchinson 2009]. The material length parameter will also depend on the choice of μ in (1-1). As noted earlier, μ should not be regarded as a material parameter. Most theoretical formulations have $\mu = 2$, explicitly or implicitly, while the case for a linear dependence on strain gradients ($\mu = 1$) has been argued in [Nix and Gao 1998; Evans and Hutchinson 2009]. Strength elevation as dependent on μ for fixed ℓ/h is illustrated in Figure 5 for the scalar version. The implication of this figure, consistent with related studies [Fleck and Hutchinson 1997; Idiart et al. 2009; Idiart and Fleck 2010], is that a length parameter calibrated against experimental data for a model having $\mu = 1$ will be somewhat smaller than that for a model with $\mu = 2$.

7. Plane strain compression or separation of a film: numerical flow theory solutions

7A. *Numerical method.* Numerical solutions based on the one-parameter flow theory are carried out incrementally using a finite element procedure based on the minimum principles devised in [Fleck and Willis 2009a; 2009b]. Details will be restricted to the scalar version. The uniaxial stress-strain relationship is taken as the Ramberg–Osgood curve

$$\varepsilon = \frac{\sigma}{E} + \left(\frac{\sigma}{\sigma_y}\right)^{1/N}. \tag{7-1}$$

---

[1]The tensor contribution, $\ell\varepsilon_P^*$, to the effective plastic strain defined in (2-3) is the same as that of the 3-parameter scalar contribution of [Fleck and Hutchinson 2001; Fleck and Willis 2009a] with $\ell_1 = \sqrt{2/3}\ell$, $\ell_2 = \sqrt{1/6}\ell$, and $\ell_3 = \ell/2$.
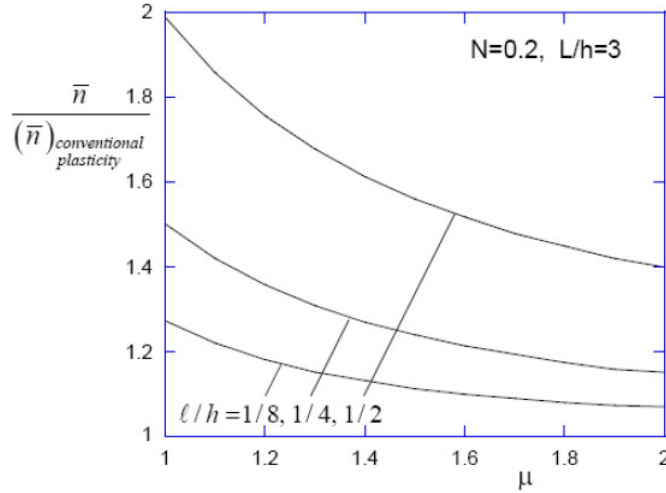
Figure 5. Elevation of average normal stress above that for conventional plasticity (at identical platen separations) due to strain gradient effects as dependent on the parameter μ characterizing the composition of the effective plastic strain measure. These results are derived for pure-power law material based on the scalar model. Plastic flow at the film/platen interfaces is not constrained.

For the present problem this will ensure that the entire block of material behaves plastically throughout the history of deformation, sidestepping the issue of a condition for initial plastic yielding. Moreover, no special consideration needs to be given to disjoint active plastic zones; $\dot{\varepsilon}_P$ can be expressed as a product of a multiplier, Λ, and a trial field, $\dot{\varepsilon}_P(x_i) = \Lambda \hat{\varepsilon}_P(x_i)$.

For each increment the solution is carried out in two steps:

1. Based on a plastic trial field, $\hat{\varepsilon}_P(x_i)$, and a specified load increment, the displacement rate, $\dot{u}_i(x_i)$ is obtained together with the plastic multiplier, Λ, by minimizing J in (3-10).

2. The plastic trial field, $\hat{\varepsilon}_P(x_i)$, is obtained by minimizing H in (3-7).

Step 1. Stationarity of J leads to

$$\int_V \{L_{ijkl}(\dot{\varepsilon}_{ij} - \Lambda \hat{\varepsilon}_P m_{ij})(\delta \dot{\varepsilon}_{kl} - \delta \Lambda \hat{\varepsilon}_P m_{kl}) + h(E_P)\hat{E}_P^2 \Lambda \delta \Lambda\} dV = \int_{S_T} (\dot{T}_i \delta \dot{u}_i + i \delta \Lambda \hat{\varepsilon}_P) dS,$$

which can be separated into the following two systems of equations:

$$\int_V L_{ijkl}(\dot{\varepsilon}_{ij} - \Lambda \hat{\varepsilon}_P m_{ij})\delta \dot{\varepsilon}_{kl} dV = \int_{S_T} \dot{T}_i \delta \dot{u}_i dS, \qquad (7\text{-}2)$$

$$\int_V \{-L_{ijkl}(\dot{\varepsilon}_{ij} - \Lambda \hat{\varepsilon}_P m_{ij})\hat{\varepsilon}_P m_{kl}\delta \Lambda + h(E_P)\hat{E}_P^2 \Lambda \delta \Lambda\} dV = \int_{S_T} i \delta \Lambda \hat{\varepsilon}_P dS. \qquad (7\text{-}3)$$

Using standard finite element interpolation with quadratic shape functions according to

$$\dot{u}_i = \sum_{N=1}^{8} N_i^N \dot{U}^N, \qquad \dot{\varepsilon}_{ij} = \sum_{N=1}^{8} B_{ij}^N \dot{U}^N,$$

or in matrix notation

$$\dot{\boldsymbol{u}} = \boldsymbol{N}\dot{\boldsymbol{U}}, \qquad \dot{\boldsymbol{\varepsilon}} = \boldsymbol{B}\dot{\boldsymbol{U}},$$

we obtain the discretized equations

$$\int_V \boldsymbol{B}^T \boldsymbol{L} \boldsymbol{B} \, dV \cdot \dot{\boldsymbol{U}} - \int_V \boldsymbol{B}^T \boldsymbol{L}(\hat{\varepsilon}_P \boldsymbol{m}) \, dV \cdot \Lambda = \int_{S_T} \dot{\boldsymbol{T}} \boldsymbol{N}^T \, dS, \qquad (7\text{-}4)$$

$$- \int_V (\hat{\varepsilon}_P \boldsymbol{m})^T \boldsymbol{L} \boldsymbol{B} \, dV \cdot \dot{\boldsymbol{U}} + \int_V \left((\hat{\varepsilon}_P \boldsymbol{m})^T \boldsymbol{L}(\hat{\varepsilon}_P \boldsymbol{m}) + h \cdot (\hat{E}^P)^2 \right) dV \cdot \Lambda = \int_{S_T} \dot{t} \hat{\varepsilon}_P \, dS. \qquad (7\text{-}5)$$

From these the incremental nodal displacement vector, $\dot{\boldsymbol{U}}$, and the plastic multiplier, Λ, are obtained. In these expressions, $\boldsymbol{L}$ and $\boldsymbol{m}$ are matrix forms of the tensors L_{ijkl} and m_{ij}, respectively.

Initially, a zero plastic trial field is assumed. At later load increments, the plastic trial field is solved for in Step 2.

Step 2. Stationarity of H results in

$$\int_V (Q\delta\dot{\varepsilon}_P + \tau_i \delta\dot{\varepsilon}_{P,i}) \, dV = \int_V \sigma_e \delta\dot{\varepsilon}_P \, dV + \int_{S_T} t^0 \delta\dot{\varepsilon}_P \, dS. \qquad (7\text{-}6)$$

An iterative procedure is used to solve for $\hat{\varepsilon}_P(x_i)$. Within a 2D planar setting $\hat{\varepsilon}_P$ is interpolated using bilinear shape functions. Hence, the trial field and its spatial derivatives can be expressed as

$$(\hat{\varepsilon}_P, \hat{\varepsilon}_{P,1}, \hat{\varepsilon}_{P,2})^T = \sum_{N=1}^{4} M_I^N \hat{\varepsilon}_P^N = \boldsymbol{M}\hat{\boldsymbol{\varepsilon}}_P,$$

with

$$\boldsymbol{M} = \begin{bmatrix} M_1 & M_2 & M_3 & M_4 \\ M_{1,1} & M_{2,1} & M_{3,1} & M_{4,1} \\ M_{1,2} & M_{2,2} & M_{3,2} & M_{4,2} \end{bmatrix} = \begin{bmatrix} \boldsymbol{M}_v \\ \boldsymbol{M}_{v,1} \\ \boldsymbol{M}_{v,2} \end{bmatrix}.$$

Introducing this in the variational statement above together with the constitutive equations, as well as requiring it to hold for all admissible variations in the plastic trial field, we obtain the discretized system of equations

$$\int_V \frac{\sigma(E_P)}{\hat{E}_P} \boldsymbol{M}^T \boldsymbol{A} \boldsymbol{M} \, dV \cdot \hat{\boldsymbol{\varepsilon}}_P = \int_V \sigma_e \boldsymbol{M}_v^T \, dV + \int_{S_T} t^0 \boldsymbol{M}_v^T \, dS. \qquad (7\text{-}7)$$

Here, for the scalar family of theories,

$$\boldsymbol{A} = \text{diag}\left(\left(\frac{\hat{\varepsilon}_P}{\hat{E}_P}\right)^{\mu-2}, \; l^2 \left(\frac{l\hat{\varepsilon}_P^*}{\hat{E}_P}\right)^{\mu-2}, \; l^2 \left(\frac{l\hat{\varepsilon}_P^*}{\hat{E}_P}\right)^{\mu-2} \right), \qquad (7\text{-}8)$$

which for the case of $\mu = 2$ reduces to $\boldsymbol{A} = \text{diag}(1, l^2, l^2)$.

This system of equations is solved iteratively for $\hat{\boldsymbol{\varepsilon}}_P$. An initial guess, $\hat{\boldsymbol{\varepsilon}}_P^0$, is taken as the solution from the former increment, except at the first increment when $\hat{\boldsymbol{\varepsilon}}_P^0$ is taken to be unity in all nodes, except on any boundary nodes where there is a constraint. The system of equations is solved iteratively ($k = 1, 2, 3, \ldots$) according to

$$\int_V \frac{\sigma(E_P)}{\hat{E}_P^{k-1}/\|\hat{E}_P^{k-1}\|_\infty} \boldsymbol{M}^T \boldsymbol{A} \boldsymbol{M} \, dV \cdot \hat{\boldsymbol{\varepsilon}}_P^k = \int_V \sigma_e \boldsymbol{M}_v^T \, dV + \int_{S_T} t^0 \boldsymbol{M}_v^T \, dS. \qquad (7\text{-}9)$$

Here, $\|\hat{E}_P^{k-1}\|_\infty$ denotes the maximum norm of the gradient enhanced effective plastic strain from the previous increment.

7B. *Flow theory predictions.* Normalized load-separation curves are presented in Figure 6 for short film slabs ($L/h = 1$) of an elastic-plastic material with $\varepsilon_Y = \sigma_Y/E = 0.01$, $\nu = 0.49$, and $N = 0.2$ for conventional J_2 flow theory and for the scalar version of the flow theory with two values of ℓ/h. The lower set of curves is for the case in which $\varepsilon_P = 0$ is enforced at the film/platen interfaces and the upper set is for the case where plastic flow is unconstrained at the interfaces. Included in each figure are the predictions based on the corresponding deformation theory for the pure power-law material with elasticity neglected.

The transition from predominantly elastic to nearly fully plastic behavior occurs for rather small normal displacements ($\Delta/\varepsilon_Y h \approx 4$) for the constrained case. For the constrained case considered in Figure 6, right, the response curves for the deformation theory and the incremental theory agree reasonably well at large deformation levels, for both the conventional predictions and the gradient dependent predictions with $\ell/h \approx \frac{1}{8}$ and $\ell/h \approx \frac{1}{4}$. In the initial range of Δ, the difference between the predictions of the two theories is due to the fact that the deformation theory solutions neglect elasticity.

For the unconstrained case in Figure 6, left, the predictions from the deformation theory and flow theory do not agree when gradient effects are important ($\ell/h > 0$), with significantly softer responses
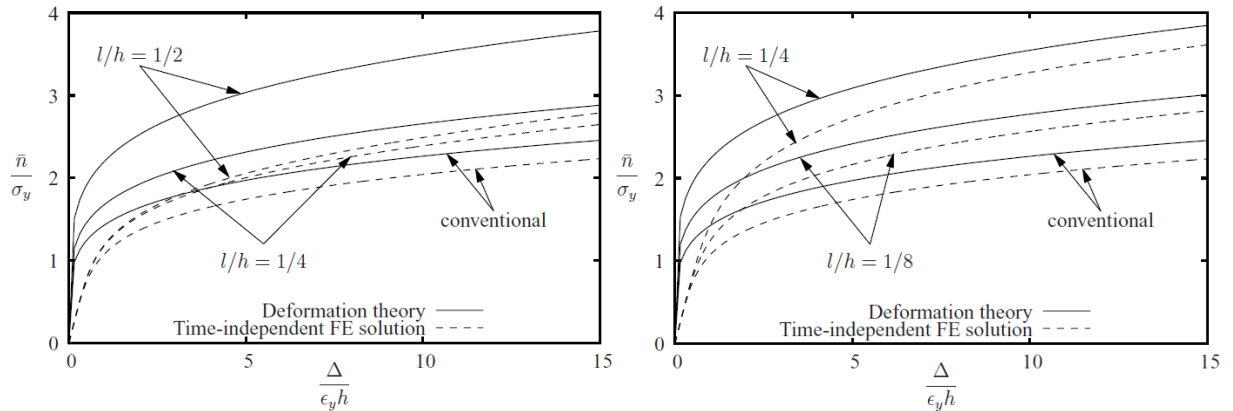
Figure 6. Normalized load-separation response computed using the scalar flow theory (left: unconstrained flow at the platens; right: constrained flow) for an elastic-plastic solid specified in the text (dashed curves) compared with predictions from the deformation theory for a pure power-law material with elasticity neglected (solid curves). All cases have $L/h = 1$, $N = 0.2$, and $\mu = 2$.

predicted for the flow theory. The discrepancy is associated with the fact that the plastic strain rates in certain regions of the slab are much larger according to the flow theory than the deformation theory. In fact, over large parts of the domain the plastic strain rates are so large in the flow theory that the von Mises effective stress decreases after initially increasing. As the slab is further deformed the effective stress approaches zero in two regions in the slab (see Figure 7). Hence, even though plastic deformation occurs, the von Mises stress decreases steadily as a result of a deviatoric stress relaxation. Eventually, this would lead to an indeterminate plastic flow direction, as the von Mises stress and the stress-deviator vanish. To circumvent this problem in the present calculations the direction of plastic flow as specified by m_{ij} is "frozen" when the von Mises stress drops below 10% of the initial yield stress. The domains where this method has been implemented are shown at the overall deformation level $\Delta/\varepsilon_Y h = 15$ with $\ell/h \approx \frac{1}{8}$, for the unconstrained (left figure) and constrained (right figure) case. The regions over which the effective stress is well below σ_Y are much larger. For both cases, these regions are connected to the center of the film/platen interface or located at the free surface. The regions where the von Mises stress has approached zero are appreciably larger for the unconstrained case, and this is even further exaggerated for increasing values of ℓ.

Also included in Figure 7 are contour lines of the plastic trial field evaluated at $\Delta/\varepsilon_Y h = 15$. It is observed for the constrained case in Figure 6, right, that the trial field goes toward zero at the constrained boundary, and that the contour lines are perpendicular to the all other parts of the exterior of the computational domain in accordance with the natural boundary condition. Correspondingly, in Figure 7, left, for the unconstrained case, the natural boundary condition is fulfilled at the entire exterior boundary of the computational domain. These trial fields provide insight as to why the deformation theory solutions
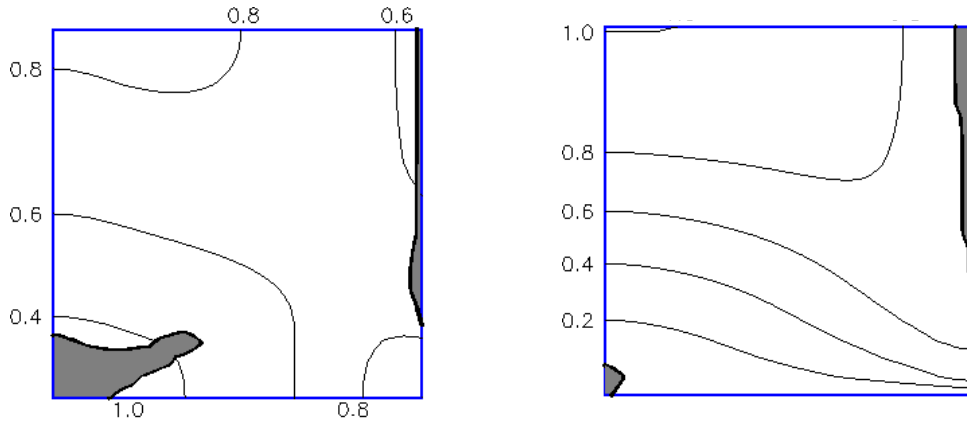
Figure 7. Regions (in gray) within which the effective stress, σ_e, has approached zero during continued plastic loading at $\Delta/\varepsilon_Y h = 15$ with $\ell/h = \frac{1}{8}$ (left: plastic strain is unconstrained at the film/platen interfaces; right: constrained case), for the scalar version with $L/h = 1$, $N = 0.2$, and $\mu = 2$. The rectangular region shown is one-quarter the entire cross-section of the film slab, extending upwards from the lower platen to the horizontal centerline and rightward from the vertical centerline to the right edge. The contour lines describe constant values of the trial function, $\hat{\varepsilon}_P$, whose maximum value has been normalized to be unity. Plastic loading occurs throughout the slab.
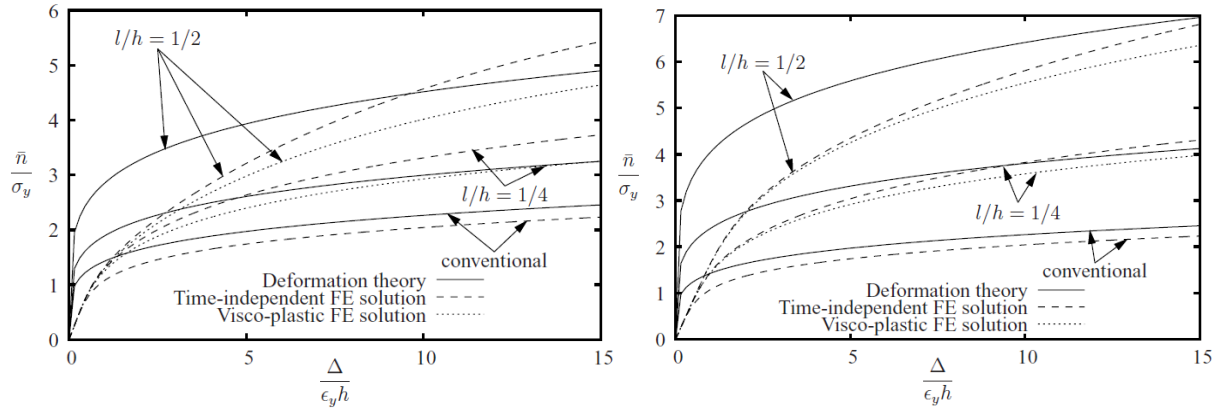
Figure 8. Normalized load-separation response computed using the tensor version of flow theory (left: unconstrained flow at the platens; right: constrained flow) for an elastic-plastic solid specified in the text (dashed and dotted curves) compared with predictions from the corresponding deformation theory for a pure power-law material with elasticity neglected (solid curves). The dashed curves are for the rate-independent formulation while the dotted curves are obtained using the viscous-plastic formulation. All cases have $L/h = 1$, $N = 0.2$, and $\mu = 2$.

and the flow theory solutions agree reasonably well for the constrained case but not for the unconstrained case — the constraint on plastic flow at the film/platen interfaces shrinks significantly the regions where the deviatoric stress is relaxed, whereas this is not the case for the unconstrained analysis.

Similar comparisons of pure-power deformation solutions and full elastic-plastic flow theory solutions for the tensor version are presented in Figure 8. Flow theory solutions are presented for both the rate-independent formulation and the viscoplastic, rate-dependent formulation [Gudmundson 2004; Gurtin and Anand 2005; Fleck and Willis 2009b], with either full constraint on plastic flow at the platens ($\varepsilon_{12}^P = \varepsilon_{11}^P = \varepsilon_{22}^P = \varepsilon_{33}^P = 0$) or no constraint on plastic flow ($t_{12} = t_{11} = t_{22} = t_{33} = 0$). The viscoplastic solutions are obtained using a method very similar to that used for the time-independent solutions. However, for the viscoplastic formulation, the Minimum Principle II in [Fleck and Willis 2009b] delivers the entire plastic strain rate field (no unknown multiplicative factor is needed), which means that the corresponding Minimum Principle I is used to solve for the displacement field alone. While the displacement interpolation is the same as that used for the scalar time-independent solutions, the plastic strain components are interpolated individually in the tensor version using bilinear shape functions, while exploiting plastic incompressibility to define one normal component in terms of the two others [Niordson and Legarth 2010]. For plane problems this results in 3 plastic strain degrees of freedom for each node as compared to just one for the scalar case. A standard viscoplastic power law has been employed using a viscoplastic exponent of 0.01 in combination with an overall loading rate on the order of the reference strain rate, thus minimizing viscous effects in the solutions obtained.

For the tensor versions, there is reasonable agreement between deformation and flow theory solutions in the fully plastic region for both the constrained and unconstrained cases. It is still true that the effective stress approaches zero as Δ increases over a nonnegligible region of the slab for both the rate-independent

and rate-dependent flow theories even though the material continues to deform plastically (plots are not shown). The components of the Cauchy deviator stress can even switch signs. However, unlike the scalar version, this behavior does not present a numerical problem because the plastic strain rate is not constrained to be codirectional with the deviator stress in the tensor version.

8. Concluding remarks

Constrained plastic compression or extension of thin metal films intrinsically involves strain gradients which, in turn, give rise to appreciable size-dependent strengthening when the film thickness is reduced below 3 to 5 times the material length parameter employed in the present formulation. Strengthening effects have clear implications for metal bonding layers, including increased resistance to plastic flow and higher stresses which are likely to affect ductility.

In this paper both deformation and flow theory versions of strain gradient plasticity have been employed. In rate-independent form, the flow theories used here are the new formulations proposed in [Fleck and Willis 2009a; 2009b]. For the thin slab problem considered here, difficulties arise with the scalar version of the flow theory wherein the strengthening effect due to the strain gradients appears to be significantly underestimated for the case in which plastic flow is not constrained at the loading platens. Although the source of the difficulties is not entirely clear, it seems likely that it stems in part from the fact that the theory predicts regions having a decreasing Cauchy deviator stress with the von Mises stress going to zero even as the plastic deformation increases. This presents numerical difficulties due to the fact that the plastic strain rate direction becomes indeterminate in the scalar formulation. That continuing plastic flow is predicted to occur at small, or even zero, von Mises stress may also be problematic from a physical standpoint. The von Mises stress also approaches zero in regions of the slab in the tensor version of the flow theory. This does not create numerical difficulties for this class of theories because the plastic strain rate is not tied to the Cauchy deviator stress. The apparently anomalous behavior in which the von Mises stress becomes small even in the presence of appreciable plastic straining has also been noted in simpler problems, such as in pure bending [Idiart et al. 2009].

The behavior noted above can arise because plastic straining is largely decoupled from the Cauchy stresses in these new flow theory formulations. Similar behavior occurs for solutions we have carried out based on the viscoplastic tensor version reported for the film slabs in Figure 8. Plastic straining is even less tightly coupled to the Cauchy stresses in viscoplastic formulations than in the rate-independent versions. These same features pertain to the general class of viscoplastic theories proposed in [Gudmundson 2004; Gurtin and Anand 2005]. Thus, for some problems this class of flow theories predicts that plastic straining can occur in regions of the body in which the elastic lattice strains are essentially zero. Such regions are predicted to be as large as one or more times the material length parameter, ℓ, that is, on the order of microns in size. It remains to be seen if such behavior is observed experimentally and whether such predictions are truly problematic from a physical standpoint.

Appendix: Formulations based on the strain measure proposed by Nix and Gao

Nix and Gao [1998] proposed an alternative way to combine the effective plastic strain, ε_P, and the magnitude of its gradient, ε_P^*, to model the effect of geometrically necessary dislocations. Evans and Hutchinson [2009] discussed differences in trends based the Nix–Gao composition from those based

on (1-1), highlighting the physical basis for each as well as their advantages and disadvantages. Here, a generalized effective strain measure will be defined based on the Nix–Gao composition. Then, that measure will be used to postulate strain gradient theories following the procedure outlined earlier using the deformation theory as the starting template. It should be noted that Nix and Gao [1998] did not propose a complete strain gradient plasticity theory — they suggested the manner in which the flow stress is enhanced by the gradient contribution. Later, Huang et al. [2000] proposed a higher-order theory based on the Nix–Gao composition; however, that proposal is quite different from what is given below. Indeed, the version suggested by Huang et al. introduces a second material length parameter tied to the higher-order stresses and unrelated to that of Nix and Gao used below. In most instances, applications of the Nix–Gao composition have been restricted to lower-order versions [Huang et al. 2004].

With σ_e as the effective flow stress, σ_0 as a reference yield stress, and $\sigma_e = \sigma_0 f(\varepsilon_P)$ as the relation between stress and plastic strain in uniaxial tension, Nix and Gao proposed that the flow stress is enhanced by the gradients according to

$$\sigma_e = \sigma_0 \sqrt{f(\varepsilon_P)^2 + \ell \varepsilon_P^*}. \tag{A.1}$$

This incorporates the gradient measure of the geometrically necessary dislocations in accord with the Taylor hardening law with a linear dependence on $\ell \varepsilon_P^*$ as the classical limit is approached, and it reduces to the uniaxial relation when the gradients can be ignored. For uniaxial behavior with monotonically increasing stress, that is, $df/d\varepsilon_P > 0$, let $f^{-1}(x)$ denote the inverse of $f(x)$, and define a generalized effective plastic strain by

$$E_P = f^{-1}\left(\sqrt{f(\varepsilon_P)^2 + \ell \varepsilon_P^*}\right). \tag{A.2}$$

With this definition, $\sigma_e = \sigma_0 f(E_P)$ is identical to (A.1) of Nix and Gao. Specifically, for a power law in uniaxial tension, $\sigma_e = \sigma_0 \varepsilon_P^N$, (A.2) is

$$E_P = (\varepsilon_P^{2N} + \ell \varepsilon_P^*)^{1/(2N)}. \tag{A.3}$$

The subsequent discussion will be restricted to (A.3).

With E_P in (A.3), the deformation theory based on (3-1) is fully defined. For the one-parameter gradient measure (2-2), one has

$$Q = \frac{\partial U_P}{\partial \varepsilon_P} = \sigma(E_P)\left(\frac{\varepsilon_P}{E_P}\right)^{2N-1}, \qquad \tau_i = \frac{\partial U_P}{\partial \varepsilon_{P,i}} = \sigma(E_P)\frac{\ell}{2N}\left(\frac{1}{E_P}\right)^{2N-1}\frac{\varepsilon_{P,i}}{\varepsilon_P^*}. \tag{A.4}$$

The definition of the flow theory is not as straightforward. Unlike the measure in (1-1), (A.3) is not homogeneous in ε_P and ε_P^*, except if $N = \frac{1}{2}$. Nevertheless, a generalized plastic strain rate can be defined that coincides with (A.3) for proportional straining. With $\dot{\varepsilon}_P = \sqrt{2\dot{\varepsilon}_{ij}^P \dot{\varepsilon}_{ij}^P/3}$, $\dot{\varepsilon}_P^*$ defined by (2-7), and total values defined by (2-10), it is easily seen that a measure with the desired property is

$$\dot{E}_P = \frac{1}{E_P^{2N-1}}\left(\varepsilon_P^{2N-1}\dot{\varepsilon}_P + \frac{\ell}{2N}\dot{\varepsilon}_P^*\right). \tag{A.5}$$

If $N = \frac{1}{2}$, (A.5) coincides with the present measure in (2-9) for $\mu = 1$ and it is otherwise similar in the sense that it is a linear composition of the rate contributions. The constitutive relations for the flow

theory are taken as

$$Q = \sigma(E_P)\left(\frac{\varepsilon_P}{E_P}\right)^{2N-1} \qquad \text{and} \qquad \tau_i = \frac{\sigma(E_P)\ell}{2N}\left(\frac{1}{E_P}\right)^{2N-1}\frac{\dot{\varepsilon}_{P,i}}{\dot{\varepsilon}_P^*} \tag{A.6}$$

such that the plastic dissipation rate is again $Q\dot{\varepsilon}_P + \tau_i\dot{\varepsilon}_{P,i} = \sigma(E_P)\dot{E}_P$. The field equation for $\dot{\varepsilon}_P$ is obtained from the higher-order equilibrium equation (2-12), and it is the Euler equation associated with Minimum Principle I defined, as before, by (3-7). One can also show that

$$\dot{Q}\dot{\varepsilon}_P + \dot{\tau}_i\dot{\varepsilon}_{P,i} = h(E_P)\dot{E}_P^2 + (1-2N)\frac{\sigma(E_P)}{E_P}\left(\dot{E}_P^2 - \left(\frac{\varepsilon_P}{E_P}\right)^{2N-1}\dot{\varepsilon}_P^2\right). \tag{A.7}$$

Thus, the functional associated with Minimum Principle II in (3-10) must be modified by replacing $h(E_P)\dot{E}_P^2$ by the right hand side of (A.7).

It is interesting to note that the deformation and flow theories proposed above based on the effective strain measure of [Nix and Gao 1998] coincide with the corresponding theories based on the Fleck–Hutchinson measure when $\mu = 1$ and $N = \frac{1}{2}$. The nice properties of solutions for pure power-law materials noted in Section 5 require a measure that is homogeneous in the plastic strain and its gradient. Therefore, the simple solution structure and its consequences in Section 5 do not hold for the theories based on the Nix–Gao measure unless $N = \frac{1}{2}$.

References

[Aifantis 1984] E. C. Aifantis, "On the microstructural origin of certain inelastic models", *J. Eng. Mater. Technol.* (*ASME*) **106**:4 (1984), 326–330.

[Evans and Hutchinson 2009] A. G. Evans and J. W. Hutchinson, "A critical assessment of theories of strain gradient plasticity", *Acta Mater.* **57**:5 (2009), 1675–1688.

[Fleck and Hutchinson 1997] N. A. Fleck and J. W. Hutchinson, "Strain gradient plasticity", *Adv. Appl. Mech.* **33** (1997), 295–361.

[Fleck and Hutchinson 2001] N. A. Fleck and J. W. Hutchinson, "A reformulation of strain gradient plasticity", *J. Mech. Phys. Solids* **49**:10 (2001), 2245–2271.

[Fleck and Willis 2009a] N. A. Fleck and J. R. Willis, "A mathematical basis for strain-gradient plasticity theory, I: Scalar plastic multiplier", *J. Mech. Phys. Solids* **57**:1 (2009), 161–177.

[Fleck and Willis 2009b] N. A. Fleck and J. R. Willis, "A mathematical basis for strain-gradient plasticity theory, II: Tensorial plastic multiplier", *J. Mech. Phys. Solids* **57**:7 (2009), 1045–1057.

[Fredriksson and Gudmundson 2005] P. Fredriksson and P. Gudmundson, "Size-dependent yield strength of thin films", *Int. J. Plast.* **21**:9 (2005), 1834–1854.

[Gudmundson 2004] P. Gudmundson, "A unified treatment of strain gradient plasticity", *J. Mech. Phys. Solids* **52**:6 (2004), 1379–1406.

[Gurtin and Anand 2005] M. E. Gurtin and L. Anand, "A theory of strain-gradient plasticity for isotropic, plastically irrotational materials, I: Small deformations", *J. Mech. Phys. Solids* **53**:7 (2005), 1624–1649.

[Gurtin and Anand 2009] M. E. Gurtin and L. Anand, "Thermodynamics applied to gradient theories involving the accumulated plastic strain: the theories of Aifantis and Fleck and Hutchinson and their generalization", *J. Mech. Phys. Solids* **57**:3 (2009), 405–421.

[Hill 1950] R. Hill, *The mathematical theory of plasticity*, Clarendon, Oxford, 1950.

[Huang et al. 2000] Y. Huang, H. Gao, W. D. Nix, and J. W. Hutchinson, "Mechanism-based strain gradient plasticity, II: Analysis", *J. Mech. Phys. Solids* **48**:1 (2000), 99–128.

[Huang et al. 2004] Y. Huang, S. Qu, K. C. Hwang, M. Li, and H. Gao, "A conventional theory of mechanism-based strain gradient plasticity", *Int. J. Plast.* **20**:4–5 (2004), 753–782.

[Idiart and Fleck 2010] M. I. Idiart and N. A. Fleck, "Size effects in the torsion of thin metal wires", *Model. Simul. Mater. Sci. Eng.* **18**:1 (2010), 015009.

[Idiart et al. 2009] M. I. Idiart, V. S. Deshpande, N. A. Fleck, and J. R. Willis, "Size effects in the bending of thin foils", *Int. J. Eng. Sci.* **47**:11-12 (2009), 1251–1264.

[Mühlhaus and Aifantis 1991] H.-B. Mühlhaus and E. C. Aifantis, "A variational principle for gradient plasticity", *Int. J. Solids Struct.* **28**:7 (1991), 845–857.

[Niordson and Legarth 2010] C. F. Niordson and B. N. Legarth, "Strain gradient effects on cyclic plasticity", *J. Mech. Phys. Solids* **58**:4 (2010), 542–557.

[Nix and Gao 1998] W. D. Nix and H. Gao, "Indentation size effects in crystalline materials: a law for strain gradient plasticity", *J. Mech. Phys. Solids* **46**:3 (1998), 411–425.

CHRISTIAN F. NIORDSON: cn@mek.dtu.dk
Department of Mechanical Engineering, Technical University of Denmark, Nils Koppels Allé, Building 403, 2800 Kgs. Lyngby, Denmark

JOHN W. HUTCHINSON: hutchinson@husm.harvard.edu
School of Engineering and Applied Sciences, Harvard University, 29 Oxford St., Cambridge, MA 02138, United States
http://www.seas.harvard.edu/hutchinson/

■msp

AN AXISYMMETRIC PARACHUTE MODEL WITH WRINKLING

YOAV OFIR, DAN GIVOLI AND AVINOAM LIBAI

We are honored to dedicate this paper to Charles Steele and to the memory of Marie-Louise Steele.
Charles was my admired teacher at Stanford, and has opened my eyes to the wonderful world of shells.
Marie-Louise was a friend who always had a few warm words for a PhD student undergoing stressful times.
I will always remember her very fondly.   *– Dan Givoli*

The deformation and stresses in a parachute canopy are analyzed, using an axisymmetric model. The canopy is modeled as an ideal or nearly ideal elastic membrane shell. Steady state is assumed. No restriction is imposed on the size of the deformation, and thus the problem is strongly nonlinear. Wrinkles which form over parts of the surface are taken into consideration approximately via the use of wrinkle fields (tension fields). The solution is obtained by a double-iteration numerical procedure, based on a shooting technique and incremental loading. First a basic parachute model is considered in which the material is assumed isotropic and Hookean, the pressure distribution is assumed uniform (but with unknown magnitude), and the canopy is assumed to have a small central hole. Later all these assumptions are relaxed. Numerical examples are presented for some representative cases.

1. Introduction

Analysis of deformation and stresses in a parachute under general conditions is extremely complicated. Treatment of the full problem in the literature is scarce; the impressive computational work of Tezduyar's group [Stein et al. 2000; 2001] and that of [Kim and Peskin 2006; 2009], which takes into account the complete fluid-structure interaction, should be mentioned in this context. The structural analysis of the canopy is considered, for example, in [Liu et al. 2001; Lu et al. 2001]. These two papers take into account the *wrinkles* which are generally formed in the parachute canopy. Information on such wrinkles is important since they may affect the deformation and stress distribution in the canopy and thus may have an influence on the overall parachute performance.

In this paper we propose a much simpler method of analysis for parachutes. Although the actual behavior of real parachutes and their design are beyond the scope of this paper, we briefly mention a few facts related to this subject, with particular attention to the issues of shape, stability, and wrinkles. The classical paper [Taylor 1919] recognized that "if stability were the chief thing to be desired, a parachute in the form of rounded cone is desirable... If lightness is the most important quality, it appears that one should make one's parachute as flat as possible, provided it will remain distended. If, however, the parachute is made too flat, radial crinkles or pleats will appear and these must mean a waste of material." Taylor then proceeds to determine the optimal deformed shape of a parachute, and obtains a theoretical shape determined by elliptic function curves. He comments that this theoretical shape differs from actual shapes of parachutes in that the curvature of the profile near the edge is smaller in the former, and he attributes this to the assumption of uniform pressure underlying his analysis.

The parachute studied in [Taylor 1919] assumes that inextensible wires woven into the fabric extend radially from the center to the edge. In between these wires the fabric bulges out, to form a lobed shape. In contrast, the parachute considered in the present paper is assumed to be *axisymmetric*, and therefore cannot account for canopy lobes. Admittedly, axisymmetry is rather a strong simplification in practical parachute analysis, since it neglects the presence of the lobes. Moreover, in some cases the axisymmetric solution may be an *unstable* equilibrium solution of the parachute problem, while a stable solution deviates from the axisymmetric configuration. Baginski et al. [2006] and Deng and Pellegrino [2008] demonstrate such symmetry-breaking clefting in the case of pumpkin balloons. In the case of smooth surfaces, wrinkles coalesce and form a small number of localized folds, as in Mylar party balloons. Both papers obtain stable solutions that deviate from the axisymmetric shape as the pressure is decreased from a nominal value by using detailed three-dimensional finite element analyses. The method of analysis proposed here is, of course, oblivious to all these effects, which is its main limitation. Despite this fact, we believe that the types of results that this simple method can produce are very helpful in preliminary design stages, and can serve as basis for more detailed asymmetric analysis.

In this context, the spectrum of analysis methods for parachutes should be recognized. On one end of this spectrum one can find the complete numerical analysis methods mentioned above, usually based on finite element discretization in space. On the other end of the spectrum are crude calculations or analytical solutions, usually associated with linear elasticity, namely with the assumption of small deformation and linear material behavior. Our proposed method is somewhere inside this interval; it does make a few basic assumptions, the most limiting of which being axisymmetry, but on the other hand it assumes neither small deformation nor linear material behavior.

The canopy may be thought of as an ideal or nearly ideal elastic membrane shell. Ideal membranes will wrinkle rather than support compressive stresses. Often this results in the formation of wrinkle zones over portions of the deformed surface. These zones, in which the principal stress resultants are everywhere nonnegative, provide a mechanism for carrying the imposed loads. Within a wrinkle zone, the crests and troughs of the wrinkles are parallel to the direction of principal tension, and the stress resultant in the direction normal to them (in the tangent plane) vanishes.

The parachute behaves essentially as a pulled axisymmetric membrane with surface of positive Gaussian curvature. To illustrate the basic phenomena involved, Figure 1 describes the behavior of a model of a membrane with surface of positive Gaussian curvature, albeit significantly simpler than that of the
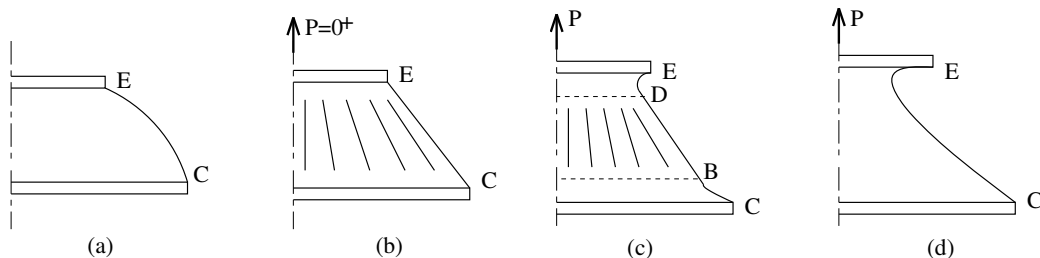
Figure 1. Wrinkles forming in an axisymmetric pulled membrane: (a) the unloaded membrane; (b) the wrinkled cone generated by a slight pulling force; (c) the wrinkle region (BD) and the two doubly tense regions (CB and ED) formed by a medium-sized pulling force; (d) the all-tense surface generated by a large pulling force.

parachute to be discussed later. The membrane is subjected to a pulling force, as was considered in [Libai and Givoli 2002]. The membrane is attached to two rigid plates along its two edges C and E, as shown in Figure 1a. The bottom plate is assumed to be fixed, whereas the top plate is pulled upwards. During the pull, the upper plate remains parallel to the lower plate and does not rotate with respect to it. The application of a slight pulling force $P = 0^+$ to the upper plate results in the formation of meridional wrinkles, and the membrane surface becomes a "wrinkled cone;" see Figure 1b. The formation of wrinkles is due to the fact that the circumferential stress resultants, which would have existed in such a membrane, must be negative, and are replaced by the wrinkle field. In addition, the requirements of equilibrium in the direction of the normal to the membrane demand that the curvature of the meridional lines be zero. Thus, the meridional wrinkle field takes on the shape of a cone.

In the vast majority of materials for membranes, a Poisson-like effect exists, such that a tensile stress in one direction is accompanied by contraction (negative strain) in the directions normal to it. If this contraction were prevented, say, along a boundary, then positive stresses would form along it and in its immediate neighborhood. The rigid plates at the upper and lower edges of the membrane prevent the transverse contraction, which would have occurred due to the pulling, and, thus, positive circumferential stresses accompany the positive pulling stresses. A biaxial state of stress is, thus, formed at and near the edges. The size of the edge zones is $o(e)$, where e is a typical meridional strain; see [Libai 1990]. The entire effect is strongly nonlinear.

An increase in the pulling force leads to a corresponding increase in e, so that the edge zones deteriorate. Thus, increasingly larger doubly tense regions are formed near the boundaries, where both the meridional and tangential (hoop) stress resultants are positive, while the central region remains wrinkled, as illustrated in Figure 1c. In the absence of internal pressure, the meridional curvature of the surface in the doubly tense regions is always negative; see, for example, [Libai and Simmonds 1998; Libai and Givoli 2002]. A further increase in the force causes the central wrinkle region to shrink, until, finally, a tensile *biaxial* state of stress exists over the entire surface; see Figure 1d.

In the case of large wrinkle zones in a membrane, direct and full treatment of the wrinkles is generally difficult. This fact gave rise to the notion of *wrinkle fields*, or *tension fields*, which represent the wrinkles in a "homogenized" sense. The basic idea in the theory of wrinkle fields is to avoid studying the wrinkle region in detail by replacing it with a smoothed-out *pseudosurface*. This pseudosurface must be in equilibrium, and the minimum principal stress resultant must vanish on it. Obviously, the stretch of the pseudosurface in the direction of the zero minimum principal stress resultant is nonphysical, in that it is not equal to that of the actual wrinkled surface.

In [Wu 1978] the first complete treatment of wrinkling in nonlinearly elastic membranes of revolution was given. Zak [1982] extended the theory to wrinkling of films of arbitrary shape. Pipkin [1986] proposed the notion of "relaxed energy density" in this context. Steigman and coworkers have made extensive theoretical studies [Steigmann and Pipkin 1989; Steigmann 1990; Li and Steigmann 1995a; 1995b] of the behavior of wrinkled and partly wrinkled membranes in various cases, including that of pressurized spherical and toroidal membranes. Libai [1990] presented a complete theoretical analysis of the transition zone between the doubly tense and wrinkle regions for pulled spherical membranes. Tait et al. [1996] and Tait and Connor [1997] solved wrinkling problems for cylindrical membranes. Roddeman et al. [1987a; 1987b], Jeong and Kwak [1992], Chiu et al. [1993], and Muttin [1996] developed finite element methods for the solution of partly wrinkled membranes. See [Libai and Simmonds 1998,

Chapters 5 and 7] for a detailed exposition on the subject and additional references. Recent publications making computational use of wrinkle fields for various applications include [Raible et al. 2005; Cavicchi et al. 2009; Mosler and Cirak 2009; Liu and Sze 2009].

In this paper we use wrinkle fields to analyze the deformation and stresses in a parachute canopy, using an axisymmetric model. We model the parachute as an ideal or nearly ideal elastic membrane shell. "Nearly ideal" means that the membrane is allowed to sustain small compressive stresses; it wrinkles only under larger compressive stresses. Thus, the minimal principal stress resultant is N_{cr}, where N_{cr} is a nonpositive parameter assumed to be a material property. For ideal membranes $N_{cr} = 0$, and for nearly ideal membranes N_{cr} has a small negative (given) value.

The idea to use $N_{cr} \neq 0$ in a membrane was suggested in [Rimrott and Cvercko 1986] in the context of pulled flat rectangular membranes. The authors claim that the formation of a *finite* number of wrinkles, which is typically observed in experiments in real membranes, is evidence of the presence of some small bending stiffness, since in the absence of such stiffness there should be an infinite number of wrinkles. Based on the experimental observation that, after the initial wrinkles form, additional deformation does not affect their number, Rimrott and Cvercko obtained a relation which connects the parameter N_{cr} to the number of wrinkles in the membrane, that is, to the wavelength of a single wrinkle. This relation can be exploited to effectively measure N_{cr} by counting the number of wrinkles.

Wong and Pellegrino [2006] derive another relation, based on a simple buckling model, in which N_{cr} depends on the material properties (Young's modulus and Poisson's ratio), the thickness of the membrane, and the half-wavelength of the wrinkle. They use this relation and the condition $N_\theta = N_{cr}$ (where N_θ is the transverse stress resultant) to calculate the wrinkle half-wavelength, namely the number of wrinkles.

In the present paper we calculate solutions for both ideal membranes ($N_{cr} = 0$) and nearly ideal membranes ($N_{cr} < 0$). In the latter case, our choice for the value of N_{cr} is rather arbitrary and does not necessarily correspond to an actual membrane material and thickness. We simply increase $|N_{cr}|$ from zero until the effect of this parameter becomes significant. Choosing N_{cr} in a more judicious way would require an additional analysis, such as that of Wong and Pellegrino.

We assume steady state, namely that the canopy is fully open and the air pressure acting on it is quasistatic. On the other hand, we pose no restriction on the size of the deformation, which may be very large, thus making the problem strongly nonlinear. We solve the problem using a double-iteration numerical procedure, based on a shooting technique and incremental loading. We demonstrate that the method works very well despite the complexity of the problem which involves the highly nonlinear membrane shell equations, different regions with interfaces which are unknown a priori, and deformation-dependent pressure magnitude.

We first consider a basic parachute model in which we assume the material to be isotropic and Hookean, the pressure distribution to be uniform, although with an unknown magnitude, and the canopy to have a small hole in its center. Later we relax all these assumptions and consider also an orthotropic canopy (which allows us to represent the higher stiffness in the meridional direction due to the canopy cords), neo-Hookean material, nonuniform pressure (with a known distribution but unknown magnitude), and canopies without a central hole. In the latter case, some numerical difficulties arise near the apex, and we overcome them by slightly modifying the calculation procedure in the apex vicinity.

Summary. In Section 2 we present the parachute model and its governing equations. In Section 3 we outline the computational scheme for the solution of the parachute problem, for the basic model. In

Section 4 we discuss the four extensions mentioned above. In Section 5 we present some representative numerical results. We end with some concluding remarks in Section 6.

2. Parachute model and governing equations

2.1. *Basic model and notation.* We consider a thin curved axisymmetric hyperelastic membrane with a positive Gaussian curvature (the canopy) subjected to a given axial pulling force P (the carried weight), as illustrated in Figure 2. The meridian of the membrane extends between points C and E. The force P is transferred to the canopy through tensed cords of given length L, which must be tangential to the meridian at point C (see Figure 2) since the membrane has no bending stiffness. The entire model is axisymmetric, and thus the cords are represented by a continuous "cone" revolving around the axis of symmetry. The tension of this "axisymmetric cord" in the deformed configuration is denoted T, and the angle it forms with the axial direction β, both unknown. The pressure acting on the membrane is assumed to be uniform, but its magnitude, denoted p, is unknown.

The upper edge of the shell, point E, is fixed to a rigid ring of given radius, which encloses and reinforces a small hole in the center of the canopy. For arbitrary external forces P and p, an axial reaction force R would act at this point. Since the parachute is known to be in equilibrium under P and p alone, we require that this reaction vanish, that is, $R = 0$, which will give us a condition that will be essential in determining the pressure p.

We expect the membrane to have doubly tense regions near the lower and upper edges, and possibly a wrinkle field in a central region of the meridian, bounded by points B and D (see Figure 2).

We define the notation needed to describe the geometry of the shell. The surface of the unloaded membrane is described by the function $r = r(z)$, where r and z are the radial and axial coordinates,
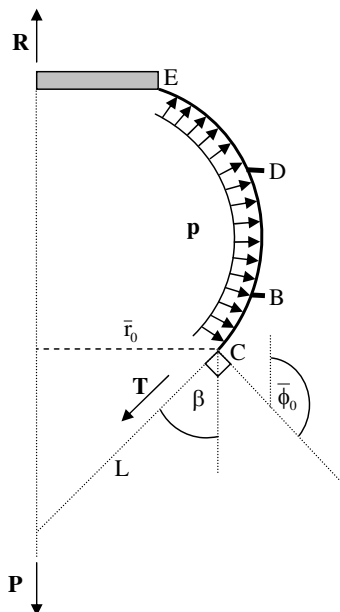
Figure 2. The basic parachute model.

respectively. The arc length coordinate along the meridian is denoted s. We use a superposed bar to denote quantities in the deformed configuration. Thus, we let $\bar{r}$, $\bar{z}$, and $\bar{s}$ denote, respectively, the radial, axial, and arc length locations of a material point in the deformed configuration. Thus, the surface of the deformed membrane is described by the function $\bar{r} = \bar{r}(\bar{z})$. We let $\bar{\phi}$ be the angle between the axial direction and the normal to the deformed surface. We denote quantities related to the lower edge of the meridian (point C) by the subscript 0. Thus, $\bar{\phi}_0$ is the angle $\bar{\phi}$ at the lower edge in the deformed configuration.

We let N_s and N_θ be the meridional and tangential stress resultants per unit undeformed length, and $\bar{N}_s$ and $\bar{N}_\theta$ be the meridional and tangential stress resultants per unit deformed length. We let λ_s and λ_θ be the principal stretches.

2.2. *Governing equations and boundary conditions.* In this section we summarize all the equations and boundary conditions that constitute our model. We first write down all the equations that hold in a doubly tense region, and then we comment on the changes that need to be made in these equations in a wrinkle region.

Basic relations connecting quantities in the deformed and undeformed configurations. Classical relations of axisymmetric membrane shell theory are used. The relations among the stress resultants are given by

$$r N_s = \bar{r} \bar{N}_s, \qquad ds\, N_\theta = d\bar{s}\, \bar{N}_\theta, \tag{1}$$

which actually *define* the stress resultants in the undeformed configuration. The principal stretches are related to the geometry through

$$\lambda_s = d\bar{s}/ds, \tag{2}$$

$$\lambda_\theta = \bar{r}/r. \tag{3}$$

Axial equilibrium. Consideration of the axial equilibrium of a portion of the membrane which lies on the lower side of a circumferential cross section yields

$$P - 2\pi \bar{r} \bar{N}_s \sin \bar{\phi} - I = 0, \tag{4}$$

where

$$I \equiv \int_0^{\bar{s}} 2\pi p \bar{r}' \cos \bar{\phi}' \, d\bar{s}'. \tag{5}$$

The integral I is the total axial force due to the pressure p acting from the lower edge of the membrane along the meridian up to the point under consideration. A prime ($'$) denotes a quantity calculated at the integral variable point $\bar{s}' \in [0, \bar{s}]$. Using (1) in (4) we also obtain

$$N_s = \frac{P - I}{2\pi r \sin \bar{\phi}}. \tag{6}$$

This equation holds in the entire membrane.

Normal equilibrium. In the doubly tense regions, we also consider the normal equilibrium of an element, which yields the equation

$$\frac{\bar{N}_\theta}{\bar{\rho}_\theta} + \frac{\bar{N}_s}{\bar{\rho}_s} = p. \tag{7}$$

Here, $\bar{\rho}_s$ and $\bar{\rho}_\theta$ are the meridional and tangential radii of curvature of the deformed membrane, which satisfy

$$\frac{1}{\bar{\rho}_s} = -\bar{\phi},_{\bar{s}}, \qquad \bar{\rho}_\theta = \frac{\bar{r}}{\sin\bar{\phi}}. \tag{8}$$

In (8) and elsewhere, a comma indicates differentiation. We use (1), (7), and (8) to obtain

$$N_\theta = \frac{r N_s \bar{\phi},_s + p\bar{r}\bar{s},_s}{\sin\bar{\phi}}. \tag{9}$$

Substituting N_s from (6) yields

$$N_\theta = \frac{p\bar{r}\bar{s},_s}{\sin\bar{\phi}} - \frac{(\cot\bar{\phi}),_s}{2\pi}(P - I). \tag{10}$$

Compatibility equation. The *compatibility relation* simply relates $\bar{r}$, $\bar{s}$, and $\bar{\phi}$ through

$$\frac{d\bar{r}}{d\bar{s}} = -\cos\bar{\phi}. \tag{11}$$

The minus sign in (11) comes from the fact that in the region where the cosine is positive, namely where $\bar{\phi} < \pi/2$ (for example, as in the segment CB shown in Figure 1c), $\bar{r}$ is a decreasing function of $\bar{s}$.

Constitutive relations. The constitutive relations complete the set of differential equations governing the membrane's behavior. It is assumed that the membrane is isotropic hyperelastic, namely it possesses a strain energy function $W(\lambda_s, \lambda_\theta)$ per unit area of the undeformed surface, such that

$$N_s = W,_{\lambda_s}, \qquad N_\theta = W,_{\lambda_\theta}, \tag{12}$$

in the doubly tense regions.

In our basic model we consider a *Hookean* (linear elastic) isotropic material. In this case, the strain energy function is quadratic:

$$W(\lambda_s, \lambda_\theta) = \frac{Eh}{1 - \nu^2}\left(\tfrac{1}{2}(\lambda_s^2 + \lambda_\theta^2) + \nu\lambda_s\lambda_\theta - (1 + \nu)(\lambda_s + \lambda_\theta - 1)\right). \tag{13}$$

Here h is the membrane thickness, E is the Young's modulus and ν is the Poisson's ratio. With this strain energy function, the relations (12) in the doubly tense regions become

$$N_s = \frac{Eh}{1 - \nu^2}(\lambda_s + \nu\lambda_\theta) - \frac{Eh}{1-\nu}, \tag{14}$$

$$N_\theta = \frac{Eh}{1 - \nu^2}(\lambda_\theta + \nu\lambda_s) - \frac{Eh}{1-\nu}. \tag{15}$$

Modified equations in the wrinkle region. Equations (1), (2), (4)–(6), and (11) hold in the entire membrane. The other equations written above hold in the doubly tense regions, and have to be modified in the wrinkle region. Equation (3) is not relevant in the wrinkle region since λ_θ is not physical there. Equation (10) is simply replaced by $N_\theta = N_{\mathrm{cr}}$ in the wrinkle region. Here N_{cr} is a nonpositive quantity assumed to be a material property. For ideal membranes $N_{\mathrm{cr}} = 0$, and for nearly ideal membranes N_{cr} has a small negative (given) value.

The constitutive equations (14) and (15) are replaced by an effective constitutive equation in the wrinkle region. Since $N_\theta = N_{\mathrm{cr}}$ there, (15) furnishes then a relation between λ_s and λ_θ. This relation

can be used to eliminate λ_θ from (14), thus obtaining a modified unidirectional constitutive relation for N_s. After some algebra, the end result is

$$N_s = Eh(\lambda_s - 1) + \nu N_{\text{cr}}. \tag{16}$$

Geometrical relations at the lower edge. A few geometrical relations may be inferred from the setup shown in Figure 2, in the deformed configuration:

$$\bar{\phi}_0 = \beta + \pi/2, \tag{17}$$

$$\bar{r}_0 = L \sin\beta, \tag{18}$$

$$P = 2\pi \bar{r}_0 T \cos\beta. \tag{19}$$

From the last two relations, a simple relation between the unknown quantities β and T can be deduced:

$$T = \frac{P}{\pi L \sin 2\beta}. \tag{20}$$

Boundary and interface conditions. To complete the statement of the problem, the equations given above must be accompanied by boundary conditions at the two ends of the membrane. At the lower end (point C in Figure 2) we have

$$(\bar{N}_s)_0 = T, \tag{21}$$

since T is tangent to the meridian at this point. At the upper end (point E in Figure 2) we have

$$r_E = \bar{r}_E \quad \text{or} \quad (\lambda_\theta)_E = 1, \tag{22}$$

since the membrane is fixed to a rigid ring there. The second equality in (22) follows from (3).

We also need interface conditions at the meridional junctions of the doubly tense and wrinkle regions (points B and D in Figure 2). These conditions are:

$$N_\theta = N_{\text{cr}} \quad \text{and} \quad \bar{\phi} \text{ is continuous} \qquad \text{at B and D}. \tag{23}$$

From (6) and (16) we conclude that N_s and λ_s are also continuous at the interfaces. This completes the statement of the problem for given pressure p.

Global equilibrium. Since the entire parachute is in self equilibrium under load P and pressure p, we must have

$$P - I_E = 0, \tag{24}$$

where I_E is the value of I defined by (5) at point E, which is the total axial force due to the pressure p along the entire meridian. This equation will serve us in determining the pressure p.

We remark that from (24) and (6) it follows that at point E either $N_s = 0$ or $\bar{\phi} = 0$. However, we do not enforce this boundary condition directly, but rather enforce the global equilibrium condition (24), as will be explained in the next section.

3. Computational scheme

3.1. *The overall scheme.* The problem involves a single *given* loading parameter, namely the carried weight P. We approach the parachute problem via *incremental loading*, by increasing the load P in small discrete steps from zero to its final value. This serves two purposes. First, for a given undeformed geometry of the canopy, this provides us with a family of solutions for various carried weights. Second, this procedure helps us to detect critical values of P for which essential changes occur in the nature of the solution.

For a given value of P, the computational scheme that we use in order to solve the problem stated in the previous section involves two nested iteration processes. The outer iteration loop may be described as follows. We make an initial guess $\tilde{p}$ for the pressure value p. Then we step into the inner iteration loop (to be described in the next subsection) and compute all deformation and stress variables along the entire meridian. Based on this solution, we calculate the reaction R (see Figure 2) by

$$R = P - I_E. \tag{25}$$

Of course, we know that R should vanish, since the whole parachute is self-equilibrated; see (24). Computationally, we check if the criterion

$$|R| \le \delta_1 P \tag{26}$$

is satisfied, where δ_1 is a small given tolerance. If this criterion is satisfied, our guess was correct, and thus $\tilde{p} \simeq p$ and the whole problem is solved. If not, we change the guess $\tilde{p}$ and repeat the inner iteration loop. This process continues, until (26) is satisfied.

Despite the strong nonlinear nature of the whole problem, numerical experiments show that the relation between the reaction R and the assumed applied pressure $\tilde{p}$ is almost linear! This is demonstrated in Figure 3, where the reaction-pressure relation is shown for a typical set of input parameters (ideal membrane, elliptic initial shape, vertical force $P = 0.1$). This surprising fact can be explained by the following calculation. Let us evaluate the integral I defined by (5) at the upper point E (see Figure 2) for a guessed pressure value $\tilde{p}$:

$$I_E = 2\pi \tilde{p} \int_0^{\bar{s}_E} \bar{r}' \cos \bar{\phi}' \, d\bar{s}'. \tag{27}$$

Defining

$$\hat{I}_E = \int_0^{\bar{s}_E} \bar{r}' \cos \bar{\phi}' \, d\bar{s}', \tag{28}$$

and using (25), we have

$$R(\tilde{p}) = P - 2\pi \tilde{p} \hat{I}_E. \tag{29}$$

We calculate $\hat{I}_E$ using (28), (3), (11), and (22):

$$\hat{I}_E = \int_0^{\bar{s}_E} \bar{r}' \cos \bar{\phi}' \, d\bar{s}' = -\int_{\bar{r}_0}^{\bar{r}_E} \bar{r}' \, d\bar{r}' = -\tfrac{1}{2}(\bar{r}_E^2 - \bar{r}_0^2) = -\tfrac{1}{2}(r_E^2 - \bar{r}_0^2) = -\tfrac{1}{2}\left(r_E^2 - r_0^2(\lambda_\theta)_0^2\right). \tag{30}$$

Now we make the assumption that the hoop strain at the lower point is small, namely that $(\lambda_\theta)_0$ is close to one. (This assumption was verified numerically in the example problems that we have solved.) We

write this assumption is the form

$$(\lambda_\theta)_0 = 1 + \epsilon f(\tilde{p}) = 1 + O(\epsilon). \tag{31}$$

Then from the last calculation we have

$$\hat{I}_E = -\tfrac{1}{2}\left(r_E^2 - r_0^2 + O(\epsilon)\right). \tag{32}$$

Using this result in (29) finally yields

$$R(\tilde{p}) = P + \pi\left(r_E^2 - r_0^2 + O(\epsilon)\right)\tilde{p}. \tag{33}$$

This is an approximate linear relation between R and $\tilde{p}$, which explains the observation above.

The fact that R is almost proportional with $\tilde{p}$ is a considerable aid in the numerical convergence of the scheme. By taking advantage of this fact, one is able to find an excellent approximation for the pressure p by making only two calculations, with two values of $\tilde{p}$, and linearly extrapolating the results as implied by Figure 3.

Incidentally, from the calculation above we can deduce another almost linear relation, namely the relation between the true pressure p and the weight P. With the true pressure we should get $R = 0$, and from (29) we thus have

$$P - 2\pi p \hat{I}_E = 0. \tag{34}$$

Therefore,

$$p = \frac{P}{2\pi \hat{I}_E}, \tag{35}$$

and using (32) we have

$$p = -\frac{P}{\pi(r_E^2 - r_0^2 + O(\epsilon))}. \tag{36}$$
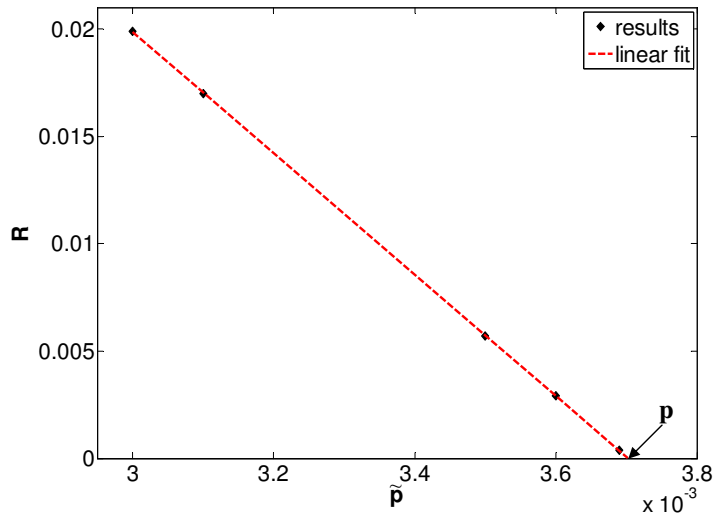
Figure 3. The reaction R as a function of the pressure $\tilde{p}$ for a typical set of input parameters. The five dots along the straight line correspond to the numerical results of five runs of the inner iteration loop with different values of $\tilde{p}$.

This proves the claim, which was also verified numerically. If the term $O(\epsilon)$ is dropped as an approximation, we get the simple result $p = P/[\pi(r_0^2 - r_E^2)]$ which can be obtained directly from summation of forces in the z direction, and is valid also in the case of no hole ($r_E = 0$).

3.2. *The inner-loop iterative procedure.* The inner-loop iteration procedure is a generalization of the one used in [Libai and Givoli 2002] to analyze a pulled membrane without internal pressure and with simpler boundary conditions. This procedure is based on a shooting technique and on approximately integrating the equations along the meridian. (The use of such a technique was suggested to the third author by D. J. Steigmann; see also [Li and Steigmann 1995a; 1995b]).

The numerical scheme is outlined as follows. The value of the pulling force P is given. Also given is the function $r(z)$ describing the shape of the unloaded membrane, where z ranges from 0 to the height of the undeformed membrane Z. We divide the z axis into M small intervals $[z_{m-1}, z_m]$ for $m = 0, \ldots, M$. For simplicity we take the intervals to be of uniform size $\Delta z \equiv z_m - z_{m-1} = Z/M$. We start from the lower edge of the membrane, that is, point C in Figure 2. To initiate the integration process, we first guess the value of the hoop stretch at point C, $(\lambda_\theta)_0$. Then we integrate the equations numerically as will be described shortly, until we reach the other edge, that is, point E. We check if the boundary condition (22) is satisfied at point E to within a desired precision; if it is, the solution process has terminated. If not, the guessed value of $(\lambda_\theta)_0$ is adjusted, and the integration starts all over again from point C. This process is repeated until the boundary condition (22) is satisfied at point E.

The scheme is now described in full detail. The value of a variable v at location z_m is denoted v_m. The pressure $\tilde{p}$ is assumed to be known, having been set in the outer loop.

(1) **The lower doubly tense zone (from point C to point B; see Figure 2).**

  (a) Choose an initial value $(\lambda_\theta)_0$ for the hoop stretch at point C.

  (b) Set $m = 0$, $s_0 = \bar{s}_0 = 0$, $z_0 = \bar{z}_0 = 0$ and $I_0 = 0$.

  (c) Use (3), (17), (18), and (20) to compute

  $$\bar{r}_0 = r_0(\lambda_\theta)_0, \quad \beta = \sin^{-1}(\bar{r}_0/L), \quad \bar{\phi}_0 = \beta + \pi/2, \quad T = P/(\pi L \sin 2\beta).$$

  (d) Use the equilibrium equation (6) to calculate

  $$(N_s)_m = (P - I_m)/(2\pi r_m \sin \bar{\phi}_m).$$

  (e) Use the constitutive equation (14) to calculate $(\lambda_s)_m$ from $(N_s)_m$ and $(\lambda_\theta)_m$:

  $$(\lambda_s)_m = \frac{1-v^2}{Eh}(N_s)_m - v(\lambda_\theta)_m + 1 + v.$$

  (f) Use the constitutive equation (15) to calculate $(N_\theta)_m$ from $(\lambda_s)_m$ and $(\lambda_\theta)_m$.

  (g) Check if $(N_\theta)_m \leq N_{cr}$. If yes, point B is reached (see Figure 2); move to step 2.

  (h) Calculate

  $$z_{m+1} = z_m + \Delta z, \quad \Delta r_m = r_{m+1} - r_m, \quad \Delta s_m = \sqrt{(\Delta r_m)^2 + (\Delta z)^2}, \quad s_{m+1} = s_m + \Delta s_m.$$

  (i) Use (2) to calculate

  $$\Delta \bar{s}_m = \Delta s_m (\lambda_s)_m, \qquad \bar{s}_{m+1} = \bar{s}_m + \Delta \bar{s}_m.$$

(j) Integrate the equilibrium equation (10) to calculate

$$\bar{\phi}_{m+1} = \cot^{-1}\left[\cot\bar{\phi}_m - \left(2\pi/(P - I_m)\right)\left(\Delta s_m (N_\theta)_m - \tilde{p}\bar{r}_m \Delta\bar{s}_m/\sin\bar{\phi}_m\right)\right].$$

(k) Use the compatibility equation (11) to calculate

$$\Delta\bar{r}_m = -\Delta\bar{s}_m \cos\bar{\phi}_{m+1}, \qquad \bar{r}_{m+1} = \bar{r}_m + \Delta\bar{r}_m.$$

(l) Use (3) to calculate

$$(\lambda_\theta)_{m+1} = \bar{r}_{m+1}/r_{m+1}.$$

(m) Calculate

$$\Delta\bar{z}_m = \sqrt{(\Delta\bar{s}_m)^2 - (\Delta\bar{r}_m)^2}, \qquad \bar{z}_{m+1} = \bar{z}_m + \Delta\bar{z}_m.$$

(n) Use (5) to calculate

$$I_{m+1} = I_m + 2\pi\tilde{p}\cos(\bar{\phi}_m)\bar{r}_m\Delta\bar{s}_m.$$

(o) Set $m \leftarrow m + 1$ and return to substep (d).

(2) **The wrinkle zone (from point B to point D; see Figure 2).**

    (a) Calculate

$$z_{m+1} = z_m + \Delta z, \qquad \Delta r_m = r_{m+1} - r_m, \qquad \Delta s_m = \sqrt{(\Delta r_m)^2 + (\Delta z)^2},$$

$$s_{m+1} = s_m + \Delta s_m, \qquad \Delta\bar{s}_m = \Delta s_m(\lambda_s)_m, \qquad \bar{s}_{m+1} = \bar{s}_m + \Delta\bar{s}_m.$$

    (b) Set $(N_\theta)_{m+1} = N_{\text{cr}}$.

    (c) Integrate the equilibrium equation (10) to calculate

$$\bar{\phi}_{m+1} = \cot^{-1}\left[\cot\bar{\phi}_m - \left(2\pi/(P - I_m)\right)(\Delta s_m N_{\text{cr}} - \tilde{p}\bar{r}_m\Delta\bar{s}_m/\sin\bar{\phi}_m)\right].$$

    (d) Use the compatibility equation (11) to calculate

$$\Delta\bar{r}_m = -\Delta\bar{s}_m \cos\bar{\phi}_{m+1}, \qquad \bar{r}_{m+1} = \bar{r}_m + \Delta\bar{r}_m.$$

    (e) Use (5) to calculate

$$I_{m+1} = I_m + 2\pi\tilde{p}\cos(\bar{\phi}_m)\bar{r}_m\Delta\bar{s}_m.$$

    (f) Use the equilibrium equation (6) to calculate

$$(N_s)_{m+1} = (P - I_{m+1})/(2\pi r_{m+1}\sin\bar{\phi}_{m+1}).$$

    (g) Use the constitutive equation (16) to calculate $(\lambda_s)_{m+1}$ from $(N_s)_{m+1}$:

$$(\lambda_s)_{m+1} = 1 + \left(1/(Eh)\right)\left((N_s)_{m+1} - \nu N_{\text{cr}}\right).$$

    (h) Calculate

$$\Delta\bar{z}_m = \sqrt{(\Delta\bar{s}_m)^2 - (\Delta\bar{r}_m)^2}, \qquad \bar{z}_{m+1} = \bar{z}_m + \Delta\bar{z}_m.$$

    (i) Use (3) to calculate a nonphysical value of λ_θ:

$$(\lambda_\theta^{\text{NP}})_{m+1} = \bar{r}_{m+1}/r_{m+1}.$$

(j) Use the constitutive equation (15) to calculate a nonphysical value of $(N_\theta)_{m+1}$, denoted by $(N_\theta^{\mathrm{NP}})_{m+1}$, from $(\lambda_s)_{m+1}$ and $(\lambda_\theta^{\mathrm{NP}})_{m+1}$.

(k) Check if $(N_\theta^{\mathrm{NP}})_{m+1} \geq N_{\mathrm{cr}}$. If yes — point D is reached (see Figure 2); skip to step 3.

(l) Set $m \leftarrow m + 1$ and return to substep (a).

(3) **The upper doubly tense zone (from point D to point E; see Figure 2).**

(a) Repeat substeps (d)–(f) and (h)–(o) of step 1. Continue the integration process till the last step $m = M$, namely till $z_m = Z$.

(b) Based on the value $(\lambda_\theta)_M$ at point E, calculate $\mu = |1 - (\lambda_\theta)_M|$. Check if $\mu < \delta_2$, where δ_2 is a predetermined small tolerance. If yes, the solution process has ended. Otherwise, adjust the value of $(\lambda_\theta)_0$ (see step 1, substep (a)) and start the whole process again.

Numerical experiments show that the adjustment of the value of $(\lambda_\theta)_0$ is easy in practice. The "output" $(\lambda_\theta)_M$ turns out to be a monotonely increasing function of the "input" $(\lambda_\theta)_0$, so by using a bisection-type technique the correct value of $(\lambda_\theta)_0$ can be found after a few trials. Numerical experiments also show that the convergence of the results as the number of subintervals M increases is fast. Moreover, the entire calculation is explicit and therefore very efficient. For example, we have used $M = 10,000$ for the problems presented in Section 5, and the running time in each case on a personal workstation was a few seconds. Thus, despite the fact that the solution process involves two nested iteration schemes, it is extremely fast.

4. Extensions

In the basic model described above the material was assumed isotropic and Hookean, the pressure distribution was assumed uniform, and the canopy was assumed to have a small hole in its center. Now we show how each of these assumptions may be relaxed.

4.1. *Orthotropic canopies.* In real parachutes, the canopy is stiffened with a finite number of meridional cords. Classical analysis [Taylor 1919] assumes that these cords are inextensible and the membrane forms lobes between each pair of adjacent cords. In the framework of our basic axisymmetric assumption, we cannot handle individual cords. Instead, we assume that the cords cause a *global* stiffening effect in the meridional direction. Thus we consider an effective canopy material which is homogeneous but orthotropic. Thus, different material properties are associated with the meridional and the circumferential directions. Of course, the only change caused by this extension in our formulation is in the constitutive equations (14)–(16).

We denote the meridional and circumferential Young's moduli by E_s and E_θ, and the corresponding Poisson's ratios by ν_s and ν_θ. Then, in the doubly tense regions, (14) and (15) are replaced by

$$N_s = \frac{E_s h}{1 - \nu_s \nu_\theta}(\lambda_s + \nu_s \lambda_\theta) - \frac{E_s h}{1 - \nu_s \nu_\theta}(1 + \nu_s), \tag{37}$$

$$N_\theta = \frac{E_\theta h}{1 - \nu_s \nu_\theta}(\lambda_\theta + \nu_\theta \lambda_s) - \frac{E_\theta h}{1 - \nu_s \nu_\theta}(1 + \nu_\theta). \tag{38}$$

In the wrinkle region, in which $N_\theta = N_{\mathrm{cr}}$, we obtain, in a similar manner to the isotropic case,

$$N_s = E_s h(\lambda_s - 1) + (E_s/E_\theta)\nu_s N_{\mathrm{cr}}, \tag{39}$$

which replaces (16).

These modified equations imply obvious changes in the algorithm presented in the previous section.

4.2. *Neo-Hookean material.* In the case of a neo-Hookean material, the strain energy function (13) is replaced by

$$W(\lambda_s, \lambda_\theta) = Ch\big(\lambda_s^2 + \lambda_\theta^2 + (\lambda_s \lambda_\theta)^{-2} - 3\big), \tag{40}$$

where C is a material constant. In this case the relations (12) in the doubly tense regions become

$$N_s = 2Ch(\lambda_s - \lambda_s^{-3}\lambda_\theta^{-2}), \tag{41}$$

$$N_\theta = 2Ch(\lambda_\theta - \lambda_s^{-2}\lambda_\theta^{-3}), \tag{42}$$

and the modified constitutive law in the wrinkle region becomes

$$N_s = 2Ch(\lambda_s - \lambda_s^{-2}). \tag{43}$$

It should be noted that this relation is valid for an ideal membrane only.

An algorithmic complication arises from the fact that the numerical scheme presented above requires (see steps 1, substep (e) and 2, substep (g)) the inversion of the constitutive equations (41) and (43), so that in the doubly tense regions λ_s is expressed explicitly in terms of N_s and λ_θ, and in the wrinkle region λ_s is expressed as a function of N_s. In the Hookean case this inversion was trivial, but in the neo-Hookean case it involves the solution of a quartic equation and a cubic equation. Closed-form formulas for the solutions of these two equations were obtained in [Libai and Givoli 2002] — an exact formula for the cubic equation and a 4th-order asymptotic solution for the quartic solution; see more details in [Libai and Givoli 2002].

4.3. *Nonuniform pressure.* The assumption of uniform pressure distribution is typically a reasonable one; in many cases the pressure on the canopy is nearly uniform except in a thin boundary layer near the outer edge of the membrane and the edge of the central hole (if any), where the pressure drops rapidly to the atmospheric pressure. However, in some cases a more significant deviation from uniformity occurs, which the analyzer may want to take into account.

To demonstrate this capability, we use as an example the pressure distribution found via computational fluid dynamics in [Sahu et al. 1995], and adopted in [Shannon 2001, p. 117]. According to this result, the internal pressure on the canopy is almost uniform but the external pressure is nonuniform, leading to a nonuniform net pressure, that reaches a maximum around the center of the meridian. By using a piecewise cubic polynomial approximation we interpolate this pressure distribution. This is shown in Figure 4 as a function $f(\chi)$ defined over the unit interval $\chi \in [0, 1]$. The actual pressure $p(s)$ can then be expressed as a function of the undeformed arc length coordinate s by

$$p(s) = p_0 f(s/S). \tag{44}$$

Here p_0 is the pressure magnitude, which is unknown as in the uniform-pressure case, and S is the total arc length of the meridian in the undeformed configuration, which can easily be computed in advance from the known geometry of the undeformed meridian.

Obvious changes have to be applied to the algorithm presented in the previous section in order to replace the uniform pressure by the nonuniform pressure (44).
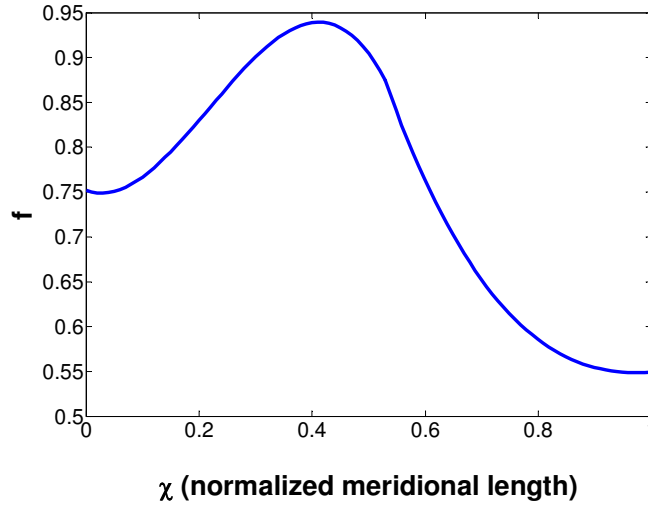
Figure 4. Nonuniform pressure distribution function, after [Sahu et al. 1995].

4.4. *Canopy without a hole.* The most involved extension to the basic setup presented in Section 2 is the consideration of a canopy without a central hole, as shown in Figure 5.

The governing equations in this case are identical to those in Section 2. However, the boundary condition (22) at point E needs to be modified. At the apex E we require that the membrane be continuous and have no kink (and thus must be flat there), namely

$$\bar{r}_E = 0, \qquad \bar{\phi}_E = 0. \tag{45}$$

From the condition $\bar{\phi}_E = 0$ and from (6), we deduce that (24) is also satisfied, and thus we do not need to require global equilibrium as an additional condition.

A numerical difficulty arises when trying to apply the inner-loop scheme presented in Section 3.2 to this case. In the close vicinity of the apex, owing to the conditions $\bar{r}_E = 0$ and $\bar{\phi}_E = 0$, we have $\bar{r} \simeq 0$, $\bar{\phi} \simeq 0$, and $P - I \simeq 0$, and we also know that $r \simeq 0$. These approximate equalities give

Figure 5. Parachute model with no hole.

rise to large roundoff errors near the apex. For example, using $(\lambda_\theta)_{m+1} = \bar{r}_{m+1}/r_{m+1}$ (see the inner-loop scheme, step 1, substep (l), which is repeated in step 3) becomes impossible, since the quantity on the right side approaches a 0/0 type limit at the apex.

To circumvent this difficulty, we approximate the deformed shape of the membrane in the close vicinity of the apex by the parabola

$$\bar{z}(\bar{r}) = a\bar{r}^2 + b\bar{r} + c, \tag{46}$$

where a, b, and c are constants. A similar technique has been used in the context of hot air balloon analysis [Irvine and Montauban 1980]. In that paper, the governing equations are simple enough to observe that in the limit $r \to 0$ an equation of a circle is obtained near the apex; therefore, for a $1°$ arc near the apex the shape of the membrane is assumed to be circular. In our case, the equations are not simple enough to observe the actual limit. However, in assuming the parabolic shape (46) we are guided by the fact that any smooth contour can be locally approximated by a parabolic curve.

We assume that (46) is the form of the deformed membrane when the condition $\bar{\phi} < \epsilon_\phi$ is satisfied, where ϵ_ϕ is a predetermined tolerance. This defines the point Q on the meridian, where the usual equations are to be replaced by modified equations based on the approximation (46). Since the parabola must be flat at the apex, we require

$$\left.\frac{\partial \bar{z}}{\partial \bar{r}}\right|_{\bar{r}=0} \equiv b = 0. \tag{47}$$

With the form (46) and (47) given, it is easy to calculate the limit of $\bar{r}/\bar{\phi}$ at the apex, and we find

$$\bar{r}/\bar{\phi} \simeq \bar{r}/\tan\bar{\phi} = \bar{r}/(-d\bar{z}/d\bar{r}) = -1/(2a). \tag{48}$$

Equations (47) and (48) are used in the scheme to pose two practical conditions on a and b:

$$|b| < \epsilon_1, \qquad 1/2 + (\bar{r}/\bar{\phi})a < \epsilon_2, \tag{49}$$

where ϵ_1 and ϵ_2 are given tolerances. We note that the second condition involves the ratio $\bar{r}/\bar{\phi}$ that is numerically problematic at the apex. However, this condition is checked only at point Q, which is on the *border* of the apex region where roundoff error is not yet significant.

In the close vicinity of the apex, the canopy behaves like a flat circular membrane under axisymmetric tension. Therefore we can assume

$$\lambda_s \simeq \lambda_\theta \simeq \text{constant}, \qquad N_s \simeq N_\theta \simeq \text{constant}. \tag{50}$$

From this and (3) we deduce

$$\lambda_s = \lambda_\theta = \frac{\bar{r}_Q}{r_Q}. \tag{51}$$

From axial equilibrium calculated from the apex to the point under consideration (see (4), which is calculated from the lower edge) we obtain, with $\eta = \bar{S} - \bar{s}$,

$$2\pi \bar{r} \bar{N}_s \sin\bar{\phi} = \int_0^{\bar{\eta}} 2\pi p \bar{r}' \cos\bar{\phi}' \, d\bar{\eta}'. \tag{52}$$

We use $\sin\bar{\phi} \simeq \bar{\phi}$, $\cos\bar{\phi} \simeq 1$, and $\bar{\eta} \simeq \bar{r}$, calculate the integral, and finally use (48) to get

$$\bar{N}_s \simeq -p/4a. \tag{53}$$

Using this result with (50), (1), and (3) we have

$$N_\theta = N_s = -\frac{p}{4a}\frac{\bar{r}_Q}{r_Q}. \tag{54}$$

Accordingly, step 3 in the inner-loop scheme of Section 3.2 should be modified for the case of no central hole in the following manner:

(3) **The upper doubly tense zone (from point D to point E; see Figure 2).**

    (a) Repeat substeps (d)–(f) and (h)–(o) of step 1. Continue the integration process till the condition $\bar{\phi} < \epsilon_\phi$ is satisfied. Once this condition is satisfied, point Q is reached; denote the current meridional-step m by m_Q, and proceed to (b).

    (b) Use the values of $\bar{r}$ and $\bar{z}$ at the last three grid points (m_Q, $m_Q - 1$, and $m_Q - 2$) to perform parabolic extrapolation and compute the constants a, b, and c in the parabola (46).

    (c) Check the condition $1/2 + (\bar{r}_Q/\bar{\phi}_Q)a < \epsilon_2$ (see (49)). If it is satisfied, proceed to (d). Otherwise, adjust the value of $(\lambda_\theta)_0$ (see step 1, substep (a)) and start the whole process again.

    (d) Calculate $\lambda_s = \lambda_\theta = \bar{r}_Q/r_Q$ and $N_\theta = N_s = -(p/4a)(\bar{r}_Q)(r_Q)$. These values are valid in the interval from Q to E.

    (e) For $m = m_Q, \ldots, M$ calculate $\bar{r}_m = \lambda_\theta r_m$, $\bar{\phi}_m = -2a\bar{r}_m$, and $\bar{z}_m = a\bar{r}_m^2 + c$.

In the *outer loop* (see Section 3.1), we check if the condition $|b| < \epsilon_1$ (see (49)) is satisfied. If it is, our guess for $\tilde{p}$ was correct, and thus $\tilde{p} \simeq p$ and the whole problem is solved. If not, we change the guess $\tilde{p}$ and repeat the inner iteration loop. This process continues, until the condition $|b| < \epsilon_1$ is satisfied.

5. Numerical examples

We apply the proposed method to two example canopies, which differ in their initial (undeformed) shape. We consider a parabolic undeformed membrane described by

$$r(z) = -1.9z^2 + 2.9z + 2, \quad z \in [0, 2], \tag{55}$$

and an elliptic undeformed membrane described by

$$r(z) = 3.3\sqrt{1 - (z-1)^2/4}, \quad z \in [0, 2.99]. \tag{56}$$

The elliptic undeformed canopy shape chosen here is inspired by [Lennon and Pellegrino 2005], and is such that the deformed shape turns out to be quite similar to the oval shape obtained in the analysis of [Taylor 1919], but without the lobes (see discussion in Section 1). Moreover, in Taylor's optimal parachute the ratio of minimum to maximum diameter is about 0.6, and this is also the ratio obtained in our case. The parabolic shape was chosen not due to physical reasons but simply because this was the family of shapes considered in our previous investigation (see [Libai and Givoli 2002]).

We assume uniformly distributed pressure, except in one case discussed at the end of this section.

We use nondimensional parameter values which were chosen based on the following normalization. Denoting dimensional quantities with a *, we first define two reference quantities: the length L^*, related to the height Z of the undeformed membrane, and the force $F^* = E^*h^*L^*$. Due to technical reasons that we omit here, we define $L^* = Z/2$ for the parabolic canopy and $L^* = Z/3$ for the elliptic canopy, although of course the choice $L^* = Z$ would have been equally legitimate. Then all the parameters and variables are normalized, according to their units, with respect to a product of powers of L^* and F^*. Thus, for example, $p = p^*/(E^*h)$ and $T = T^*/(E^*h^*)$. This normalization method always gives the nondimensional quantity $Eh = 1$. All parameters that appear in the examples below are unitless, having been defined by using this method of normalization. In all the examples shown here we use the Hookean material law, and $\nu = 0.5$.
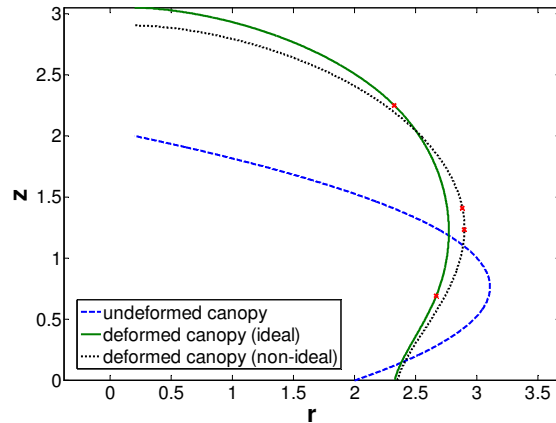
Figure 6. Undeformed and deformed shapes of a parabolic canopy with a central hole, under the applied load $P = 0.4$.

The computational tolerances needed for the double iteration loop are set as follows. In the case of a canopy with a central hole we use the tolerances $\delta_1 = 0.01$ and $\delta_2 = 10^{-3}$, while in the case of a canopy without a central hole we use $\epsilon_\phi = 0.07$, $\epsilon_1 = 10^{-4}$, and $\epsilon_2 = 5 \cdot 10^{-3}$. In all cases we take 10,000 steps in the z direction. This choice is based on practical convergence analysis; see [Ofir 2009] for the details and a convergence graph.

Figure 6 shows the undeformed and deformed shapes of the parabolic canopy with a central hole, under the load $P = 0.4$. The deformed shapes of both an ideal membrane and a nearly ideal membrane with $N_{cr} = -0.05$ are shown. Here and in subsequent figures, the boundaries of the wrinkle zone are marked on the deformed meridian by two × marks. It is clear that the wrinkle zone is much larger in the ideal membrane, as it should be. We also see that the deformed membranes are almost flat at the upper edge, which is consistent with the discussion at the end of Section 2.2.

Figure 7 shows the size of the wrinkle zone, relative to the entire meridian length, as a function of the applied load P, for four values of N_{cr}. In the case of the ideal membrane ($N_{cr} = 0$), the wrinkle
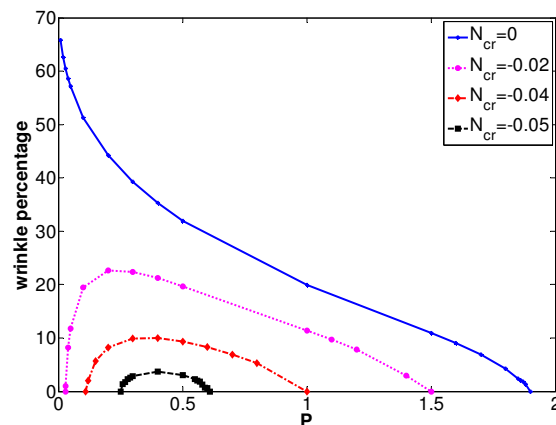
Figure 7. The wrinkle zone size as a function of the applied load P in a parabolic canopy with a central hole.
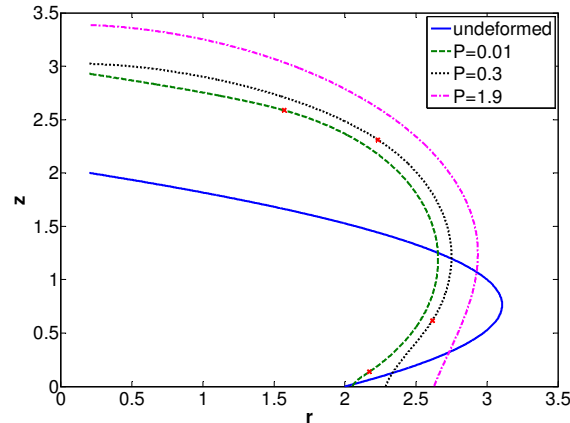
Figure 8. Response of a parabolic canopy with a central hole: undeformed and deformed shapes for three values of the vertical load P.

zone shrinks as P increases, up to a load value of about $P = 1.9$, beyond which the entire membrane becomes doubly tense. In the case of a nearly ideal membrane (that is, for *any* nonzero value of N_{cr}), the wrinkling behavior is considerably different. When the applied forces are very small, no wrinkles are present since the membrane has a certain capability of carrying compressional hoop stresses. When the applied loads are large enough, wrinkles start to form. Increasing the load further causes the wrinkle zone to expand, up to a maximal size. Increasing the load even further causes the wrinkle zone to shrink until it totally vanishes. The wrinkled regime shrinks when the value of $|N_{cr}|$ is increased. The latter behavior is the more physical one, since real membranes are not ideal. Additional numerical examples in this context can be found in [Ofir 2009].

Figures 8 and 9 show the responses of the ideal parabolic canopy with a central hole, under an increasing applied load. In Figure 8 we show the undeformed and deformed shapes, and in Figure 9 the distribution of the stress resultants N_s and N_θ and the distribution of the stretches λ_s and λ_θ. Note that in the wrinkle zone $N_\theta \equiv 0$, and the hoop stretch λ_θ becomes nonphysical and is therefore not shown in this zone. For the three load values considered in these figures, the lowest one ($P = 0.01$) yields a very large wrinkle zone, the intermediate one ($P = 0.3$) gives a smaller wrinkle zone, and the highest one ($P = 1.9$) causes the membrane to be doubly tense along the entire meridian.

Figure 10 summarizes the response of a parabolic canopy with a central hole as a function of the applied load P. We show the maximal values along the meridian of the stress resultants N_s and N_θ and of the stretches λ_s and λ_θ. These functions are monotone but are far from being linear.

Figures 11 and 12 show the responses of a nearly ideal elliptic canopy with no central hole, under an increasing applied load. Here we set $N_{cr} = -0.002$. Again we show the undeformed and deformed shapes (Figure 11), and the distributions of the stress resultants N_s and N_θ and of the stretches λ_s and λ_θ (Figure 12). In this case the response of the membrane is more sensitive to changes of the load value than in the previous case. Already for $P = 0.4$ the entire meridian is doubly tense.

Figure 13 shows the response of the ideal elliptic canopy with a central hole in two cases: one in which the pressure along the meridian is uniformly distributed, and the other in which it is distributed nonuniformly according to the function shown in Figure 4. We show the two deformed shapes in Figure 13a and the two meridional stress resultants in Figure 13b, for $P = 0.3$. The difference between the two
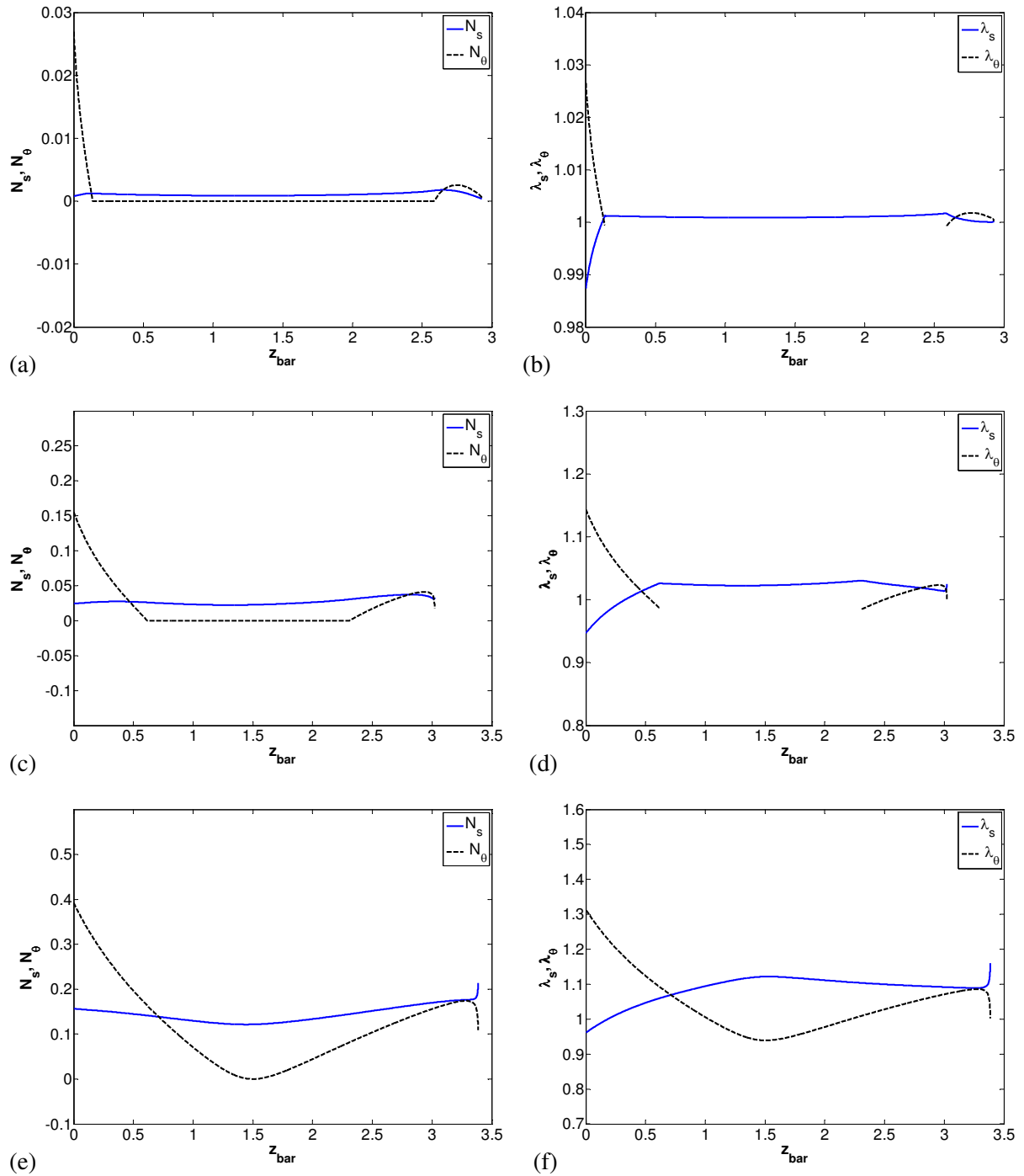
Figure 9. Response of a parabolic canopy with a central hole, stress resultant, and stretch distribution: (a) $P = 0.01$, stress resultants, (b) $P = 0.01$, stretches, (c) $P = 0.3$, stress resultants, (d) $P = 0.3$, stretches, (e) $P = 1.9$, stress resultants, (f) $P = 1.9$, stretches.
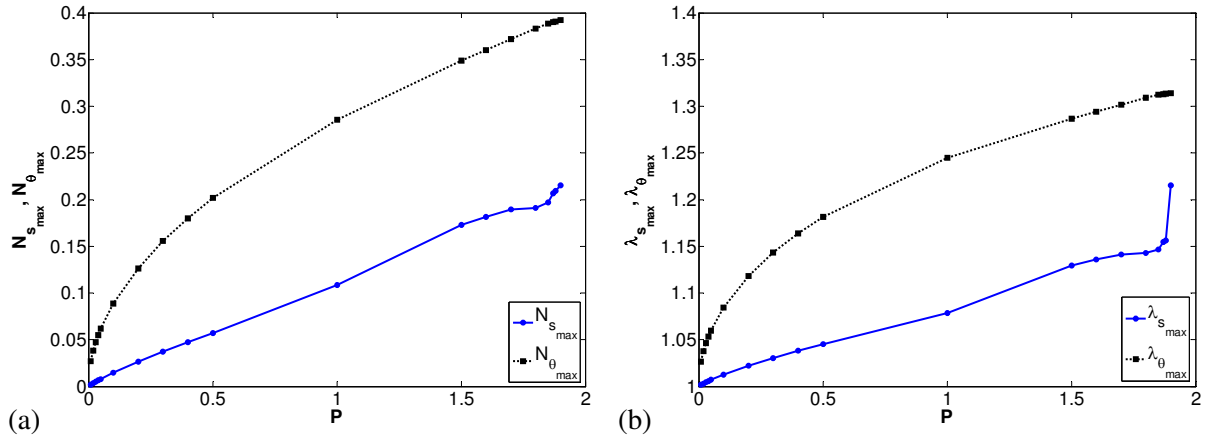
Figure 10. Response of a parabolic canopy with a central hole as a function of the applied load P: (a) maximal stress resultants, (b) maximal stretches.
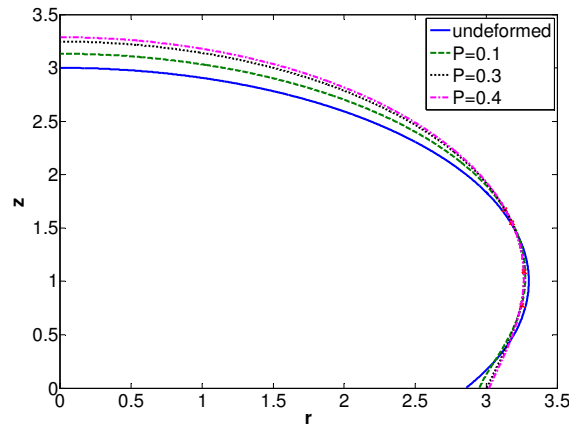
Figure 11. Response of a nearly ideal elliptic canopy with no central hole: undeformed and deformed shapes for three values of the vertical load P.

deformed shapes is small, although wrinkles are formed only in the case of uniform pressure distribution. However, the meridional stress in the vicinity of the hole is much larger in the case of uniform pressure. This shows that ignoring the effect of nonuniformity of pressure distribution on the mechanical response of the canopy is not always justified.

Finally we consider the case of an orthotropic canopy (see Section 4.1). We use an ideal membrane with the elliptic undeformed shape and a central hole, and with the vertical load $P = 0.3$. The nondimensional Young's modulus in the circumferential direction is still $E_\theta h = 1$, while in the meridional direction we take $E_s = 2E_\theta = 2$. The Poisson ratio is taken as 0.5 in both directions. In Figures 14a and 14b we compare the meridional stress resultant N_s and stretch λ_s, respectively, in the isotropic and orthotropic cases. We see that the orthotropy does not affect N_s along the meridian except near the hole surface, where the effect is very strong. The stretch λ_s is affected almost uniformly along the meridian. The deformation and circumferential stress and stretch (not shown here) are affected much less by the orthotropy.
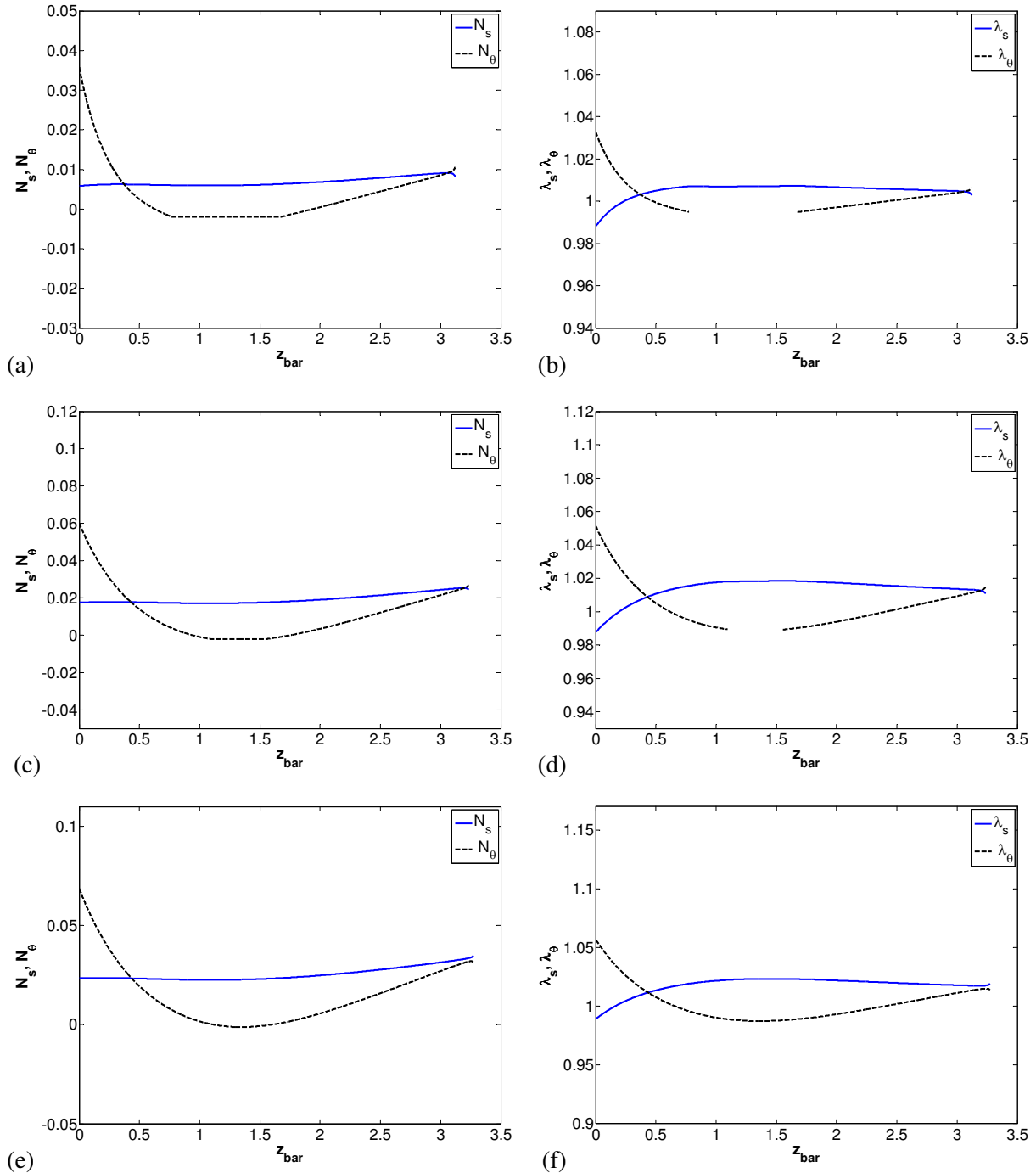
Figure 12. Response of a nearly ideal elliptic canopy with no central hole, stress resultants, and stretch distribution: (a) $P = 0.1$, stress resultants, (b) $P = 0.1$, stretches, (c) $P = 0.3$, stress resultants, (d) $P = 0.3$, stretches, (e) $P = 0.4$, stress resultants, (f) $P = 0.4$, stretches.
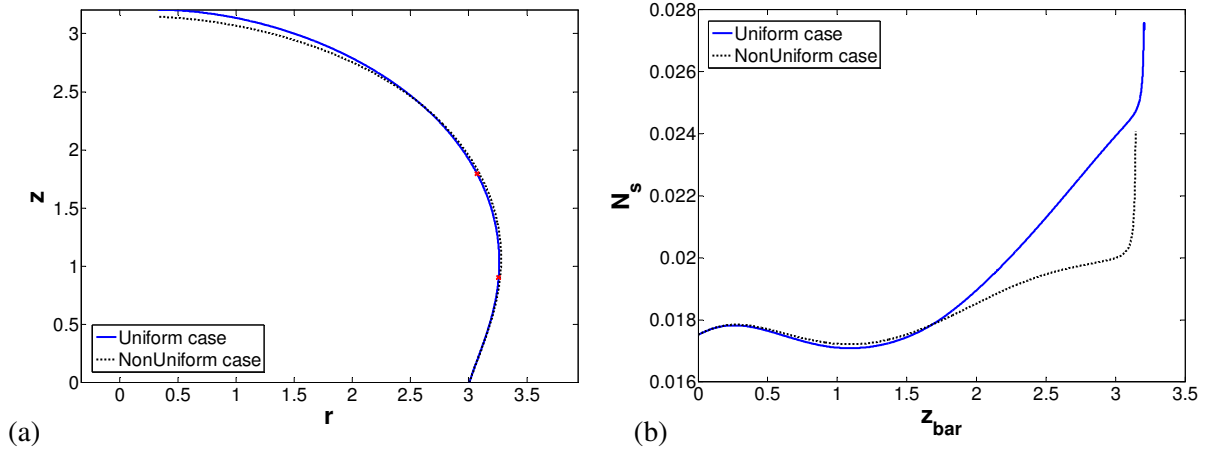
Figure 13. An ideal elliptic canopy with a central hole — responses in the cases of *uniformly and nonuniformly distributed pressure*: (a) deformed shapes, (b) meridional stress resultant.
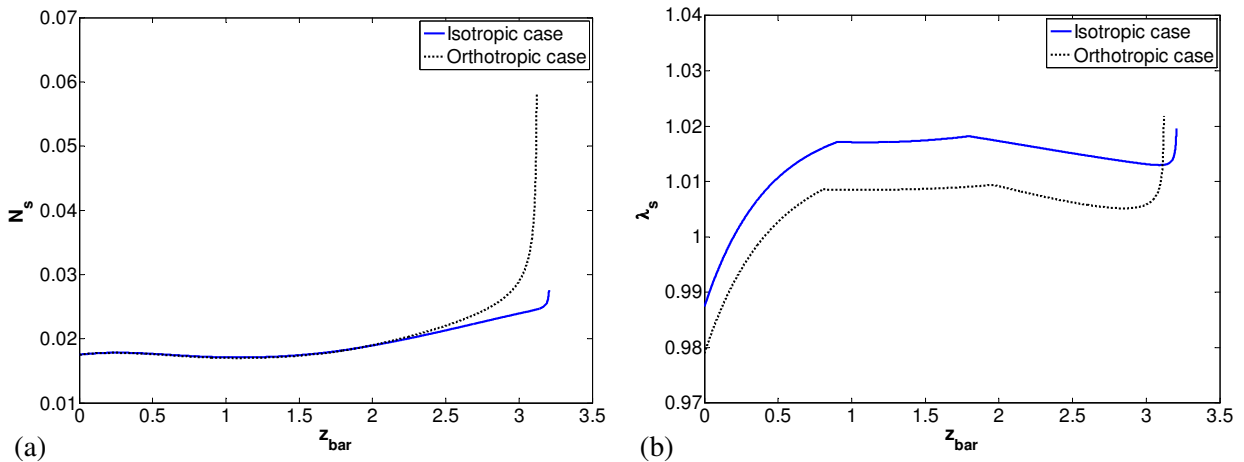
Figure 14. Comparison between the responses of isotropic and orthotropic canopies: (a) meridional stress resultant N_s, (b) meridional stretch λ_s.

6. Concluding remarks

The problem considered in this paper, of the quasistatic behavior of parachute canopies including the effects of large deformation and wrinkling, belongs to an important class of problems in nonlinear membrane shell theory. Other examples of such problems can be found, for example, in biology. In the present paper we have extended [Libai and Givoli 2002], which used wrinkle fields (also called tension fields) to model the wrinkle zones, but considered a much simplified setup.

The model introduced here, which represents a parachute under axisymmetric geometry, material properties, and loading, was analyzed via a double-iteration numerical procedure, based on a shooting technique and incremental loading. We first considered a canopy with a small central hole, made of an

isotropic Hookean material, and we assumed uniform pressure distribution. Later we showed how to relax these assumptions.

The advantage of the proposed numerical scheme employed here, over, say, the finite element method (FEM), is that it is quite simple, very easy to implement, and exhibits fast convergence, despite the complexity of the problem which involves the highly nonlinear membrane shell equations, different regions with interfaces which are unknown a priori, and deformation-dependent pressure magnitude. An attack of the same class of axisymmetric problems by FEM would be much more complicated. However, our method is restricted by the strong assumption of axial symmetry. It may be possible to extend it for more realistic nonaxisymmetric scenarios by using Fourier decomposition and considering the modal response to sinusoidal loading (see, for example, [Givoli and Libai 1995]). However, it may well be the case that for such problems one should resort to FEM.

References

[Baginski et al. 2006] F. Baginski, K. A. Brakke, and W. W. Schur, "Stability of cyclically symmetric strained pumpkin balloons and the formation of undesired equilibria", *J. Aircr.* **43**:5 (2006), 1414–1423.

[Cavicchi et al. 2009] A. Cavicchi, L. Gambarotta, and R. Massabò, "Computational modeling of reconstructive surgery: the effects of the natural tension on skin wrinkling", *Finite Elem. Anal. Des.* **45**:8-9 (2009), 519–529.

[Chiu et al. 1993] H. C. Chiu, R. C. Benson, M. D. Fiscella, and S. J. Burns, "Mechanical and thermal wrinkling of polymer membranes", pp. 1–4 in *Proc. 1993 ASME Winter Annual Meeting* (New Orleans, LA), ASME, New YorK, 1993.

[Deng and Pellegrino 2008] X. Deng and S. Pellegrino, "Computation of partially inflated shapes of stratospheric balloon structures", in *Proc. 49th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference* (Schaumburg, IL), 2008. paper AIAA-2008-2133.

[Givoli and Libai 1995] D. Givoli and A. Libai, "Incremental stresses in loaded orthotropic circular membrane tubes, II: numerical solution", *Int. J. Solids Struct.* **32**:13 (1995), 1927–1947.

[Irvine and Montauban 1980] H. M. Irvine and P. H. Montauban, "On hot air balloons", *Int. J. Mech. Sci.* **22**:10 (1980), 637–649.

[Jeong and Kwak 1992] D. G. Jeong and B. M. Kwak, "Complementarity problem formulation for the wrinkled membrane and numerical implementation", *Finite Elem. Anal. Des.* **12**:2 (1992), 91–104.

[Kim and Peskin 2006] Y. Kim and C. S. Peskin, "2-D parachute simulation by the immersed boundary method", *SIAM J. Sci. Comp.* **28**:6 (2006), 2294–2312.

[Kim and Peskin 2009] Y. Kim and C. S. Peskin, "3-D parachute simulation by the immersed boundary method", *Comput. Fluids* **38**:6 (2009), 1080–1090.

[Lennon and Pellegrino 2005] A. Lennon and S. Pellegrino, "Structural mechanics of lobed inflatable structures", pp. 1–12 in *Proc. European Conf. on Spacecraft Structures, Materials & Mechanical Testing* (Noordwijk, The Netherlands), European Space Agency, Noordwijk, The Netherlands, 2005. ESA SP-581.

[Li and Steigmann 1995a] X. Li and D. J. Steigmann, "Finite deformation of a pressurized toroidal membrane", *Int. J. Non-Linear Mech.* **30**:4 (1995), 583–595.

[Li and Steigmann 1995b] X. Li and D. J. Steigmann, "Point loads on a hemispherical elastic membrane", *Int. J. Non-Linear Mech.* **30**:4 (1995), 569–581.

[Libai 1990] A. Libai, "The transition zone near wrinkles in pulled spherical membranes", *Int. J. Solids Struct.* **26**:8 (1990), 927–939.

[Libai and Givoli 2002] A. Libai and D. Givoli, "Analysis of pulled axisymmetric membranes with wrinkling", *Int. J. Solids Struct.* **39**:5 (2002), 1259–1274.

[Libai and Simmonds 1998] A. Libai and J. G. Simmonds, *The nonlinear theory of elastic shells*, Cambridge University Press, 1998.

[Liu and Sze 2009] X. H. Liu and K. Y. Sze, "A corotational interpolatory model for fabric drape simulation", *Int. J. Numer. Methods Eng.* **77**:6 (2009), 799–823.

[Liu et al. 2001] X. Liu, C. H. Jenkins, and W. W. Schur, "Large deflection analysis of pneumatic envelopes using a penalty parameter modified material model", *Finite Elem. Anal. Des.* **37**:3 (2001), 233–251.

[Lu et al. 2001] K. Lu, M. Accorsi, and J. Leonard, "Finite element analysis of membrane wrinkling", *Int. J. Numer. Methods Eng.* **50**:5 (2001), 1017–1038.

[Mosler and Cirak 2009] J. Mosler and F. Cirak, "A variational formulation for finite deformation wrinkling analysis of inelastic membranes", *Comput. Methods Appl. Mech. Eng.* **198**:27-29 (2009), 2087–2098.

[Muttin 1996] F. Muttin, "A finite element for wrinkled curved elastic membranes, and its application to sails", *Commun. Numer. Methods Eng.* **12**:11 (1996), 775–785.

[Ofir 2009] Y. Ofir, "Analysis of stresses and wrinkles in a parachute canopy model using "wrinkle fields"", M.Sc. thesis, Dept. of Aerospace Engineering, Technion, Haifa, Israel, 2009, Available at http://www.graduate.technion.ac.il/Theses/Abstracts.asp?Id=24297.

[Pipkin 1986] A. C. Pipkin, "The relaxed energy density for isotropic elastic membranes", *IMA J. Appl. Math.* **36**:1 (1986), 85–99.

[Raible et al. 2005] T. Raible, K. Tegeler, S. Löhnert, and P. Wriggers, "Development of a wrinkling algorithm for orthotropic membrane materials", *Comput. Methods Appl. Mech. Eng.* **194**:21-24 (2005), 2550–2568.

[Rimrott and Cvercko 1986] F. P. J. Rimrott and M. Cvercko, "Wrinkling in thin plates due to inplane body forces", pp. 19–48 in *Inelastic behaviour of plates and shells: IUTAM symposium* (Rio de Janeiro, 1985), edited by L. Bevilacqua et al., Springer-Verlag, Berlin, 1986.

[Roddeman et al. 1987a] D. G. Roddeman, J. Drukker, C. W. J. Oomens, and J. D. Janssen, "The wrinkling of thin membranes, I: theory", *J. Appl. Mech. (ASME)* **54**:4 (1987), 884–887.

[Roddeman et al. 1987b] D. G. Roddeman, J. Drukker, C. W. J. Oomens, and J. D. Janssen, "The wrinkling of thin membranes, II: numerical analysis", *J. Appl. Mech. (ASME)* **54**:4 (1987), 888–892.

[Sahu et al. 1995] J. Sahu, G. Cooper, and R. Benney, "3D parachute descent analysis using coupled CFD and structural codes", *AIAA paper* (1995), 269–277. A95-1580-CP.

[Shannon 2001] M. P. Shannon, "Experimental analysis of the pressure distribution on a 35-foot personal parachute", *AIAA paper* (2001), 114–121. A01-29265.

[Steigmann 1990] D. J. Steigmann, "Tension-field theory", *Proc. R. Soc. Lond. A* **429**:1876 (1990), 141–173.

[Steigmann and Pipkin 1989] D. J. Steigmann and A. C. Pipkin, "Axisymmetric tension fields", *Z. Angew. Math. Phys.* **40**:4 (1989), 526–542.

[Stein et al. 2000] K. R. Stein, R. J. Benney, V. Kalro, T. E. Tezduyar, J. W. Leonard, and M. Accorsi, "Parachute fluid-structure interactions: 3-D computation", *Comput. Methods Appl. Mech. Eng.* **190**:3-4 (2000), 373–386.

[Stein et al. 2001] K. R. Stein, R. J. Benney, T. E. Tezduyar, J. W. Leonard, and M. Accorsi, "Fluid-structure interactions of a round parachute: modeling and simulation techniques", *J. Aircr.* **38**:5 (2001), 800–808.

[Tait and Connor 1997] R. J. Tait and P. Connor, "On the expansion of a deformed cylindrical elastic membrane", *IMA J. Appl. Math.* **59**:3 (1997), 231–243.

[Tait et al. 1996] R. J. Tait, D. J. Steigmann, and J. L. Zhong, "Finite twist and extension of a cylindrical elastic membrane", *Acta Mech.* **117**:1-4 (1996), 129–143.

[Taylor 1919] G. I. Taylor, "On the shapes of parachutes", technical report written for the U.K. Advisory Committee for Aeronautics, 1919. Reprinted as pp. 26–37 in *The scientific papers of sir Geoffrey Ingram Taylor*, vol. **3**, edited by G. K. Batchelor, Cambridge University Press, Cambridge, 1958.

[Wong and Pellegrino 2006] Y. W. Wong and S. Pellegrino, "Wrinkled membranes, II: analytical models", *J. Mech. Mater. Struct.* **1**:1 (2006), 27–61.

[Wu 1978] C. H. Wu, "Nonlinear wrinkling of nonlinear membranes of revolution", *J. Appl. Mech.* (*ASME*) **45** (1978), 533–538.

[Zak 1982] M. Zak, "Statics of wrinkling films", *J. Elasticity* **12**:1 (1982), 51–63.

YOAV OFIR: joav.ofir@gmail.com
Department of Aerospace Engineering, Technion – Israel Institute of Technology, Haifa 32000, Israel

DAN GIVOLI: givolid@aerodyne.technion.ac.il
Department of Aerospace Engineering, Technion – Israel Institute of Technology, Haifa 32000, Israel

AVINOAM LIBAI: aer7501@technion.ac.il
Department of Aerospace Engineering, Technion – Israel Institute of Technology, Haifa 32000, Israel

msp

CONICAL INDENTATION OF THICK ELASTIC SPHERICAL SHELLS

NKEM OGBONNA AND ALAN NEEDLEMAN

Indentation of spherical shells by a rigid conical indenter is analyzed. Because the shape change associated with overall bending provides a main deviation from the corresponding indentation behavior of a half-space, finite deformation analyses are carried out. The shell material is characterized by a hypoelastic relation that reduces to isotropic linear elasticity for small strains and is an elastic relation at finite strains for fixed principal directions. Axisymmetric deformations are assumed. For indentation of shells there is a complex interaction between the indentation response and the overall structural response of the shell. The aim of the analyses is to give insight into the range of indentation depths for which the indentation response dominates so that an indentation analysis can be used to extract the elastic properties. Shells having various thicknesses are analyzed. It is found that for spherical shells with thickness to radius ratios less than or equal to 0.4 and for sufficiently shallow indentation depths, the scaling of the indentation load with contact area has the same form as that for indentation into a half-space but with a different scaling constant.

1. Introduction

Indentation is a convenient method for investigating the mechanical properties of solids that are difficult to characterize by other mechanical testing methods. It is relatively nondestructive and is routinely used to measure the elastic stiffness and hardness of a wide variety of solids at a variety of scales; see, for example, [Doerner and Nix 1986; Oliver and Pharr 1992; Oliver and Pharr 2004]. Indentation has also been explored for determining the yield strength and strain hardening index of thin films [Cheng and Cheng 1998; Ma et al. 1998].

An indentation test is typically performed on a flat surface of a solid. An indenter of prescribed geometry is pressed into the solid under prescribed load and the depth of penetration and/or the size of the contact are measured. For depth-sensing indentation tests, the mechanical properties are determined directly from load and displacement measurement. Various indenter profiles are used: the most common ones include a cone (Rockwell test), a sphere (Brinell test), a rectangular pyramid (Vickers test) and a triangular pyramid (Berkovich test).

Much is known about indentation of solids, particularly elastic and elastic-plastic solids [Johnson 1985; Barber and Ciavarella 2000]. For indentation of a conical indenter into an isotropic linear elastic solid that can be appropriately idealized as being a half space, there is no characteristic dimension other than the indentation depth. Hence, from dimensional considerations,

$$\frac{P}{Eh^2} = K,$$

where P is the indentation force, E is Young's modulus, h is the indentation depth and K is a constant. The value of K depends on the shape of the cone cross section, the cone angle and Poisson's ratio, and can be determined from the solution of an elasticity problem. With K known, measuring the indentation force and depth enables E to be directly determined. A recent comprehensive overview of indentation into a planar linear elastic solid has been given in [Poon et al. 2008]. Much of the interest in indentation of linear elastic solids has been associated with characterizing the unloading response of elastic-plastic solids that have been indented [Oliver and Pharr 2004].

Some solids, such as natural and biological materials, may not have flat surfaces available for indentation. For example, the outer surface of the human eye is dome-shaped and, in order to determine intraocular pressure, doctors usually subject it to an indentation test using an instrument called a tonometer. The growing interest in natural materials, which often involve shell-type structures, as potential renewable resources for high performance and cost effective structural applications provides another motivation for examining the indentation behavior of solids with curved surfaces.

There is a literature on the deformation of thin elastic shells subject to point type loads that can give rise to complex deformation patterns; see for example [Vaziri and Mahadevan 2008]. However, the focus here is on relatively thick shells for which the deformation pattern remains axisymmetric when subject to conical indentation. We also neglect any effect of adhesion which can play an important role in contact of elastic solids as shown in [Johnson et al. 1971]. In this paper, we explore the extent to which simple scaling relations can be used to extract elastic moduli of isotropic elastic solids from indentation of spherical shells by a rigid conical indenter. In particular, we consider indentation of a rigid conical indenter with a circular cross-section into a thick walled isotropic linear elastic spherical shell. Indentation of a cone into a solid sphere has been considered in [Fu 2007]. A fundamental difference between indentation of a cone into a half space and its indentation into a spherical shell is that for the shell geometry there are characteristic lengths other than the indentation depth h. For a spherical shell with internal radius R_i and external radius R_o,

$$\frac{P}{Eh^2} = C\left(\frac{h}{R_i}, \frac{R_i}{R_o}\right),$$

so that extracting the value of Young's modulus E from a measurement of indentation force and depth is not necessarily as straightforward as for a half space. Although our interest is in linear elastic material response, we carry out finite deformation analyses because for shells a main deviation from linearity is associated with a change in shape due to bending.

2. Problem formulation

The problem we analyze is indentation of a linear elastic spherical shell by a rigid conical indenter allowing for the possibility of finite shape changes. The finite element implementation is based on a convected coordinate Lagrangian formulation of the field equations, with the initial unstressed state taken as reference. All field quantities are taken to be functions of convected coordinates, y^i, which serve as particle labels, and time, t. With attention confined to quasi-static deformations and with body

forces neglected, the rate form of the principal of virtual work is

$$\int_V [\dot{\tau}^{ij}\delta E_{ij} + \tau^{ij}\dot{u}^k_{,i}\delta u_{k,j}]\, dV = \int_S \dot{T}^i \delta u_i \, dS. \tag{2-1}$$

Here, V and S are the volume and surface, respectively, of the body in the reference configuration and $(\dot{\ }) = \partial()/\partial t$ at fixed y^i. The contravariant components of the Kirchhoff stress, τ^{ij}, are given by $\tau^{ij} = J\sigma^{ij}$, where J is the ratio of current to reference volume of a material element and σ^{ij} are the contravariant components of the Cauchy (or true) stress on the deformed convected coordinate net.

The nominal traction components, T^i, and the Lagrangian strain components, E_{ij}, are given by

$$T^i = (\tau^{ij} + \tau^{kj}u^i_{,k})\nu_j, \tag{2-2}$$

$$E_{ij} = \tfrac{1}{2}(u_{i,j} + u_{j,i} + u^k_{,i}u_{k,j}), \tag{2-3}$$

where u_j are the components of the displacement vector on base vectors in the reference configuration, $(\ )_{,i}$ denotes covariant differentiation in the reference frame and ν_j are the surface normal components in the reference configuration.

The material is described by a constitutive relation of the form

$$\hat{\tau}^{ij} = L^{ijkl}\dot{E}_{kl}, \tag{2-4}$$

where $\hat{\tau}^{ij}$ is the Jaumann derivative of the Kirchhoff stress and

$$L^{ijkl} = \frac{E}{1+\nu}\left[\frac{1}{2}\left(G^{ik}G^{jl} + G^{il}G^{jk}\right) + \frac{1}{1-2\nu}G^{ij}G^{kl}\right], \tag{2-5}$$

with G^{ij} being the contravariant components of the metric tensor in the current configuration. The relation in (2-4) and (2-5) is a hypoelastic relation but does reduce to isotropic linear elasticity for small strains and is an elastic relation at finite strains for fixed principal directions.

Figure 1 is a sketch of the indenter and the definition of the angle β. The contact radius a is defined as the actual contact radius accounting for sink in. The corresponding indentation depth h is then defined by

$$h = a\tan\beta \tag{2-6}$$

The spherical shell has internal radius R_i and external radius R_o so that the shell thickness is $t = R_o - R_i$. Axisymmetric deformations are assumed. Figure 2 shows the region analyzed and finite element mesh used for a shell with $t/R_o = 0.4$.
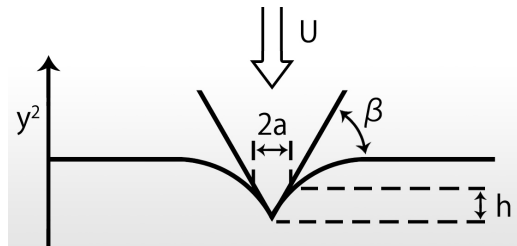
Figure 1. A sketch of the indenter geometry.
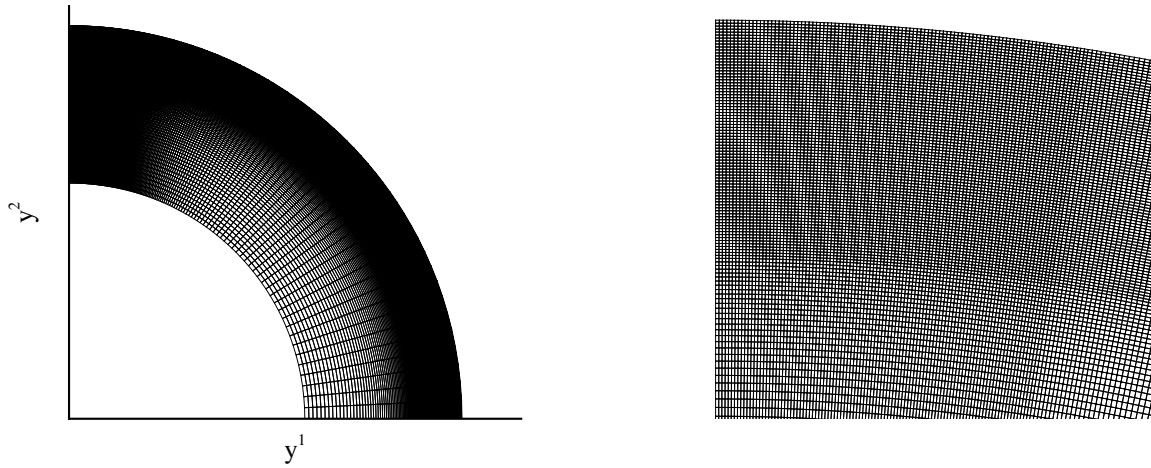
Figure 2. The finite element mesh for a spherical shell with $t/R_o = 0.4$. Each quadrilateral consists of four "crossed" triangular linear displacement elements. Left: the full mesh. Right: the fine mesh region.

Perfect sticking is assumed as soon as the shell comes into contact with the indenter, so that the rate boundary conditions are

$$\dot{u}_1 = 0 \qquad \dot{u}_2 = -\dot{U} \qquad \text{on} \quad S_{\text{contact}}, \tag{2-7}$$

where S_{contact} denotes the portion of the shell surface in contact with the indenter and U is the imposed displacement of the rigid indenter. Note that U is not equal to h as defined here where h is the depth associated with the contact radius. Initially, S_{contact} is the single point $y^1 = 0$, $y^2 = R_o$. As additional points come into contact, S_{contact} expands.

The symmetry conditions on the y^2-axis are

$$\dot{u}_1 = 0 \qquad \dot{T}^2 = 0 \quad \text{on} \quad y^1 = 0. \tag{2-8}$$

Symmetry is also assumed about the y^1-axis so that

$$\dot{u}_2 = 0 \qquad \dot{T}^1 = 0 \quad \text{on} \quad y^2 = 0, \tag{2-9}$$

and the remaining external surfaces are traction free so that $\dot{T}^1 = \dot{T}^2 = 0$.

The finite element discretization of (2-1) is based on linear displacement triangles arranged into "crossed" triangle quadrilaterals and the deformation history is calculated in an incremental manner. Calculations are carried out for four values of the thickness to radius ratio, t/R_o, 0.1, 0.2, 0.4 and 0.6. Different finite element meshes are used for each value of t/R_o. For $t/R_o = 0.6$, a 155×220 quadrilateral mesh is used with a uniform 100×100 quadrilateral mesh in a region $t/R_o = 0.1 \times \Delta\theta = 9°$ near the indenter. Here, $R = \sqrt{(y^1)^2 + (y^2)^2}$ and $\theta = \tan^{-1}(y^2/y^1)$. This mesh has 34100 quadrilaterals and 68592 degrees of freedom after static condensation is used to eliminate the central node in each quadrilateral. The mesh for $t/R_o = 0.4$, shown in Figure 2, has 140×220 quadrilaterals, giving 62322

degrees of freedom after static condensation and the uniform mesh is 80×100 in a $t/R_o = 0.1 \times \Delta\theta = 9°$ region near the indenter. The corresponding values for $t/R_o = 0.2$ are 100×220, 44642 and an 80×100 uniform mesh in a $t/R_o = 0.1 \times \Delta\theta = 9°$ uniform region and for $t/R_o = 0.1$, 100×200, 40602 and an 90×100 uniform mesh in a $t/R_o = 0.15 \times \Delta\theta = 9°$ uniform region.

3. Results

For indentation of a rigid conical indenter into a flat surface we have (see [Sneddon 1948; Johnson 1985])

$$\frac{P}{Ea^2} = \frac{\pi}{2} \frac{1}{(1-v^2)} \tan \beta. \tag{3-1}$$

Alternatively, expressed in terms of the contact depth $h = a \tan \beta$,

$$\frac{P}{Eh^2} = \frac{\pi}{2} \frac{1}{(1-v^2)} \frac{1}{\tan \beta} = \frac{\pi}{2} \frac{1}{(1-v^2)} \cot \beta. \tag{3-2}$$

The relation (3-2) is sometimes expressed in terms of the displacement of the rigid indenter, here termed U, for example in [Poon et al. 2008]. For indentation of a rigid cone into a linear elastic half-space $h = (2/\pi)U$; see [Bower et al. 1993], for example.

With $\beta = 20°$, (3-2) gives $P/(Eh^2) = 4.74$. A finite element calculation using the formulation in Section 2 but for indentation into a flat surface gave $P/(Eh^2) = 4.90$.

For spherical shells, calculations were carried out for $t/R_o = 0.1, 0.2, 0.4, 0.6$. In all calculations, Poisson's ratio v, was taken to be 0.3 and, unless otherwise stated, the indenter angle β is fixed at 20°. The calculations were carried out for prescribed increasing displacement U.

The relation between h and U for the spherical shells analyzed is shown in Figure 3. For U/R_o sufficiently small, the half-space relation provides an accurate characterization and for $h/R_o = 0.4$ and
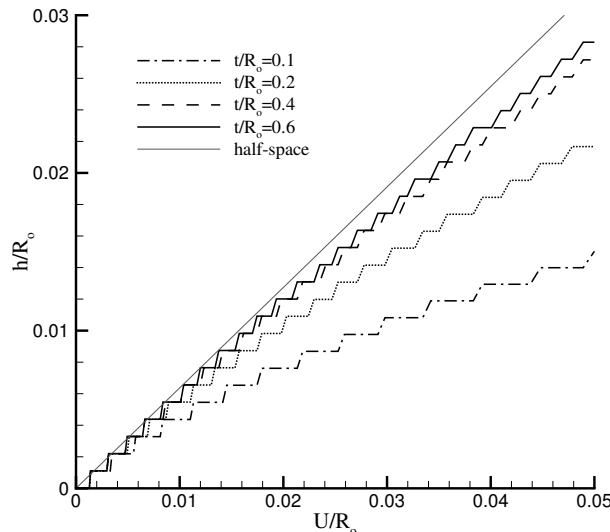
Figure 3. Variation of the indentation depth h associated with the current contact radius a versus the displacement of the rigid indenter U.
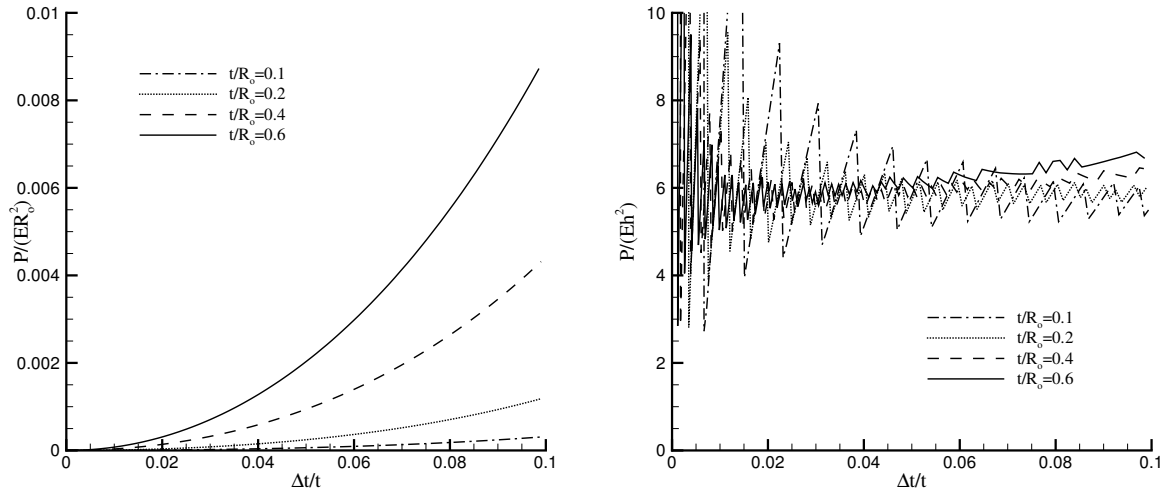
Figure 4. Variation of normalized contact force P/ER_o^2 (left) and normalized contact pressure P/Eh^2 with the change in shell thickness under the indenter (i.e., at $y^1 = 0$) for spherical shells with $t/R_o = 0.1, 0.2, 0.4, 0.6$.

0.6 that relation is approximately satisfied up about $U/R_o = 0.02$. For U/R_o greater than about 0.02 there are significant deviations with the deviation increasing with decreasing shell thickness.

Curves of contact force normalized by Young's modulus E and the outer sphere radius R_o, P/ER_o^2, plotted against $\Delta t/t$, are shown in Figure 4. The quantity $\Delta t/t$ is the change in thickness divided by the original thickness of the shell at $y^1 = 0$ and is convenient for representing the indentation depth. There is a clear dependence of the indentation force on the shell thickness, with the thinner shells requiring a smaller contact force for a given indentation depth. The indentation force at a given value of $\Delta t/t$ does not quite scale with t^2; for example, at $\Delta t/t = 0.08$, the indentation force with $t/R_o = 0.6$ is 7.8 times that for $t/R_o = 0.2$ and 30.4 times that for $t/R_o = 0.1$.

However, when the contact depth $h = a \tan \beta$ is used to normalize the indentation force there is a range where the value of P/Eh^2 is nearly the same for all four values of the shell thickness. The oscillations are a numerical artifact due to the contact radius increasing by an element length when a new node comes into contact with the indenter whereas the contact force increases smoothly. When only a few nodes are in contact with the indenter, the contact area is too small for an accurate value of P/Eh^2 to be obtained. A finer finite element mesh would permit accurate values of P/Eh^2 to be obtained for smaller indentation depths $\Delta t/t$. At larger values of indentation depth, a reasonably accurate estimate of P/Eh^2 can be obtained. For example, $\Delta t/t = 0.035$, the value of P/Eh^2 lies between 5.9 and 6.1 for all four values of t/R_o. The mean value of P/Eh^2 for $t/R_o = 0.1$ and 0.2 is 5.9. The mean values of P/Eh^2 are a bit higher and, furthermore, clearly increase with increasing indentation depth for $\Delta t/t$ greater than about 0.05. The value of normalized contact force, P/Eh^2, for the spherical shells is significantly greater than for indentation into a flat surface for the same material and the same indenter angle, ≈ 6 versus the numerically obtained value of 4.9 for indentation of a flat surface.

The dependence of the indentation pressure on the angle β is shown in Figure 5 for a spherical shell having $t/R_o = 0.4$. The evolution of the normalized contact force P/Eh^2 is plotted against the contact
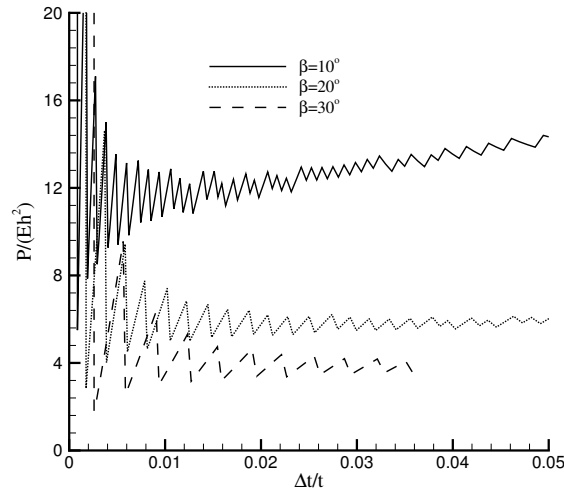
Figure 5. Variation of normalized contact pressure, P/Eh^2, with indenter angle β for spherical shells with $t/R_o = 0.4$.

depth as measured by $\Delta t/t$. For $\beta = 30°$, the mean value of P/Eh^2 is 3.8, for $\beta = 20°$ it is 5.9 and for $\beta = 10°$ it is about 12. For the most flat cone geometry, $\beta = 10°$, the mean value of P/Eh^2 increases with $\Delta t/t$ over the entire range for which a reliable estimate of P/Eh^2 can be obtained. This is not surprising since in the limit of a flat surface pressing against a spherical shell, the deformation is dominated by overall shell bending. Nevertheless, for all three values of β, the dependence of P/Eh^2 on $\tan \beta$ follows the scaling in (3-2). The ratio $\cot 30°/\cot 20°$ equals 0.631, while the corresponding ratio of normalized contact force in Figure 5 is 0.644. Even for $\beta = 10°$ we have $\cot 10°/\cot 20° = 2.06$, while $12/5.9 = 2.03$.

Distributions of mean normal stress $\tau_k^k/3$, where $\tau_k^k = G_{km}\tau^{km}$, normalized by Young's modulus E, are shown in Figure 6 for $t/R_o = 0.6$ and 0.2. These plots also show the deformed shape of the spherical shells at $\Delta_t/t = 0.1$. The mean normal stress is negative and a significant fraction of Young's modulus
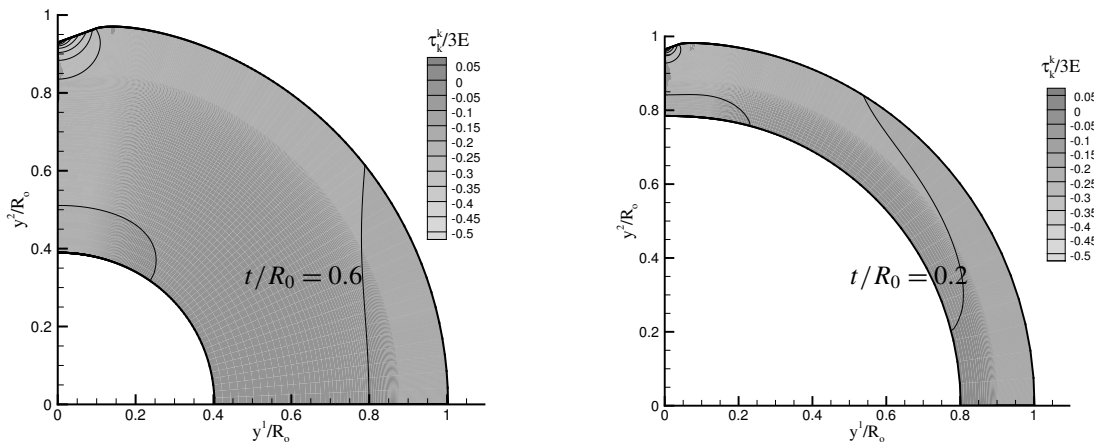
Figure 6. Contours of $\tau_k^k/3E$ at $\Delta t/t = 0.1$ for shells with $t/R_o = 0.6$ and $t/R_o = 0.2$.
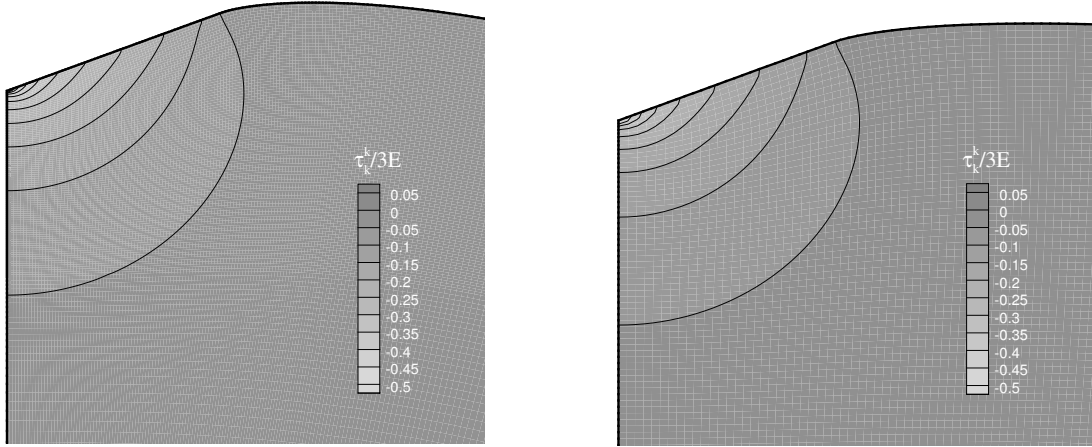
Figure 7. Contours of $\tau_k^k/3E$ at $\Delta t/t = 0.1$ in the vicinity of the indentation for the two cases shown in Figure 6: shells with $t/R_o = 0.6$ (left) and $t/R_o = 0.2$ (right).

E directly under the indenter. In fact, the stress is singular at the indenter tip and the peak value obtained there depends on the finite element discretization. In both cases, the values of τ_k^k are positive in a vicinity near $y^1 = 0$, $y^2 = R_i$ and in a vicinity near $y^1 = R_o$, $y^2 = 0$ which are indications of bending of the spherical shell. As can also be seen in the figure, the inner surface of the initially spherical shell does not quite remain spherical. The overall bending of the spherical shell gives rise to the deviation of P/Eh^2 from its value for indentation into a flat surface.

The distributions of $\tau_k^k/3E$ for these two cases in the vicinity of the indenter are shown in the bottom row in Figure 7. Near the indenter these distributions are very similar to those for indentation into a flat surface.

4. Concluding remarks

We have carried out finite deformation analyses of indentation of a rigid cone into linear elastic (actually linear hypoelastic) spherical shells. Our results indicate that if the indenter is not too flat and if the shell is not too thick, say $t/R_o \leq 0.4$, there is a regime of indentation depths where the scaling of the normalized contact force is of the same form as that for a flat surface. Hence, for isotropic elastic spherical shells in this regime we can write

$$\frac{P}{Sh^2} = C \cot \beta \quad \text{or} \quad \frac{P}{Sa^2} = C \tan \beta, \tag{4-1}$$

with $S = E/(1 - \nu^2)$ and $C \approx 2$ as opposed to $\pi/2 \approx 1.57$ for indentation into a flat surface, as in (3-1) and (3-2). Our results suggest that if the actual contact area ($a = h \cot \beta$) can be measured accurately, indentation into elastic shells may be able to be used to obtain the material's elastic stiffness even though the load is not a "pure" indentation load but also includes the effects of overall shell deformation.

For a spherical shell with $t/R_o \leq 0.4$, we find that C in (4-1) is only weakly dependent on the shell thickness. For a sufficiently thin spherical shell a complex non-axisymmetric deformation pattern, for example as discussed in [Vaziri and Mahadevan 2008], occurs and our axisymmetric analysis is no longer

applicable. This lower limit for the applicability of (4-1) cannot be ascertained by an axisymmetric analysis as carried out here. However, in the range where the scaling relation in (4-1) is applicable, a relatively good estimate of the elastic stiffness can be obtained without carrying out a finite element analysis. It is of interest to learn whether or not such simple scaling relations can be used to obtain an estimate of elastic stiffness for shells in more general circumstances than considered here.

Acknowledgment

References

[Barber and Ciavarella 2000] J. R. Barber and M. Ciavarella, "Contact mechanics", *Int. J. Solids Struct.* **37**:1-2 (2000), 29–43.

[Bower et al. 1993] A. F. Bower, N. A. Fleck, A. Needleman, and N. Ogbonna, "Indentation of a power law creeping solid", *Proc. R. Soc. Lond. A* **441**:1911 (1993), 97–124.

[Cheng and Cheng 1998] Y.-T. Cheng and C.-M. Cheng, "Further analysis of indentation loading curves: effects of tip rounding on mechanical property measurements", *J. Mater. Res.* **13**:4 (1998), 1059–1064.

[Doerner and Nix 1986] M. F. Doerner and W. D. Nix, "A method for interpreting the data from depth-sensing indentation instruments", *J. Mater. Res.* **1**:4 (1986), 601–609.

[Fu 2007] G. Fu, "An extension of Hertz's theory in contact mechanics", *J. Appl. Mech.* (*ASME*) **74**:2 (2007), 373–374.

[Johnson 1985] K. L. Johnson, *Contact mechanics*, Cambridge University Press, Cambridge, 1985.

[Johnson et al. 1971] K. L. Johnson, K. Kendall, and A. D. Roberts, "Surface energy and the contact of elastic solids", *Proc. R. Soc. Lond. A* **324**:1558 (1971), 301–313.

[Ma et al. 1998] D. Ma, K. Xu, and J. He, "Numerical simulation for determining the mechanical properties of thin metal films using depth-sensing indentation technique", *Thin Solid Films* **323**:1–2 (1998), 183–187.

[Oliver and Pharr 1992] W. C. Oliver and G. M. Pharr, "An improved technique for determining hardness and elastic modulus using load and displacement sensing indentation experiments", *J. Mater. Res.* **7**:6 (1992), 1564–1583.

[Oliver and Pharr 2004] W. C. Oliver and G. M. Pharr, "Measurement of hardness and elastic modulus by instrumented indentation: advances in understanding and refinements to methodology", *J. Mater. Res.* **19**:1 (2004), 3–20.

[Poon et al. 2008] B. Poon, D. Rittel, and G. Ravichandran, "An analysis of nanoindentation in linearly elastic solids", *Int. J. Solids Struct.* **45**:24 (2008), 6018–6033.

[Sneddon 1948] I. N. Sneddon, "Boussinesq's problem for a rigid cone", *Math. Proc. Cambridge Philos. Soc.* **44**:4 (1948), 492–507.

[Vaziri and Mahadevan 2008] A. Vaziri and L. Mahadevan, "Localized and extended deformations of elastic shells", *Proc. Nat. Acad. Sci. USA* **105**:23 (2008), 7913–7918.

NKEM OGBONNA: nkem42@yahoo.com
Department of Mathematics, Statistics and Computer Science, Michael Okpara University of Agriculture, Umudike, Abia State, Nigeria

ALAN NEEDLEMAN: needle@unt.edu
Department of Materials Science and Engineering, College of Engineering and Center for Advanced Scientific Computing and Modeling (CASCaM), University of North Texas, 1155 Union Circle #305310, Denton, TX 76203-50175, United States

∎msp

PROPAGATION OF WAVES IN AN INCOMPRESSIBLE TRANSVERSELY ISOTROPIC ELASTIC SOLID WITH INITIAL STRESS: BIOT REVISITED

RAY W. OGDEN AND BALJEET SINGH

In this paper, the general constitutive equation for a transversely isotropic hyperelastic solid in the presence of initial stress is derived, based on the theory of invariants. In the general finite deformation case for a compressible material this requires 18 invariants (17 for an incompressible material). The equations governing infinitesimal motions superimposed on a finite deformation are then used in conjunction with the constitutive law to examine the propagation of both homogeneous plane waves and, with the restriction to two dimensions, Rayleigh surface waves. For this purpose we consider incompressible materials and a restricted set of invariants that is sufficient to capture both the effects of initial stress and transverse isotropy. Moreover, the equations are specialized to the undeformed configuration in order to compare with the classical formulation of Biot. One feature of the general theory is that the speeds of homogeneous plane waves and surface waves depend nonlinearly on the initial stress, in contrast to the situation of the more specialized isotropic and orthotropic theories of Biot. The speeds of (homogeneous plane) shear waves and Rayleigh waves in an incompressible material are obtained and the significant differences from Biot's results for both isotropic and transversely isotropic materials are highlighted with calculations based on a specific form of strain-energy function.

1. Introduction

Initial stresses in solids have an important influence on the mechanical response of the material from an initially stressed configuration. Applications range from geophysics to the components of engineering structures and the behavior of soft biological tissues. The term *initial stresses* embraces situations in which the stress is accompanied by finite deformation from an unstressed configuration, in which case the term *pre-stresses* is commonly used, as, for example, in [Ieşan 1989], and situations in which the initial stress arises from some other process, such as manufacturing or growth, and is present in the absence of applied loads. In this latter case the initial stress is referred to as *residual stress* according to the definition of [Hoger 1985].

In the present paper we are concerned with the effect of initial stress on the propagation of small amplitude (linearized) elastic waves. A static theory of initial stress was developed long ago by Biot [1939]. He then extended it to wave propagation problems [1940]; this work is summarized in [Biot 1965]. In [Biot 1940] he states: "No assumption is made on how the initial state of stress is produced" — he requires only that it satisfy the equilibrium equations. Biot's theory has since formed the basis of many contributions to the literature, particularly in the geophysical context; see, for example, [Tolstoy

1982] and the more recent [Dey and De 1999; Sharma and Garg 2006]. The latter was concerned with an initially stressed anisotropic material and further references can be found therein. In the context of modern continuum theory, however, Biot's formulation of the equations is not straightforward, and part of the purpose of the present work is to show how Biot's formulation fits into a more general and more transparent framework.

Surface waves in elastic solids were first studied by Lord Rayleigh [1885] for an isotropic elastic solid. The extension of surface wave analysis and other wave propagation problems to anisotropic elastic materials has been the subject of many studies; see, for example, [Musgrave 1959; Anderson 1961; Stoneley 1963; Chadwick and Smith 1977; Royer and Dieulesaint 1984; Barnett and Lothe 1985; Mozhaev 1995; Nair and Sotiropoulos 1997; 1999; Destrade 2001a; 2001b; Destrade et al. 2002; Ting 2002a; 2002c; 2002b; Destrade 2003; Ogden and Vinh 2004]. For problems involving surface waves in a finitely deformed pre-stressed elastic solid (strain-induced anisotropy) we refer to [Hayes and Rivlin 1961; Flavin 1963; Chadwick and Jarvis 1979; Dowaikh and Ogden 1990; 1991; Norris and Sinha 1995 (concerning a solid/fluid interface); Chadwick 1997; Prikazchikov and Rogerson 2004 (concerning pre-stressed transversely isotropic solids); Destrade et al. 2005; Edmondson and Fu 2009]; see also [Song and Fu 2007]. As representatives of other works concerning waves in initially stressed elastic solids we cite [Norris 1983] on plane waves, the review [Guz 2002] and the analysis [Akbarov and Guz 2004] of waves in circular cylinders. Here we shall study the effect of initial stress on the propagation of surface waves based on a general formulation of the constitutive law of an elastic material that would be transversely isotropic in the absence of initial stress.

In Section 2, the equations governing small amplitude waves in a deformed transversely isotropic elastic solid with initial stress are derived, both for compressible and incompressible materials, the transverse isotropy being associated with a preferred direction in the initially stressed reference configuration. The constitutive law of the material is based on a strain-energy function (defined per unit reference volume) which depends on the combined invariants of the right Cauchy–Green deformation tensor, the initial stress tensor and the preferred direction. For a compressible material there are 18 such independent invariants in the general three-dimensional case, a number which reduces to 17 for an incompressible material. Expressions for the Cauchy stress and nominal stress tensors and the elasticity tensor are given in general forms but, because the large number of invariants makes the theory unwieldy in general, their forms are made explicit only for a restricted number of invariants, and attention is then confined largely to incompressible materials with seven invariants.

In Section 3, the equations of motion are specialized in order to study the effect of initial stress on the wave speed of homogeneous plane waves. It is noted, in particular, that the wave speed depends in a nonlinear fashion on the initial stress. In order to make contact with Biot's theory and to see how it sits within the general framework considered here, we give, in Appendix B, a derivation of Biot's equations, their connection with the equations herein, and a formula for the relation between the components of the elasticity tensor used here (which depend nonlinearly on the initial stress) and the components of Biot's elasticity tensor. In Section 4, the theory is specialized to two-dimensional motions (for incompressible materials) and then applied, in Section 5, to the study of Rayleigh waves in a half-space subject to initial stress parallel to its boundary with the preferred direction of transverse isotropy either parallel or normal to the boundary. The secular equation is derived and then specialized to give corresponding results for Biot's isotropic and orthotropic theories. The final section, Section 6, provides numerical results that

show the significant differences between the predictions of the general theory and of Biot's theory, and some concluding discussion is contained therein.

2. Equations of motion

We consider an elastic body whose initial geometry defines a *reference configuration*, which we denote by $\mathcal{B}_r$. In this configuration the body is in equilibrium and may in general be subject to a stress distribution, and we denote the Cauchy stress in $\mathcal{B}_r$ by T. If there is a body force b_r per unit mass acting then the equation of equilibrium is

$$\text{Div } T + \rho_r b_r = 0, \tag{1}$$

where ρ_r is the mass density of the material in $\mathcal{B}_r$ and Div is the divergence operator in $\mathcal{B}_r$, i.e., with respect to position vector X in $\mathcal{B}_r$. If the traction on the boundary $\partial\mathcal{B}_r$ of $\mathcal{B}_r$ vanishes pointwise then T is referred to as a *residual stress*, and it is necessarily non-uniform [Hoger 1985; Ogden 2003]. If the traction is not zero then we refer to T as an *initial stress* or *pre-stress*, and in general this may be accompanied by some prior deformation required to reach the configuration $\mathcal{B}_r$ from an unstressed state. Here we shall not be concerned with how the initial stress is produced.

A motion of the body from $\mathcal{B}_r$ may be described by a function χ so that the current position x of the material point initially at X is given by $x = \chi(X, t)$, where t is time. The deformation gradient tensor, denoted F, is given by $F = \text{Grad } \chi(X, t)$, where Grad is the gradient operator in $\mathcal{B}_r$. Let $\mathcal{B}$ be the configuration occupied by the body at time t.

The constitutive law of an elastic material may be described in terms of a strain-energy function, which is a function of F and defined per unit volume in $\mathcal{B}_r$. We denote this by $W(F)$, but note that in general W depends also on the initial stress T and on implicit material symmetries, which are suppressed for the present. Let S denote the nominal stress tensor in the configuration $\mathcal{B}$. Then, for a material not subject to any internal mechanical constraints S is given by

$$S = \frac{\partial W}{\partial F}. \tag{2}$$

For an incompressible material the constraint

$$\det F = 1 \tag{3}$$

is enforced and (2) is modified to

$$S = \frac{\partial W}{\partial F} - p F^{-1}, \tag{4}$$

where p is a Lagrange multiplier associated with the constraint.

The motion χ is governed by the equation

$$\text{Div } S + \rho_r b = \rho_r x_{,tt}, \tag{5}$$

where b is the body force acting in the current configuration, which may in general be different from b_r, and a subscript t following a comma signifies the material time derivative, i.e., the time derivative at fixed X, so that $x_{,t}$ is the particle velocity and $x_{,tt}$ the acceleration.

Although this is not strictly necessary, we now assume, for simplicity, that the body force is uniform and independent of the deformation. Then, $\boldsymbol{b} = \boldsymbol{b}_r$ and we get from (1) and (5)

$$\text{Div}(\boldsymbol{S} - \boldsymbol{T}) = \rho_r \boldsymbol{x}_{,tt}. \tag{6}$$

Now consider a finitely deformed equilibrium configuration $\mathcal{B}_0$ defined by $\boldsymbol{\chi}_0(\boldsymbol{X})$ and let $\boldsymbol{F}_0$ be the deformation gradient in this configuration and $\boldsymbol{S}_0$ the corresponding nominal stress. Then, $\text{Div}(\boldsymbol{S}_0 - \boldsymbol{T}) = \boldsymbol{0}$. Next, we consider an incremental motion from this latter configuration with displacement $\dot{\boldsymbol{x}} = \boldsymbol{\chi}(\boldsymbol{X}, t) - \boldsymbol{\chi}_0(\boldsymbol{X})$, and corresponding increment $\dot{\boldsymbol{F}} = \text{Grad}\,\dot{\boldsymbol{x}} = \boldsymbol{F} - \boldsymbol{F}_0$ in the deformation gradient. Let $\dot{\boldsymbol{S}} = \boldsymbol{S} - \boldsymbol{S}_0$ denote the increment in the nominal stress. Then,

$$\text{Div}\,\dot{\boldsymbol{S}} = \rho_r \dot{\boldsymbol{x}}_{,tt}. \tag{7}$$

For a compressible material the linearized form of $\dot{\boldsymbol{S}}$ is

$$\dot{\boldsymbol{S}} = \mathcal{A}\dot{\boldsymbol{F}}, \quad \dot{S}_{\alpha i} = \mathcal{A}_{\alpha i \beta j}\dot{F}_{j\beta}, \tag{8}$$

while for an incompressible material

$$\dot{\boldsymbol{S}} = \mathcal{A}\dot{\boldsymbol{F}} + p\boldsymbol{F}^{-1}\dot{\boldsymbol{F}}\boldsymbol{F}^{-1} - \dot{p}\boldsymbol{F}^{-1}, \tag{9}$$

where $\dot{p}$ is the increment in p and $\mathcal{A}$ is the elasticity tensor, which, for either a compressible or an incompressible material, is defined by

$$\mathcal{A} = \frac{\partial^2 W}{\partial \boldsymbol{F}\,\partial \boldsymbol{F}}, \quad \mathcal{A}_{\alpha i \beta j} = \frac{\partial^2 W}{\partial F_{i\alpha}\partial F_{j\beta}} = \mathcal{A}_{\beta j \alpha i}, \tag{10}$$

evaluated in the configuration $\mathcal{B}_0$. The convention that Greek indices refer to the configuration $\mathcal{B}_r$ and Roman indices to $\mathcal{B}_0$ is adopted here. The linearized incompressibility condition is

$$\text{tr}(\dot{\boldsymbol{F}}\boldsymbol{F}_0^{-1}) = 0. \tag{11}$$

For details of the background on the theory of incremental deformations superimposed on a finite deformation we refer to [Ogden 1984; 2007], for example.

It is convenient to work with $\mathcal{B}_0$ as the reference configuration, which requires that all quantities are updated, i.e., pushed forward, from $\mathcal{B}_r$ to $\mathcal{B}_0$ and incremental quantities are treated as functions of $\boldsymbol{x}$ and t. In particular, we define the Eulerian form of the displacement vector by $\boldsymbol{u}(\boldsymbol{x}, t) = \dot{\boldsymbol{x}}(\boldsymbol{X}, t)$ via the connection $\boldsymbol{x} = \boldsymbol{\chi}_0(\boldsymbol{X})$. The updated forms of the incremental constitutive laws (8) and (9) are, respectively,

$$\dot{\boldsymbol{S}}_0 = \mathcal{A}_0 \boldsymbol{L}, \quad \dot{S}_{0pi} = \mathcal{A}_{0piqj} L_{jq}, \tag{12}$$

and

$$\dot{\boldsymbol{S}}_0 = \mathcal{A}_0 \boldsymbol{L} + p\boldsymbol{L} - \dot{p}\boldsymbol{I}, \tag{13}$$

where $\boldsymbol{L} = \text{grad}\,\boldsymbol{u}$ is the displacement gradient, $\boldsymbol{I}$ is the identity tensor and a subscript 0 indicates a pushed forward quantity. In particular,

$$J_0 \mathcal{A}_{0piqj} = F_{0p\alpha}F_{0q\beta}\mathcal{A}_{\alpha i \beta j}, \tag{14}$$

where $J_0 = \det \boldsymbol{F}_0 = \rho_r/\rho_0$ and ρ_0 is the density in $\mathcal{B}_0$. For an incompressible material $J_0 = 1$ and $\rho_0 = \rho_r$. The (linearized) incremental form of the incompressibility constraint (11) is then expressed as

$$\operatorname{tr} \boldsymbol{L} \equiv \operatorname{div} \boldsymbol{u} = 0. \tag{15}$$

For a compressible material, the (Cartesian) component form of the equation of motion is

$$(\mathcal{A}_{0piqj} u_{j,q}),_p = \rho_0 u_{i,tt}, \tag{16}$$

and for an incompressible material

$$(\mathcal{A}_{0piqj} u_{j,q}),_p - \dot{p},_i + p,_j u_{j,i} = \rho_r u_{i,tt}, \quad \text{with } u_{i,i} = 0. \tag{17}$$

Note that as well as possessing the major symmetry $\mathcal{A}_{0piqj} = \mathcal{A}_{0qjpi}$ induced by (10), $\mathcal{A}$ has the property

$$\mathcal{A}_{0piqj} + \delta_{jp}\sigma_{0iq} = \mathcal{A}_{0ipqj} + \delta_{ij}\sigma_{0pq} \tag{18}$$

for a compressible material, and

$$\mathcal{A}_{0piqj} + \delta_{jp}(\sigma_{0iq} + p\delta_{iq}) = \mathcal{A}_{0ipqj} + \delta_{ij}(\sigma_{0pq} + p\delta_{pq}) \tag{19}$$

for an incompressible material, where σ_{0ij} are the components of the Cauchy stress tensor $\boldsymbol{\sigma}_0$ in $\mathcal{B}_0$. These are easily established by considering the symmetry of Cauchy stress expressed in the form $\boldsymbol{FS} = \boldsymbol{S}^{\mathrm{T}}\boldsymbol{F}^{\mathrm{T}}$, taking the increment of this and then updating the reference configuration to $\mathcal{B}_0$ to obtain

$$\dot{\boldsymbol{S}}_0 + \boldsymbol{L}\boldsymbol{\sigma}_0 = \dot{\boldsymbol{S}}_0^{\mathrm{T}} + \boldsymbol{\sigma}_0\boldsymbol{L}^{\mathrm{T}}. \tag{20}$$

Suppose now that the material is homogeneous. This requires, in particular, that the initial stress $\boldsymbol{T}$ is uniform. We recall, however, that a residual stress cannot be uniform [Hoger 1985; Ogden 2003], so the following analysis does not apply if the initial stress is a residual stress. Suppose further that $\boldsymbol{\chi}_0(\boldsymbol{X})$ is a homogeneous deformation. Then, the configuration $\mathcal{B}_0$ is uniform and hence $\mathcal{A}$, $\mathcal{A}_0$ and p are constant, and the equations of motion (16) and (17) reduce to

$$\mathcal{A}_{0piqj} u_{j,pq} = \rho_0 u_{i,tt} \tag{21}$$

and

$$\mathcal{A}_{0piqj} u_{j,pq} - \dot{p},_i = \rho_r u_{i,tt}, \quad u_{i,i} = 0, \tag{22}$$

respectively.

In general $\mathcal{A}$, and hence $\mathcal{A}_0$, depends on the deformation through $\boldsymbol{F}$, on the initial stress $\boldsymbol{T}$ and on any material symmetry present in the configuration $\mathcal{B}_r$. Here we consider a transversely isotropic material with preferred direction $\boldsymbol{M}$ in $\mathcal{B}_r$, where $\boldsymbol{M}$ is a unit vector. To make the dependence of $\mathcal{A}$ (and $\mathcal{A}_0$) on these quantities explicit we consider the scalar invariants of the tensors involved.

2.1. *Invariant formulation.* By objectivity, the dependence of the strain-energy function W on $\boldsymbol{F}$ is through the right Cauchy–Green tensor $\boldsymbol{C}$, which is defined by $\boldsymbol{C} = \boldsymbol{F}^{\mathrm{T}}\boldsymbol{F}$, and we therefore consider W to depend on the invariants of the three tensors $\boldsymbol{C}$, $\boldsymbol{T}$ and $\boldsymbol{M} \otimes \boldsymbol{M}$ since the material properties are assumed to be independent of the sense of $\boldsymbol{M}$.

The invariants of C most commonly used are the *principal* invariants, defined by

$$I_1 = \operatorname{tr} C, \quad I_2 = \tfrac{1}{2}[(\operatorname{tr} C)^2 - \operatorname{tr}(C^2)], \quad I_3 = \det C. \tag{23}$$

The (anisotropic) invariants associated with M and C are usually taken as

$$I_4 = M \cdot (CM), \quad I_5 = M \cdot (C^2 M). \tag{24}$$

The notation $I_1, \ldots, I_5$ is fairly standard for these invariants; see, for example, [Merodio and Ogden 2002; 2003]. In the reference configuration $\mathcal{B}_r$ these reduce to $I_1 = I_2 = 3$, $I_3 = I_4 = I_5 = 1$. A set of independent invariants of T that do not involve M may be taken as

$$\operatorname{tr} T, \quad \operatorname{tr}(T^2), \quad \operatorname{tr}(T^3), \quad \operatorname{tr}(TC), \quad \operatorname{tr}(TC^2), \quad \operatorname{tr}(T^2 C), \quad \operatorname{tr}(T^2 C^2), \tag{25}$$

invariants of T independent of C as

$$M \cdot (TM), \quad M \cdot (T^2 M), \tag{26}$$

and invariants depending on C, M and T as

$$M \cdot (TCM), \quad M \cdot (TC^2 M), \quad M \cdot (T^2 CM), \quad M \cdot (T^2 C^2 M). \tag{27}$$

These are the only independent invariants, 18 in total. For an incompressible material we have $I_3 = 1$ and hence there are 17 independent invariants in this case. In the reference configuration $\mathcal{B}_r$ the fourth, fifth, sixth and seventh invariants in (25) and the invariants (27) reduce to the first two in (25) and the two in (26). For full discussion of the relevant background on invariants of tensors we refer to [Adkins 1960; Spencer 1971; Zheng 1994]. A set of invariants equivalent to the above has been used by [Hoger 1993a; 1996]. For related work concerned with the constitutive equations and material symmetry for a residually stressed elastic material we refer to [Coleman and Noll 1964; Hoger 1986; 1993b; Man and Lu 1987; Johnson and Hoger 1993; Man 1998; Saravanan 2008; Tanuma and Man 2008].

We have not for the moment defined particular notation for the invariants (25)–(27). In the general case suppose there are N invariants, which we denote by I_i, $i = 1, 2, \ldots, N$. Then, the expressions for the stress and elasticity tensors require the calculation of

$$\frac{\partial W}{\partial F} = \sum_{i=1}^{N} W_i \frac{\partial I_i}{\partial F} \tag{28}$$

and

$$\frac{\partial^2 W}{\partial F \partial F} = \sum_{i=1}^{N} W_i \frac{\partial^2 I_i}{\partial F \partial F} + \sum_{i=1}^{N} \sum_{j=1}^{N} W_{ij} \frac{\partial I_i}{\partial F} \otimes \frac{\partial I_j}{\partial F}, \tag{29}$$

where we have used the shorthand notations $W_i = \partial W / \partial I_i$, $W_{ij} = \partial^2 W / \partial I_i \partial I_j$, $i, j = 1, 2, \ldots, N$.

Such a large number of invariants is impractical for applications, so for simplicity we restrict attention to incompressible materials and to the following invariants, which capture the main features of the combined anisotropy and initial stress. In particular, we omit invariants that are nonlinear in T. Thus,

we consider

$$I_1 = \operatorname{tr} \boldsymbol{C}, \qquad I_4 = \boldsymbol{M} \cdot (\boldsymbol{C}\boldsymbol{M}), \qquad I_5 = \boldsymbol{M} \cdot (\boldsymbol{C}^2 \boldsymbol{M}), \tag{30}$$

$$I_6 = \operatorname{tr}(\boldsymbol{T}\boldsymbol{C}), \qquad I_7 = \operatorname{tr}(\boldsymbol{T}\boldsymbol{C}^2), \qquad I_8 = \boldsymbol{M} \cdot (\boldsymbol{T}\boldsymbol{C}\boldsymbol{M}), \qquad I_9 = \boldsymbol{M} \cdot (\boldsymbol{T}\boldsymbol{C}^2\boldsymbol{M}), \tag{31}$$

which identify the invariants I_6, I_7, I_8, I_9, while the standard notation I_1, I_4, I_5 is retained and, since we are considering incompressible materials, $I_3 \equiv 1$. We may also include $\operatorname{tr} \boldsymbol{T}$ and $\boldsymbol{M} \cdot (\boldsymbol{T}\boldsymbol{M})$ in the functional dependence of W since they do not depend on $\boldsymbol{C}$ and hence their derivatives with respect to $\boldsymbol{F}$ do not contribute to the stress or elasticity tensors. In the following we give explicit expressions for the stress tensors and the elasticity tensor based on this restricted set of invariants.

2.2. Stress tensors. The strain-energy function W is now taken to depend on the seven deformation-dependent invariants $I_1, I_4, \ldots, I_9$ together (possibly) with $\operatorname{tr} \boldsymbol{T}$ and $\boldsymbol{M} \cdot (\boldsymbol{T}\boldsymbol{M})$. For the considered incompressible material we have

$$\boldsymbol{S} = \frac{\partial W}{\partial \boldsymbol{F}} - p\boldsymbol{F}^{-1} = \sum_{\substack{1 \le i \le 9 \\ i \ne 2,3}} W_i \frac{\partial I_i}{\partial \boldsymbol{F}} - p\boldsymbol{F}^{-1}, \tag{32}$$

and the corresponding Cauchy stress is

$$\boldsymbol{\sigma} = \boldsymbol{F}\frac{\partial W}{\partial \boldsymbol{F}} - p\boldsymbol{I} = \sum_{\substack{1 \le i \le 9 \\ i \ne 2,3}} W_i \boldsymbol{F} \frac{\partial I_i}{\partial \boldsymbol{F}} - p\boldsymbol{I}, \tag{33}$$

where $W_i = \partial W/\partial I_i$, $i = 1, 4, \ldots, 9$. The required expressions for $\partial I_i/\partial \boldsymbol{F}$ are listed for convenience in Appendix A in component form. These enable the Cauchy stress to be expanded as

$$\begin{aligned}
\boldsymbol{\sigma} = {} & -p\boldsymbol{I} + 2W_1 \boldsymbol{B} + 2W_4 \boldsymbol{m} \otimes \boldsymbol{m} + 2W_5(\boldsymbol{m} \otimes \boldsymbol{B}\boldsymbol{m} + \boldsymbol{B}\boldsymbol{m} \otimes \boldsymbol{m}) \\
& + 2W_6 \boldsymbol{\Sigma} + 2W_7(\boldsymbol{B}\boldsymbol{\Sigma} + \boldsymbol{\Sigma}\boldsymbol{B}) + W_8(\boldsymbol{F}\boldsymbol{T}\boldsymbol{M} \otimes \boldsymbol{m} + \boldsymbol{m} \otimes \boldsymbol{F}\boldsymbol{T}\boldsymbol{M}) \\
& + W_9(\boldsymbol{F}\boldsymbol{T}\boldsymbol{M} \otimes \boldsymbol{B}\boldsymbol{m} + \boldsymbol{B}\boldsymbol{m} \otimes \boldsymbol{F}\boldsymbol{T}\boldsymbol{M} + \boldsymbol{F}\boldsymbol{C}\boldsymbol{T}\boldsymbol{M} \otimes \boldsymbol{m} + \boldsymbol{m} \otimes \boldsymbol{F}\boldsymbol{C}\boldsymbol{T}\boldsymbol{M}), \tag{34}
\end{aligned}$$

where $\boldsymbol{m} = \boldsymbol{F}\boldsymbol{M}$, $\boldsymbol{\Sigma} = \boldsymbol{F}\boldsymbol{T}\boldsymbol{F}^{\mathrm{T}}$ and $\boldsymbol{B} = \boldsymbol{F}\boldsymbol{F}^{\mathrm{T}}$ is the left Cauchy–Green deformation tensor.

In the reference configuration $\mathcal{B}_r$ the Cauchy stress must be equal to the initial stress. Thus, when (34) is evaluated in $\mathcal{B}_r$, we obtain

$$\begin{aligned}
\boldsymbol{T} = {} & (2W_1 - p)\boldsymbol{I} + 2(W_4 + 2W_5)\boldsymbol{M} \otimes \boldsymbol{M} + 2(W_6 + 2W_7)\boldsymbol{T} \\
& + (W_8 + 2W_9)(\boldsymbol{T}\boldsymbol{M} \otimes \boldsymbol{M} + \boldsymbol{M} \otimes \boldsymbol{T}\boldsymbol{M}), \tag{35}
\end{aligned}$$

where $W_1, W_4, \ldots, W_9$ are evaluated in the reference configuration, where $\boldsymbol{F} = \boldsymbol{I}$, and may in general depend on $\operatorname{tr} \boldsymbol{T}$ and $\boldsymbol{M} \cdot (\boldsymbol{T}\boldsymbol{M})$. For consistency it is therefore appropriate to set

$$p = 2W_1, \quad W_4 + 2W_5 = 0, \quad W_6 + 2W_7 = \tfrac{1}{2}, \quad W_8 + 2W_9 = 0, \tag{36}$$

in the reference configuration. Indeed, if (35) holds for all possible $\boldsymbol{T}$ then these conditions necessarily follow.

2.3. *The elasticity tensor.* Next, we note that the elasticity tensor $\mathcal{A}$ is given by

$$\mathcal{A} = \frac{\partial^2 W}{\partial \boldsymbol{F} \partial \boldsymbol{F}} = \sum_{\substack{1 \le i \le 9 \\ i \ne 2,3}} W_i \frac{\partial^2 I_i}{\partial \boldsymbol{F} \partial \boldsymbol{F}} + \sum_{\substack{1 \le i \le 9 \\ i \ne 2,3}} \sum_{\substack{1 \le j \le 9 \\ j \ne 2,3}} W_{ij} \frac{\partial I_i}{\partial \boldsymbol{F}} \otimes \frac{\partial I_j}{\partial \boldsymbol{F}}. \tag{37}$$

This requires expressions for the second derivatives of the invariants. In component form these are given in Appendix A. Since the resulting formula for the components of $\mathcal{A}$ is quite long we do not give it here. Instead, we give its specialization to the situation in which there is no finite deformation and $\mathcal{B}_0$ coincides with $\mathcal{B}_r$. The subscript 0 on $\mathcal{A}_0$ may now be omitted, and taking account of the conditions (36) prevailing in the reference configuration the components of $\mathcal{A}$ in the reference configuration can be arranged in the (still fairly lengthy) form

$$
\begin{aligned}
\mathcal{A}_{piqj} = {}& 2W_1 \delta_{ij} \delta_{pq} + 2W_5 (\delta_{ij} M_p M_q + \delta_{pq} M_i M_j + \delta_{iq} M_j M_p + \delta_{jp} M_i M_q) + \delta_{ij} T_{pq} \\
& + 2W_7 (\delta_{ij} T_{pq} + \delta_{pq} T_{ij} + \delta_{iq} T_{jp} + \delta_{jp} T_{iq}) + W_9 [\delta_{ij} (M_p T_{qr} + M_q T_{pr}) \\
& + \delta_{pq} (M_i T_{jr} + M_j T_{ir}) + \delta_{iq} (M_j T_{pr} + M_p T_{jr}) + \delta_{jp} (M_i T_{qr} + M_q T_{ir})] M_r \\
& + 4W_{11} \delta_{ip} \delta_{jq} + 4(W_{44} + 4W_{45} + 4W_{55}) M_i M_j M_p M_q \\
& + 4(W_{14} + 2W_{15})(\delta_{ip} M_j M_q + \delta_{jq} M_i M_p) + 4(W_{16} + 2W_{17})(\delta_{ip} T_{jq} + \delta_{jq} T_{ip}) \\
& + 2(W_{18} + 2W_{19})[\delta_{ip} (M_j T_{qr} + M_q T_{jr}) + \delta_{jq} (M_i T_{pr} + M_p T_{ir})] M_r \\
& + 4(W_{46} + 2W_{47} + 2W_{56} + 4W_{57})(M_i M_p T_{jq} + M_j M_q T_{ip}) \\
& + 2(W_{48} + 2W_{49} + 2W_{58} + 4W_{59})(T_{ir} M_j M_p M_q + T_{jr} M_i M_p M_q \\
& + T_{pr} M_i M_j M_q + T_{qr} M_i M_j M_p) M_r + 4(W_{66} + 4W_{67} + 4W_{77}) T_{ip} T_{jq} \\
& + 2(W_{68} + 2W_{69} + 2W_{78} + 4W_{79})[T_{ip} (T_{qr} M_j + T_{jr} M_q) + T_{jq} (T_{pr} M_i + T_{ir} M_p)] M_r \\
& + (W_{88} + 4W_{89} + 4W_{99})(M_i T_{pr} + M_p T_{ir})(M_j T_{qs} + M_q T_{js}) M_r M_s, \tag{38}
\end{aligned}
$$

all the derivatives of W being evaluated in $\mathcal{B}_r$. With the restricted set of invariants adopted these depend in general on tr $\boldsymbol{T}$ and $\boldsymbol{M} \cdot (\boldsymbol{T} \boldsymbol{M})$. Note that when (38) is substituted into the equation of motion all the terms involving δ_{jp} or δ_{jq} disappear by virtue of the incompressibility condition.

We now consider three special cases that will be used subsequently: (1) the underlying material is isotropic; (2) $\boldsymbol{T} \boldsymbol{M} = \boldsymbol{0}$; (3) $\boldsymbol{T} = T \boldsymbol{M} \otimes \boldsymbol{M}$.

Case 1: Isotropy. If there is no preferred direction in $\mathcal{B}_r$ and in the absence of initial stress the material is isotropic then in the presence of initial stress (38) reduces simply to

$$
\begin{aligned}
\mathcal{A}_{piqj} = {}& 2W_1 \delta_{ij} \delta_{pq} + \delta_{ij} T_{pq} + 2W_7 (\delta_{ij} T_{pq} + \delta_{pq} T_{ij} + \delta_{iq} T_{jp} + \delta_{jp} T_{iq}) + 4W_{11} \delta_{ip} \delta_{jq} \\
& + 4(W_{16} + 2W_{17})(\delta_{ip} T_{jq} + \delta_{jq} T_{ip}) + 4(W_{66} + 4W_{67} + 4W_{77}) T_{ip} T_{jq}; \tag{39}
\end{aligned}
$$

that is, all terms in which W has a subscript 4, 5, 8 or 9 are omitted.

Case 2: $\boldsymbol{T} \boldsymbol{M} = \boldsymbol{0}$. In this case all the terms in (38) in which W has a subscript 8 or 9 vanish and (38) reduces to

$$
\begin{aligned}
\mathcal{A}_{piqj} = {}& 2W_1\delta_{ij}\delta_{pq} + 2W_5(\delta_{ij}M_pM_q + \delta_{pq}M_iM_j + \delta_{iq}M_jM_p + \delta_{jp}M_iM_q) + \delta_{ij}T_{pq} \\
& + 2W_7(\delta_{ij}T_{pq} + \delta_{pq}T_{ij} + \delta_{iq}T_{jp} + \delta_{jp}T_{iq}) + 4W_{11}\delta_{ip}\delta_{jq} \\
& + 4(W_{44} + 4W_{45} + 4W_{55})M_iM_jM_pM_q + 4(W_{14} + 2W_{15})(\delta_{ip}M_jM_q + \delta_{jq}M_iM_p) \\
& + 4(W_{16} + 2W_{17})(\delta_{ip}T_{jq} + \delta_{jq}T_{ip}) + 4(W_{46} + 2W_{47} + 2W_{56} \\
& + 4W_{57})(M_iM_pT_{jq} + M_jM_qT_{ip}) + 4(W_{66} + 4W_{67} + 4W_{77})T_{ip}T_{jq}.
\end{aligned}
$$

We note, in particular, the connections

$$
\mathcal{A}_{ijji} = \mathcal{A}_{jiij} = \mathcal{A}_{ijij} - (2W_1 + T_{ii}) = \mathcal{A}_{jiji} - (2W_1 + T_{jj}), \quad i \neq j, \tag{40}
$$

$$
\mathcal{A}_{jijj} = \mathcal{A}_{ijjj} - T_{ij}, \quad i \neq j, \quad \mathcal{A}_{ijjk} = \mathcal{A}_{ijkj} - T_{ik}, \tag{41}
$$

where i, j, k are distinct.

Case 3: $\boldsymbol{T} = T\boldsymbol{M} \otimes \boldsymbol{M}$. In this case the components of $\mathcal{A}$ have a relatively simple structure and can be written compactly as

$$
\begin{aligned}
\mathcal{A}_{piqj} = {}& 2W_1\delta_{ij}\delta_{pq} + 4W_{11}\delta_{ip}\delta_{jq} + A(\delta_{ij}M_pM_q + \delta_{pq}M_iM_j + \delta_{iq}M_jM_p + \delta_{jp}M_iM_q) \\
& + B(\delta_{ip}M_jM_q + \delta_{jq}M_iM_p) + CM_iM_jM_pM_q + T\delta_{ij}M_pM_q, \quad (42)
\end{aligned}
$$

where

$$
A = 2[W_5 + T(W_7 + W_9)], \tag{43}
$$

$$
B = 4[W_{14} + 2W_{15} + T(W_{16} + 2W_{17}) + T^2(W_{18} + 2W_{19})], \tag{44}
$$

$$
\begin{aligned}
C = {}& 4[W_{44} + 4W_{45} + 4W_{55} + T(W_{46} + 2W_{47} + W_{48} + 2W_{49} + 2W_{56} + 4W_{57} + 2W_{58} + 4W_{59}) \\
& + T^2(W_{66} + 4W_{67} + 4W_{77} + W_{68} + 2W_{69} + 2W_{78} + 4W_{79} + W_{88} + 4W_{89} + 4W_{99})]. \quad (45)
\end{aligned}
$$

It is noteworthy that in each of the three cases the components $\mathcal{A}_{piqj}$ are quadratic in the components of the initial stress even though we have not included in the model invariants that are nonlinear in the initial stress.

3. Homogeneous plane waves

With the focus on incompressible materials we now apply the equation of motion and the incompressibility condition in (22) to the analysis of homogeneous plane waves. In particular, we consider the incremental displacement $\boldsymbol{u}$ and Lagrange multiplier $\dot{p}$ to have the forms

$$
\boldsymbol{u} = f(\boldsymbol{n}\cdot\boldsymbol{x} - vt)\boldsymbol{d}, \quad \dot{p} = g(\boldsymbol{n}\cdot\boldsymbol{x} - vt), \tag{46}
$$

where $\boldsymbol{d}$ is a constant unit (polarization) vector, the unit vector $\boldsymbol{n}$ is the direction of propagation of the plane wave, v is the wave speed, f is a function that need not be made explicit, but is subject to the restriction $f'' \neq 0$, and g is a function related to f. A prime on f or g indicates differentiation with respect to its argument.

Substitution of (46) into (22) then yields

$$
[\boldsymbol{Q}(\boldsymbol{n})\boldsymbol{d} - \rho v^2\boldsymbol{d}]f'' - g'\boldsymbol{n} = \boldsymbol{0}, \quad \boldsymbol{d}\cdot\boldsymbol{n} = 0, \tag{47}
$$

where the (symmetric) *acoustic tensor* $\boldsymbol{Q}(\boldsymbol{n})$ is defined by

$$Q_{ij}(\boldsymbol{n}) = \mathcal{A}_{piqj} n_p n_q. \tag{48}$$

On taking the dot product of $(47)_1$ with $\boldsymbol{n}$ we obtain $g' = [\boldsymbol{Q}(\boldsymbol{n})\boldsymbol{d}] \cdot \boldsymbol{n} f''$, and on substituting back into $(47)_1$ and eliminating the factor f'' we may arrange the resulting *propagation condition* in the form

$$[\boldsymbol{Q}(\boldsymbol{n}) - \boldsymbol{n} \otimes \boldsymbol{Q}(\boldsymbol{n})\boldsymbol{n}]\boldsymbol{d} = \rho v^2 \boldsymbol{d}, \tag{49}$$

or equivalently as

$$\bar{\boldsymbol{Q}}(\boldsymbol{n})\boldsymbol{d} = \rho v^2 \boldsymbol{d}, \tag{50}$$

where $\bar{\boldsymbol{Q}}(\boldsymbol{n})$ is the projection of $\boldsymbol{Q}(\boldsymbol{n})$ on to the plane normal to $\boldsymbol{n}$ defined by

$$\bar{\boldsymbol{Q}}(\boldsymbol{n}) = \bar{\boldsymbol{I}}\boldsymbol{Q}(\boldsymbol{n})\bar{\boldsymbol{I}}, \tag{51}$$

where $\bar{\boldsymbol{I}} = \boldsymbol{I} - \boldsymbol{n} \otimes \boldsymbol{n}$. The symmetrization (51) was originally derived in [Scott and Hayes 1985]. Thus, $\bar{\boldsymbol{Q}}(\boldsymbol{n})$ is symmetric and the two-dimensional eigenvalue problem (50) (in the plane normal to $\boldsymbol{n}$) therefore has two real solutions for ρv^2. The wave speeds are real if the two eigenvalues are positive, and this is guaranteed if the strong ellipticity condition holds. This is expressed as

$$[\bar{\boldsymbol{Q}}(\boldsymbol{n})\boldsymbol{d}] \cdot \boldsymbol{d} > 0 \quad \text{for all nonzero vectors } \boldsymbol{d} \text{ and } \boldsymbol{n} \text{ such that } \boldsymbol{d} \cdot \boldsymbol{n} = 0. \tag{52}$$

For given $\boldsymbol{n}$ and $\boldsymbol{d}$ the wave speed v is obtained from

$$\rho v^2 = [\bar{\boldsymbol{Q}}(\boldsymbol{n})\boldsymbol{d}] \cdot \boldsymbol{d} = [\boldsymbol{Q}(\boldsymbol{n})\boldsymbol{d}] \cdot \boldsymbol{d}. \tag{53}$$

As a first application we consider Case 1 in Section 2.3. Then we obtain

$$\bar{\boldsymbol{Q}}(\boldsymbol{n}) = [2W_1 + (2W_7 + 1)(\boldsymbol{T}\boldsymbol{n}) \cdot \boldsymbol{n}]\bar{\boldsymbol{I}} + 2W_7 \bar{\boldsymbol{T}} + 2(W_{66} + 4W_{67} + 4W_{77})\bar{\boldsymbol{I}}\boldsymbol{T}\boldsymbol{n} \otimes \bar{\boldsymbol{I}}\boldsymbol{T}\boldsymbol{n}, \tag{54}$$

where $\bar{\boldsymbol{T}} = \bar{\boldsymbol{I}}\boldsymbol{T}\bar{\boldsymbol{I}}$ is the projection of $\boldsymbol{T}$ on to the plane normal to $\boldsymbol{n}$. The wave speed is then easily calculated from (53). For illustration we consider the initial stress to be uniaxial and along the x_1 axis with $T_{11} = T$. Then, if we consider motion in the (x_1, x_2) plane with $n_1 = \cos\theta$, $n_2 = \sin\theta$, $d_1 = -\sin\theta$, $d_2 = \cos\theta$, we obtain

$$\rho v^2 = 2W_1 + 2W_7 T + T\cos^2\theta + (W_{66} + 4W_{67} + 4W_{77})T^2 \sin^2 2\theta. \tag{55}$$

Clearly, except for the angles $\theta = 0$ and $\theta = \pi/2$ the squared wave speed depends quadratically (at least) on the initial stress T, bearing in mind that in general the coefficients $W_1, W_7, W_{66}, \ldots$ could depend on T. If all the invariants had been included in W then the nonlinearity would be quartic (at least). This is in contrast to the predictions of Biot's theory, and we now make contact with that for comparison. For this purpose we have provided in Appendix B a derivation of the connection between the general theory used here and Biot's theory. If we denote by $\mathcal{B}_{ijkl}$ the material constants used in Biot's theory then, from Appendix B, we have

$$\mathcal{A}_{ijkl} = \mathcal{B}_{ijkl} - \tfrac{1}{2}\delta_{il}T_{jk} - \tfrac{1}{2}\delta_{ik}T_{jl} - \tfrac{1}{2}\delta_{jk}T_{il} + \tfrac{1}{2}\delta_{jl}T_{ik} + \delta_{kl}T_{ij}. \tag{56}$$

For the present special problem the relevant $\mathcal{B}_{ijkl}$ components are given by

$$\mathcal{B}_{1111} = \mu - T, \quad \mathcal{B}_{2222} = \mu, \quad \mathcal{B}_{1122} = -\mu - T, \quad \mathcal{B}_{2211} = -\mu, \quad \mathcal{B}_{1212} = \mu, \tag{57}$$

where $\mu > 0$ is the shear modulus. It follows that

$$\mathcal{A}_{1111} = \mu - T, \quad \mathcal{A}_{2222} = \mu = -\mathcal{A}_{1122}, \quad \mathcal{A}_{1212} = \mu + \tfrac{1}{2}T, \quad \mathcal{A}_{2121} = \mu - \tfrac{1}{2}T = \mathcal{A}_{1221}. \quad (58)$$

The wave speed in this case is then given by

$$\rho v^2 = \mu - \tfrac{1}{2}T + T\cos^2\theta. \quad (59)$$

Clearly, this result can be recovered from (55) if we take $2W_1 = \mu$, $W_7 = -\tfrac{1}{4}$ (and hence, by $(36)_3$, $W_6 = 1$), and $W_{66} + 4W_{67} + 4W_{77} = 0$. The individual components in (58) are recovered from the general expression (39) by also setting $W_{11} = 0$ and $W_{16} + 2W_{17} = 0$ in (39), except that we obtain $\mathcal{A}_{1122} = 0$ instead of $-\mu$ and $\mathcal{A}_{1221} = -T/2$ instead of $\mu - T/2$. However, this difference is of no consequence since the sum of these two terms is the same in each case and it is only their sum, as the coefficient of $2d_1 d_2 n_1 n_2$, that contributes to the expression for ρv^2 in (53). In fact, because of the incompressibility constraint, there is an intrinsic non-uniqueness in the components $\mathcal{A}_{ijkl}$ (and hence in $\mathcal{B}_{ijkl}$) since a term of the form $p^* \delta_{il} \delta_{jk} + q^* \delta_{ij} \delta_{kl}$ may be added to $\mathcal{A}_{ijkl}$, where p^* and q^* are arbitrary scalars, possibly dependent on tr T and $M \cdot (TM)$ in the general case. The term in q^* disappears from the incremental constitutive relation by incompressibility and the term in p^* vanishes identically in the incremental equation of motion, again by incompressibility.

As a second example we consider Case 3 in Section 2.3, for which we obtain

$$\bar{Q}(n) = [2W_1 + (A + T)(M \cdot n)^2]\bar{I} + [A + C(M \cdot n)^2]\bar{M} \otimes \bar{M}, \quad (60)$$

where $\bar{M} = \bar{I}M$, and hence from (53) we obtain

$$\rho v^2 = 2W_1 + (A + T)(M \cdot n)^2 + [A + C(M \cdot n)^2](M \cdot d)^2. \quad (61)$$

Let us take, for illustration, M to be along the x_1 axis with the same d and n as in the previous example, so the motion is confined to the (x_1, x_2) plane. Then

$$\rho v^2 = 2W_1 + A + T\cos^2\theta + C\sin^2\theta\cos^2\theta, \quad (62)$$

where A and C are given by (43) and (45), respectively. We note, in particular, that C is quadratic in T.

Again we compare with the corresponding formula from Biot's theory, which involves two material constants N and Q for this (two-dimensional orthotropic) situation, for which we obtain

$$\mathcal{A}_{1111} = N - T, \quad \mathcal{A}_{2222} = N = -\mathcal{A}_{1122}, \quad \mathcal{A}_{1212} = Q + \tfrac{1}{2}T, \quad \mathcal{A}_{2121} = Q - \tfrac{1}{2}T = \mathcal{A}_{1221}, \quad (63)$$

and hence

$$\rho v^2 = Q - \tfrac{1}{2}T + T\cos^2\theta + 4(N - Q)\sin^2\theta\cos^2\theta. \quad (64)$$

This can be obtained from (62) by setting

$$2W_1 = N, \quad 2W_1 + 2W_5 = Q, \quad W_7 = -\tfrac{1}{4}, \quad (65)$$

$$C \equiv 4(W_{44} + 4W_{45} + 4W_{55}) = 4(N - Q), \quad (66)$$

and all the other derivatives of W to zero. The case of isotropy is recovered by taking $N = Q = \mu$.

Case 2 produces very similar results when the motion is again restricted to the (x_1, x_2) plane. If we take M now to be aligned along the x_2 axis and T as before then we obtain

$$\rho v^2 = 2W_1 + A' + T\cos^2\theta + C'\sin^2\theta\cos^2\theta. \quad (67)$$

Compared with (62) A' is obtained from A in (43) by omitting the term in W_9, while C' is given by

$$C' = 4[W_{44} + 4W_{45} + 4W_{55} - 2T(W_{46} + 2W_{47} + 2W_{56} + 4W_{57}) + T^2(W_{66} + 4W_{67} + 4W_{77})], \quad (68)$$

which differs from C in (45).

For each of the cases considered above, the appropriate specialization of the strong ellipticity condition (52) requires $\rho v^2 > 0$, and conditions on the material parameters and the initial stress for this to be satisfied for all angles θ may be inferred. For example, from (67), if $T = 0$ then necessary and sufficient conditions for $\rho v^2 > 0$ are $W_1 + W_5 > 0$ and $2W_1 + 2W_5 + W_{44} + 4W_{45} + 4W_{55} > 0$. On the other hand, if $T \neq 0$ then restrictions are imposed on the permissible range of values of T. For example, by considering $\theta = \pi/2$ and $\theta = 0$, we deduce that $2W_1 + A' > 0$ and $2W_1 + A' + T > 0$, respectively, are necessary conditions for $\rho v^2 > 0$, but to obtain necessary and sufficient conditions for $\rho v^2 > 0$ for all θ requires consideration of several possible cases, depending on the signs of the material coefficients W_7, W_{46}, $W_{66}, \ldots$, and since this is algebraically lengthy the details are omitted.

4. Equations governing two-dimensional motions

In this section we restrict attention to two-dimensional motions of incompressible materials. Specifically, we consider motions in the (x_1, x_2) plane with displacement components $\boldsymbol{u} = (u_1, u_2, 0)$, where u_1 and u_2 depend only on x_1, x_2 and t. The two in-plane components of the equations of motion (14) are

$$\mathcal{A}_{p1qj} u_{j,pq} - \dot{p}_{,1} = \rho u_{1,tt}, \quad \mathcal{A}_{p2qj} u_{j,pq} - \dot{p}_{,2} = \rho u_{2,tt}, \quad (69)$$

where the summation is over indices 1 and 2. In general, the third component of the equations of motion should also be considered; here this has the form $\mathcal{A}_{p3qj} u_{j,pq} - \dot{p}_{,3} = 0$. However, henceforth we specialize the form of $\boldsymbol{T}$ and the orientation of $\boldsymbol{M}$ so that this reduces to $\dot{p}_{,3} = 0$, i.e., $\dot{p}$ depends only on x_1, x_2 and t. In particular, we take the (x_1, x_2) plane to be a principal plane of $\boldsymbol{T}$, with $\boldsymbol{M}$ either lying parallel to the (x_1, x_2) plane or aligned with the x_3 direction. Then, it is easy to show that $\mathcal{A}_{p3qj} u_{j,pq}$ vanishes identically.

Elimination of $\dot{p}$ by cross differentiation in (69) yields

$$\mathcal{A}_{p1qj} u_{j,pq2} - \mathcal{A}_{p2qj} u_{j,pq1} = \rho(u_{1,2tt} - u_{2,1tt}). \quad (70)$$

By incompressibility we can introduce a scalar function ψ such that

$$u_1 = \psi_{,2}, \quad u_2 = -\psi_{,1}, \quad (71)$$

and substitution into (70) then yields an equation for ψ, namely

$$\mathcal{A}_{1212}\psi_{,1111} + 2(\mathcal{A}_{1222} - \mathcal{A}_{1211})\psi_{,1112} + (\mathcal{A}_{1111} + \mathcal{A}_{2222} - 2\mathcal{A}_{1122} - 2\mathcal{A}_{1221})\psi_{,1122}$$
$$+ 2(\mathcal{A}_{1121} - \mathcal{A}_{2221})\psi_{,1222} + \mathcal{A}_{2121}\psi_{,2222} = \rho(\psi_{,11tt} + \psi_{,22tt}). \quad (72)$$

We now apply these equations to Cases 2 and 3 specialized to the (x_1, x_2) plane for $\boldsymbol{T} = T\boldsymbol{e}_1 \otimes \boldsymbol{e}_1$, with $\boldsymbol{M} = \boldsymbol{e}_2$ (Case 2) and $\boldsymbol{M} = \boldsymbol{e}_1$ (Case 3).

Case 2: $\boldsymbol{T} = T\boldsymbol{e}_1 \otimes \boldsymbol{e}_1$ and $\boldsymbol{M} = \boldsymbol{e}_2$. In this case (72) reduces to

$$(a + T)\psi_{,1111} + (2a + T + b)\psi_{,1122} + a\psi_{,2222} = \rho(\psi_{,11tt} + \psi_{,22tt}), \quad (73)$$

where

$$a = \mathcal{A}_{1212} - T = \mathcal{A}_{2121} = 2W_1 + A' = 2W_1 + 2W_5 + 2TW_7, \tag{74}$$

$$\begin{aligned} b &= \mathcal{A}_{1111} + \mathcal{A}_{2222} - 2\mathcal{A}_{1122} - 2\mathcal{A}_{1221} - \mathcal{A}_{1212} - \mathcal{A}_{2121} = C' \\ &= 4(W_{44} + 4W_{45} + 4W_{55}) - 8T(W_{46} + 2W_{47} + 2W_{56} + 4W_{57}) + 4T^2(W_{66} + 4W_{67} + 4W_{77}). \end{aligned} \tag{75}$$

Case 3: $\boldsymbol{T} = T\boldsymbol{e}_1 \otimes \boldsymbol{e}_1$ and $\boldsymbol{M} = \boldsymbol{e}_1$. Similarly to the previous case, (72) reduces to

$$(\alpha + T)\psi_{,1111} + (2\alpha + T + \beta)\psi_{,1122} + \alpha\psi_{,2222} = \rho(\psi_{,11tt} + \psi_{,22tt}), \tag{76}$$

where

$$\alpha = \mathcal{A}_{1212} - T = \mathcal{A}_{2121} = 2W_1 + A = 2W_1 + 2W_5 + 2TW_7 + 2TW_9, \tag{77}$$

$$\begin{aligned} \beta &= \mathcal{A}_{1111} + \mathcal{A}_{2222} - 2\mathcal{A}_{1122} - 2\mathcal{A}_{1221} - \mathcal{A}_{1212} - \mathcal{A}_{2121} = C \\ &= 4[W_{44} + 4W_{45} + 4W_{55} + T(W_{46} + 2W_{47} + 2W_{56} + 4W_{57} + W_{48} + 4W_{49} + 2W_{58} + 4W_{59}) \\ &\quad + T^2(W_{66} + 4W_{67} + 4W_{77} + W_{68} + 2W_{69} + 2W_{78} + 4W_{79} + W_{88} + 4W_{89} + 4W_{99})]. \end{aligned} \tag{78}$$

If we compare the coefficients of Case 2 and Case 3, it may be remarked that $\alpha = a + 2TW_9$, while β and b differ significantly, although their structures and the equations are very similar. If there is no initial stress then $T = 0$ and the two sets of equations are the same except that I_4 and I_5 will be different for the two cases because the directions of $\boldsymbol{M}$ are different. If there is initial stress but no preferred direction (underlying isotropy) then all the derivatives of W with a subscript 4, 5, 8 or 9 vanish and the two equations are identical, with coefficients reducing to $a = \alpha = 2W_1 + 2TW_7$ and $b = \beta = 4T^2(W_{66} + 4W_{67} + 4W_{77})$.

5. Rayleigh surface waves

We now consider a half-space occupying the region $x_2 < 0$ in the reference configuration with boundary $x_2 = 0$ and surface waves propagating along the direction x_1. The initial stress and preferred direction are confined to the (x_1, x_2) plane so that the equations of motion (69) are applicable and the third equation is again satisfied trivially. Furthermore, we take the initial stress to have the form $\boldsymbol{T} = T\boldsymbol{e}_1 \otimes \boldsymbol{e}_1$. Then the components $\mathcal{A}_{1112}, \mathcal{A}_{1121}, \mathcal{A}_{2212}, \mathcal{A}_{2221}$ all vanish, as for Cases 2 and 3 in Section 4, and the equations simplify accordingly. On the surface we take the incremental surface traction to vanish, so that $\dot{S}_{021} = 0, \dot{S}_{022} = 0$ on $x_2 = 0$. For the considered incompressible material this yields

$$\mathcal{A}_{2121}u_{1,2} + (\mathcal{A}_{2112} + p)u_{2,1} = 0, \quad \mathcal{A}_{2211}u_{1,1} + (\mathcal{A}_{2222} + p)u_{2,2} - \dot{p} = 0 \quad \text{on } x_2 = 0, \tag{79}$$

and with $u_1 = \psi_{,2}$ and $u_2 = -\psi_{,1}$ these boundary conditions become

$$\mathcal{A}_{2121}\psi_{,22} - (\mathcal{A}_{2112} + p)\psi_{,11} = 0, \quad \mathcal{A}_{2211}\psi_{,12} - (\mathcal{A}_{2222} + p)\psi_{,12} - \dot{p} = 0 \quad \text{on } x_2 = 0. \tag{80}$$

The second of these equations may be expressed in terms of ψ by differentiating with respect to x_1 and then using (69)$_1$, appropriately specialized, to eliminate $\dot{p}_{,1}$. The result is

$$(\mathcal{A}_{1111} + \mathcal{A}_{2222} - 2\mathcal{A}_{1122} - 2\mathcal{A}_{2112} + \mathcal{A}_{2121})\psi_{,112} + \mathcal{A}_{2121}\psi_{,222} = \rho\psi_{,2tt}. \tag{81}$$

The two boundary conditions can now be written compactly as

$$a(\psi_{,22} - \psi_{,11}) = 0, \quad (3a + b + T)\psi_{,112} + a\psi_{,222} = \rho\psi_{,2tt} \quad \text{on } x_2 = 0. \tag{82}$$

These apply for Case 2. The corresponding equations for Case 3 are obtained by replacing a and b by α and β, respectively. Thus, in the following we work with the parameters a and b.

We consider harmonic waves propagating in the x_1 direction and we write ψ in the form

$$\psi(x_1, x_2, t) = \phi(z) \exp[ik(x_1 - vt)], \tag{83}$$

where k is the wave number, v is the wave speed, $z = kx_2$, and the function ϕ is to be determined. Substituting this into the equation of motion (73) we obtain

$$a\phi'''' - (2a + b + T - \rho v^2)\phi'' + (a + T - \rho v^2)\phi = 0, \tag{84}$$

wherein and in the following equations a prime on ϕ denotes differentiation with respect to z. In terms of ϕ the boundary conditions (82) become

$$\phi''(0) + \phi(0) = 0, \quad a\phi'''(0) - (3a + b + T - \rho v^2)\phi'(0) = 0. \tag{85}$$

The factor a has been omitted from the first of these equations on the assumption that $a \neq 0$. For the solution of (84) we require the decay condition $\phi(x_2) \to 0$ as $x_2 \to -\infty$ to hold, and the general solution satisfying this condition is

$$\phi(z) = c_1 \exp(s_1 z) + c_2 \exp(s_2 z), \tag{86}$$

where c_1 and c_2 are constants and s_1, s_2 are the solutions of

$$as^4 - (2a + b + T - \rho v^2)s^2 + (a + T - \rho v^2) = 0 \tag{87}$$

with positive real parts. From the latter it follows that

$$s_1^2 + s_2^2 = \frac{2a + b + T - \rho v^2}{a}, \quad s_1^2 s_2^2 = \frac{a + T - \rho v^2}{a}. \tag{88}$$

If s_1^2 and s_2^2 are real, then they must be positive to ensure that s_1 and s_2 have positive real part. If they are complex then they are conjugate. In either case the product $s_1^2 s_2^2$ must be positive. Assuming that $a > 0$, a real (surface) wave speed v satisfies the inequality

$$0 < \rho v^2 < a + T. \tag{89}$$

Note that if $T \leq 0$ then this cannot be satisfied unless $a > 0$, which requires $W_1 + W_5 + T W_7 > 0$. In Biot's theory, for example, this, in conjunction with (89), yields $-2Q < T < 2Q$.

Substitution of the solution (86) into the boundary conditions (85) yields the equations

$$(1 + s_1^2)c_1 + (1 + s_2^2)c_2 = 0, \tag{90}$$

$$[as_1^3 - (3a + b + T - \rho v^2)s_1]c_1 + [as_2^3 - (3a + b + T - \rho v^2)s_2]c_2 = 0. \tag{91}$$

In order to construct a nontrivial solution of this system, corresponding to vanishing of the determinant of coefficients, it is convenient to introduce the notations

$$\bar{T} = \frac{T}{a}, \quad \xi = \frac{\rho v^2}{a}, \quad \bar{b} = \frac{b}{a}, \quad \eta = \sqrt{1 + \bar{T} - \xi}, \tag{92}$$

and then we obtain the cubic equation

$$\eta^3 + \eta^2 + (3 + \bar{b})\eta - 1 = 0 \tag{93}$$

for η. Once this is solved for (positive) η the wave speed is obtained from the formula

$$\rho v^2 = a\xi = a + T - a\eta^2. \tag{94}$$

The inequalities satisfied by η corresponding to (89) are $0 < \eta < \sqrt{1 + \bar{T}}$.

Note that (93) can be written as

$$(4 + \bar{b} + \bar{T} - \xi)\eta + \bar{T} - \xi = 0. \tag{95}$$

If $T = 0$ then this can only be satisfied for $\xi > 0$ if $\xi < 4 + \bar{b}$, and hence we must have $4 + \bar{b} > 0$ in this case. If $T \neq 0$ then since, by definition, $\eta \geq 0$, (95) can only be satisfied if either

$$\bar{T} < \xi < 4 + \bar{b} + \bar{T}, \tag{96}$$

which requires $4 + \bar{b} > 0$, or

$$4 + \bar{b} + \bar{T} < \xi < \bar{T}, \tag{97}$$

which requires $4 + \bar{b} < 0$, or the transitional case

$$\xi = \bar{T}, \tag{98}$$

which, since then $\eta = 1$, corresponds to $4 + \bar{b} = 0$. Note that (96) can hold for either positive or negative T provided $4 + \bar{b} + \bar{T} > 0$, but (97) and (98) are only possible if $T > 0$.

If we denote the function on the left-hand side of (93) by $f(\eta)$ and note that $f(0) = -1$ and $f''(\eta) > 0$ for $\eta \geq 0$ we see that (93) has a unique positive solution for $\eta \in (0, \sqrt{1 + \bar{T}}]$ if $f(\sqrt{1 + \bar{T}}) \geq 0$. If $T = 0$ then this is guaranteed if $4 + \bar{b} > 0$, and if this inequality holds a unique positive solution also exists for $T > 0$ (subject to $a > 0$). On the other hand, if $T < 0$ then there is a value $\bar{T}_0$ of $\bar{T}$ with $-1 < \bar{T}_0 < 0$ such that a unique positive solution exists only for $\bar{T} > \bar{T}_0$. The value of $\bar{T}_0$ is determined as the value of $\bar{T}$ satisfying $f(\sqrt{1 + \bar{T}}) = 0$, bearing in mind that in general $\bar{b}$ depends on T. The exact details depend in a fairly complicated way on the form of the strain-energy function W and are not discussed here. Instead we illustrate the results using a specific form of W in the following section.

6. Numerical results and discussion

We remain focused on Case 2 in order to compare the results of the general theory with those of the Biot theory for both anisotropic and isotropic models. In particular, we compare the wave speeds for both plane (shear) waves and Rayleigh waves. In each case this involves the initial stress T and material parameters a and b, which also depend on T. For the general constitutive law considered in Case 2 we record here the expressions for a and b from Section 4 for ease of reference:

$$a = 2W_1 + 2W_5 + 2T W_7, \tag{99}$$

$$b = 4(W_{44} + 4W_{45} + 4W_{55}) - 8T(W_{46} + 2W_{47} + 2W_{56} + 4W_{57}) + 4T^2(W_{66} + 4W_{67} + 4W_{77}). \tag{100}$$

If the second derivative terms in W involving a subscript 6 or 7 are identically zero then these reduce to

$$a = 2W_1 + 2W_5 + 2T W_7, \quad b = 4(W_{44} + 4W_{45} + 4W_{55}), \tag{101}$$

which can be specialized to Biot's anisotropic case by setting $W_7 = -\frac{1}{4}$, $2W_1 + 2W_5 = Q$ and $W_{44} + 4W_{45} + 4W_{55} = N - Q$ in terms of Biot's constants N and Q. Finally, the case of isotropy is then recovered by also dropping terms with a subscript 4 or 5, yielding

$$a = 2W_1 + 2TW_7, \quad b = 0, \tag{102}$$

and Biot's case is obtained by setting $2W_1 = \mu$ and $W_7 = -\frac{1}{4}$, where μ is the shear modulus.

For plane waves in the (x_1, x_2) plane the wave speed v is given by

$$\rho v^2 = a + T\cos^2\theta + b\sin^2\theta\cos^2\theta, \tag{103}$$

while the corresponding formula for the Rayleigh wave speed is

$$\rho v^2 = a + T - a\eta^2, \tag{104}$$

where η is the unique positive solution of the equation

$$\eta^3 + \eta^2 + (3 + b/a)\eta - 1 = 0, \quad 0 < \eta < \sqrt{1 + T/a}. \tag{105}$$

In order to capture the effect of the T^2 term in b, which provides the main distinction from the classical theory, it suffices for purposes of illustration to consider a specific form of the strain-energy function W. In the absence of initial stress an incompressible transversely isotropic linear elastic solid is characterized in terms of three elastic constants. Thus, for the transversely isotropic part of W three material constants are needed. The connections between the derivatives of W with respect to the invariants I_1, I_2, I_4 and I_5 evaluated in the undeformed configuration and the three classical constants were provided in [Merodio and Ogden 2005] and can be specialized to the present situation in which I_2 is omitted. We denote the transversely isotropic constants here by μ, ν and κ and consider a transversely isotropic strain-energy function of the form

$$W = \tfrac{1}{2}\mu(I_1 - 3) + \tfrac{1}{2}\nu(I_4 - I_5)^2 - 2\kappa(I_4 - 1) + \kappa(I_5 - 1), \tag{106}$$

which consists of an isotropic neo-Hookean term with constant μ and two anisotropic terms associated with the preferred direction and involving two anisotropic constants ν and κ. The condition $W_4 + 2W_5 = 0$ identified in (36) is then satisfied in the reference configuration $\mathcal{B}_r$, $W_{45} + W_{55} = 0$ and $W_{44} = \nu$.

The initial stress is next incorporated in the model by introducing two additional material constants λ and γ and the invariants I_6 and I_7, and ensuring that the condition $W_6 + 2W_7 = \frac{1}{2}$ in (36) is satisfied in the configuration $\mathcal{B}_r$. A simple example of this inclusion, which we adopt here, extends (106) to the form

$$W = \tfrac{1}{2}\mu(I_1 - 3) + \tfrac{1}{2}\nu(I_4 - I_5)^2 - 2\kappa(I_4 - 1) + \kappa(I_5 - 1) + \tfrac{1}{2}\gamma I_6^2 - \gamma(\operatorname{tr}\boldsymbol{T})I_6$$
$$+ (\tfrac{1}{2} - 2\lambda)I_6 + \lambda I_7 + \tfrac{1}{2}\gamma(\operatorname{tr}\boldsymbol{T})^2 + (\lambda - \tfrac{1}{2})\operatorname{tr}\boldsymbol{T}, \tag{107}$$

the final two terms being included merely to ensure that W vanishes in $\mathcal{B}_r$. They do not contribute to the stress. The only nonzero derivatives of W when evaluated in $\mathcal{B}_r$ are the first derivatives

$$W_1 = \tfrac{1}{2}\mu, \quad W_4 = -2\kappa, \quad W_5 = \kappa, \quad W_6 = \tfrac{1}{2} - 2\lambda, \quad W_7 = \lambda, \tag{108}$$

and the second derivatives

$$W_{44} = \nu, \quad W_{45} = -\nu, \quad W_{55} = \nu, \quad W_{66} = \gamma. \tag{109}$$

The expressions for a and b become

$$a = \mu + 2\kappa + 2\lambda T, \quad b = 4(\nu + \gamma T^2). \tag{110}$$

Clearly, if $\lambda = -\frac{1}{4}$ then the term in T within a has the same form as in the Biot theory. Inclusion of the dimensionless constant λ allows more flexibility in the model.

In what follows we present results in dimensionless form by defining the dimensionless quantities

$$\xi = \rho v^2/\mu, \ \ a^* = a/\mu, \ \ b^* = b/\mu, \ \ \kappa^* = \kappa/\mu, \ \ \nu^* = \nu/\mu, \ \ \gamma^* = \gamma\mu, \ \ T^* = T/\mu, \tag{111}$$

and for consistency of notation for the constants we set $\lambda^* = \lambda$.

Thus, for plane waves,

$$\xi = 1 + 2\kappa^* + 2\lambda^* T^* + T^* \cos^2\theta + 4(\nu^* + \gamma^* T^{*2}) \sin^2\theta \cos^2\theta. \tag{112}$$

In Figure 1 we plot ξ as a function of T^* for the values $0, \pi/6, \pi/4, \pi/3, \pi/2$ of θ and specifically we take $\kappa^* = 0.4, \nu^* = 2.2, \gamma^* = 0.5$. The results are qualitatively very similar for other values of these parameters. The parameter λ^* has been set at $-\frac{1}{4}$. By considering the values $\theta = 0$ and $\theta = \pi/2$ it can be deduced from (112) that T^* must be restricted to the range of values $-2(1 + 2\kappa^*) < T^* < 2(1 + 2\kappa^*)$. Also in Figure 1 we show, for comparison, the results corresponding to $\kappa^* = 0.4, \nu^* = 2.2, \gamma^* = 0$ (the thick dashed lines — essentially the specialization to Biot's anisotropic theory) and to $\kappa^* = 0, \nu^* = 0$,
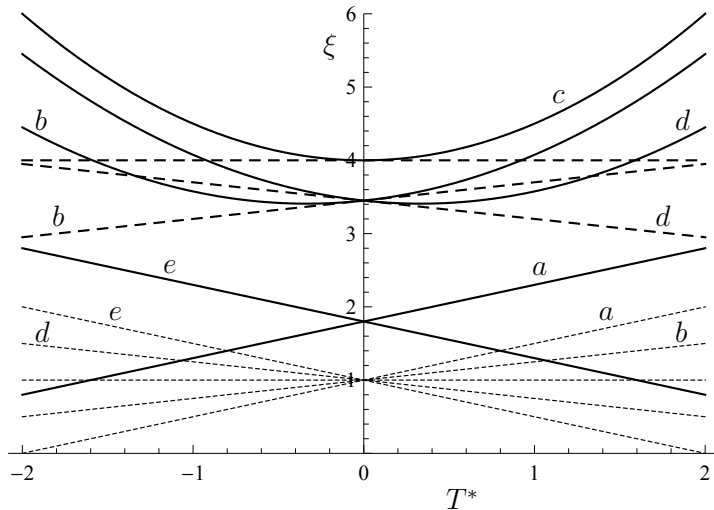
Figure 1. Plot of dimensionless squared wave speed $\xi = \rho v^2/\mu$ against dimensionless initial stress T^* for $\theta = 0$ (a), $\pi/6$ (b), $\pi/4$ (c), $\pi/3$ (d), $\pi/2$ (e), based on (112): continuous curves for parameter values $\kappa^* = 0.4, \nu^* = 2.2, \gamma^* = 0.5$; thick dashed lines for $\kappa^* = 0.4, \nu^* = 2.2, \gamma^* = 0$; thin dashed lines for $\kappa^* = 0, \nu^* = 0, \gamma^* = 0$ (isotropy). In each case $\lambda^* = -\frac{1}{4}$. The horizontal dashed lines correspond to $\theta = \pi/4$.
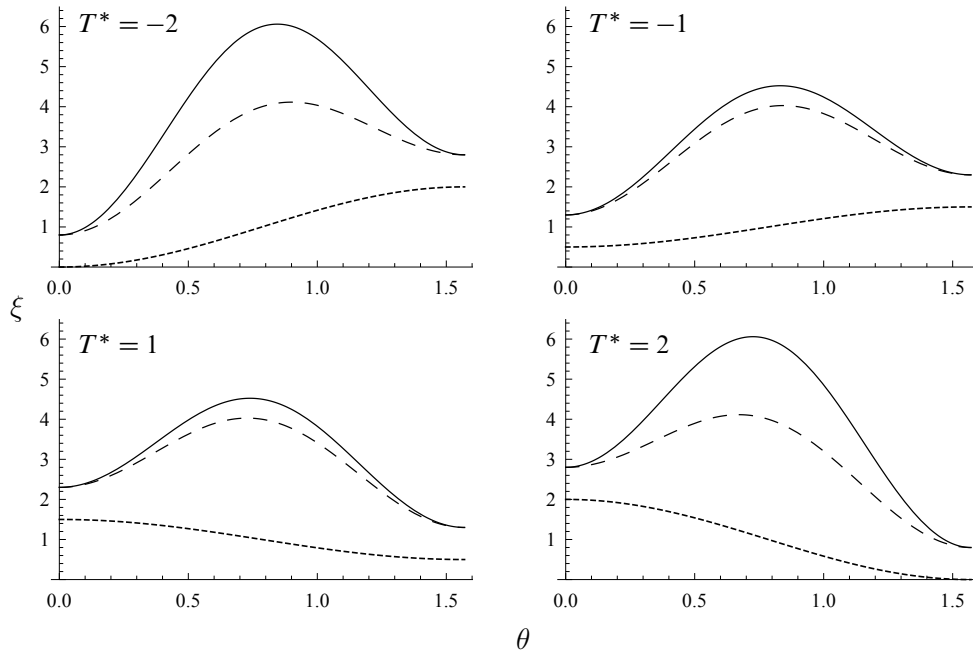
Figure 2. Dimensionless squared wave speed $\xi = \rho v^2/\mu$ versus the angle $\theta \in [0, \pi/2]$, based on (112), for $\lambda^* = -\frac{1}{4}$ and various values of the dimensionless initial stress T^*. Remaining parameter values are $\kappa^* = 0.4$, $\nu^* = 2.2$, $\gamma^* = 0.5$ (solid curves); $\kappa^* = 0.4$, $\nu^* = 2.2$, $\gamma^* = 0$ (thick dashed curves); $\kappa^* = 0$, $\nu^* = 0$, $\gamma^* = 0$ (thin dashed curves).

$\gamma^* = 0$ (the thin dashed lines — Biot's isotropic theory). For $\theta = 0$ and $\theta = \pi/2$ there is no difference between the two anisotropic models. For intermediate values of θ there is a very significant difference between the results for the general model and Biot's anisotropic model due to the term in T^{*2}. Note that in the isotropic case the wave speed vanishes for $T^* = \pm 2$, which correspond to the extreme values of T^* identified previously.

An alternative view of the results is shown in Figure 2, in which ξ is plotted against the angle θ separately for four values of T^*. For each value of T^* the curves corresponding to the three models are shown. Again the significant difference between the general model and Biot's anisotropic model should be noted, which is particularly strong for the larger values of T^*.

The parameter λ^* also has a significant effect and we illustrate this in Figure 3 in which ξ is plotted against T^* for three different values of λ^* and for $\kappa^* = 0.4$, $\nu^* = 2.2$, $\gamma^* = 0.2$ and $\theta = \pi/3$. Even though the (nonlinear) strain-energy function (107) is a considerable specialization of the most general such model it nevertheless demonstrates that the effect of initial stress on the speed of plane waves can be much stronger than is the case with the classical Biot theory.

Next we illustrate the effect of the model (107) on the Rayleigh wave speed based on the solution of (105), with $\xi = \rho v^2/\mu$ then given by (104). In Figure 4, ξ is plotted against T^* for representative values of the parameters ($\kappa^* = 0.4$, $\nu^* = 2.2$, $\gamma^* = 0.5$) and two values of λ^*: -0.15 and -0.3. In each case the dependence on T^* is effectively linear, and this is found also to be the case even when the coefficient γ^* of T^{*2} is quite large. Also shown for comparison are the specializations corresponding to Biot's
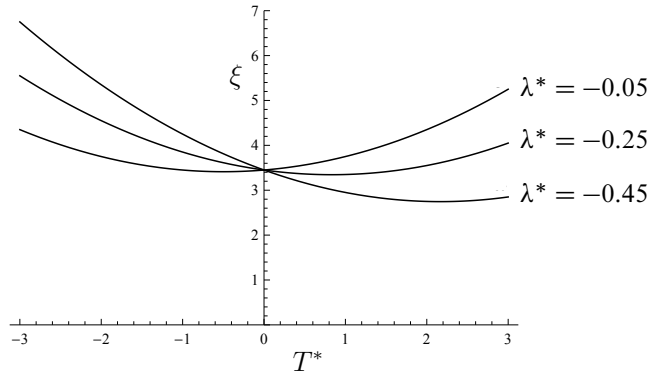
Figure 3. Dimensionless squared wave speed $\xi = \rho v^2/\mu$ versus T^*, based on (112), for $\theta = \pi/3$, $\kappa^* = 0.4$, $\nu^* = 2.2$, $\gamma^* = 0.2$ and different values of λ^*.
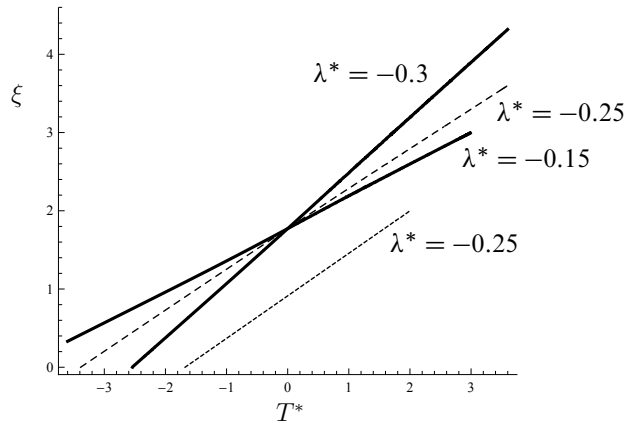
Figure 4. Dimensionless squared Rayleigh wave speed $\xi = \rho v^2/\mu$ versus T^*, based on (104) and (105), for different values of λ^*. Remaining parameter values are $\kappa^* = 0.4$, $\nu^* = 2.2$, $\gamma^* = 0.5$ (solid lines); $\kappa^* = 0.4$, $\nu^* = 2.2$, $\gamma^* = 0$ (dashed line); $\kappa^* = 0$, $\nu^* = 0$, $\gamma^* = 0$ (dotted line).

anisotropic and isotropic results, for which, respectively, the parameter values are taken as $\kappa^* = 0.4$, $\nu^* = 2.2$, $\gamma^* = 0$, $\lambda^* = -0.25$ and $\kappa^* = 0$, $\nu^* = 0$, $\gamma^* = 0$, $\lambda^* = -0.25$.

In the isotropic case we have

$$\xi = 1 + \tfrac{1}{2}T^* - (1 - \tfrac{1}{2}T^*)\eta_0^2, \tag{113}$$

where η_0 is the unique positive root of (105) when $b = 0$ and is given approximately as $\eta_0 = 0.2956$. As noted earlier, a positive real upper limit for η requires that $1 + T^*/a^* > 0$, and for isotropy this gives $-2 < T^* < 2$. The upper limit corresponds to $\xi = 2$, as can be seen in the figure, and this is a cut-off value beyond which a Rayleigh surface wave cannot propagate. At the lower limit $\xi < 0$ and there is also a cut-off value of T^*, approximately -1.679. These limiting values of T^* are associated with the onset of instability of the underlying initially stressed configuration. For a related discussion of
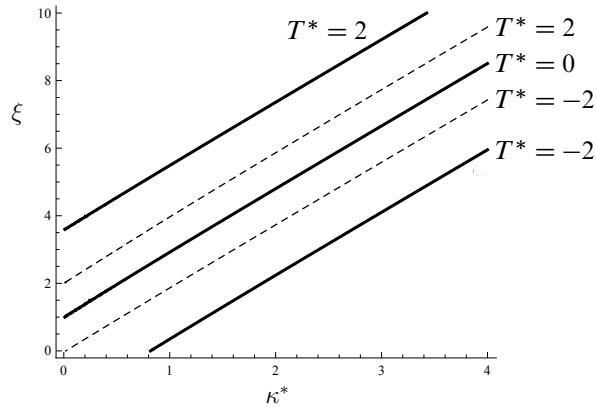
Figure 5. Dimensionless squared Rayleigh wave speed $\xi = \rho v^2/\mu$ versus the parameter κ^* for various values of T^*, with $\nu^* = 2.2$, $\gamma^* = 0$, and $\lambda^* = 0.15$ (solid lines) or $\lambda^* = -0.25$ (dashed lines).

such instability in the case of a pre-stressed (and deformed) half-space we refer to [Dowaikh and Ogden 1990]. Analogous cut-off values of T^* are evident also for the anisotropic models.

As a further illustration of the influence of the parameter λ^*, Figure 5 shows ξ plotted as a function of the parameter κ^* for each of the initial stress values $T^* = -2, 0, 2$ and for the set of parameters $\nu^* = 2.2$, $\gamma^* = 0$, $\lambda^* = 0.15$ compared with results for the set $\nu^* = 2.2$, $\gamma^* = 0$, $\lambda^* = -0.25$. Clearly, λ^* has a significant effect on the value of the wave speed (except, of course, for $T^* = 0$). Thus, as is the case with plane waves, the initial stress dependence of the nonlinear model (107) can have a strong influence on the value of the Rayleigh wave speed compared with its classical linear specialization. Moreover, other specific choices of the nonlinear model within the general framework outlined here can equally and even more substantially affect the wave speed, when, for example, invariants that are nonlinear in the initial stress are included in the form of W.

In this paper we have developed a general theory of transversely isotropic hyperelasticity incorporating initial stress and used a particular specialization of the theory to calculate the elasticity tensor for an undeformed initially stressed configuration. This was then used to examine the propagation of homogeneous plane waves and Rayleigh surface waves with particular reference to the effect of the initial stress. The results, which involve nonlinear terms in the initial stress, are significantly different from those based on the classical theory of Biot, which can be recovered as a special case of the present formulation.

Appendix A. Derivatives of the invariants

First derivatives. The first derivatives of the invariants $I_1, I_4, \ldots, I_9$ with respect to $\boldsymbol{F}$ are obtained, in component form, as

$$\frac{\partial I_1}{\partial F_{i\alpha}} = 2F_{i\alpha}, \quad \frac{\partial I_4}{\partial F_{i\alpha}} = 2M_\alpha F_{i\beta} M_\beta, \quad \frac{\partial I_5}{\partial F_{i\alpha}} = 2(M_\alpha F_{i\beta} C_{\beta\gamma} + C_{\alpha\beta} M_\beta F_{i\gamma}) M_\gamma,$$

$$\frac{\partial I_6}{\partial F_{i\alpha}} = 2T_{\alpha\beta} F_{i\beta}, \quad \frac{\partial I_7}{\partial F_{i\alpha}} = 2(T_{\alpha\beta} C_{\beta\gamma} + C_{\alpha\beta} T_{\beta\gamma}) F_{i\gamma},$$

$$\frac{\partial I_8}{\partial F_{i\alpha}} = (T_{\alpha\beta} M_\beta F_{i\gamma} + M_\alpha F_{i\beta} T_{\beta\gamma}) M_\gamma,$$

$$\frac{\partial I_9}{\partial F_{i\alpha}} = (T_{\alpha\beta} M_\beta F_{i\gamma} C_{\gamma\delta} + C_{\alpha\beta} M_\beta F_{i\gamma} T_{\gamma\delta} + M_\alpha F_{i\beta} C_{\beta\gamma} T_{\gamma\delta} + F_{i\beta} M_\beta C_{\alpha\gamma} T_{\gamma\delta}) M_\delta.$$

When evaluated in the reference configuration these reduce to

$$\frac{\partial I_1}{\partial F_{i\alpha}} = 2\delta_{i\alpha}, \quad 2\frac{\partial I_4}{\partial F_{i\alpha}} = 4M_i M_\alpha = \frac{\partial I_5}{\partial F_{i\alpha}}, \quad 2\frac{\partial I_6}{\partial F_{i\alpha}} = 4T_{i\alpha} = \frac{\partial I_7}{\partial F_{i\alpha}},$$

$$2\frac{\partial I_8}{\partial F_{i\alpha}} = 2(M_i T_{\alpha\beta} + M_\alpha T_{i\beta}) M_\beta = \frac{\partial I_9}{\partial F_{i\alpha}}.$$

Since the latter apply in the reference configuration there is strictly no distinction between Greek and Roman indices in this case.

Second derivatives. The second derivatives of the invariants $I_1, I_4, \ldots, I_9$ with respect to $\boldsymbol{F}$ are, in component form,

$$\frac{\partial^2 I_1}{\partial F_{i\alpha} \partial F_{j\beta}} = 2\delta_{ij}\delta_{\alpha\beta}, \quad \frac{\partial^2 I_4}{\partial F_{i\alpha} \partial F_{j\beta}} = 2\delta_{ij} M_\alpha M_\beta, \quad \frac{\partial^2 I_6}{\partial F_{i\alpha} \partial F_{j\beta}} = 2\delta_{ij} T_{\alpha\beta},$$

$$\frac{\partial^2 I_5}{\partial F_{i\alpha} \partial F_{j\beta}} = 2\delta_{ij}(M_\alpha C_{\beta\gamma} + M_\beta C_{\alpha\gamma}) M_\gamma$$
$$+ 2B_{ij} M_\alpha M_\beta + 2\delta_{\alpha\beta} F_{i\gamma} M_\gamma F_{j\delta} M_\delta + 2(F_{i\beta} F_{j\gamma} M_\alpha + F_{j\alpha} F_{i\gamma} M_\beta) M_\gamma,$$

$$\frac{\partial^2 I_7}{\partial F_{i\alpha} \partial F_{j\beta}} = 2[\delta_{ij}(T_{\alpha\gamma} C_{\beta\gamma} + T_{\beta\gamma} C_{\alpha\gamma}) + B_{ij} T_{\alpha\beta} + \delta_{\alpha\beta} \Sigma_{ij} + F_{i\beta} F_{j\gamma} T_{\alpha\gamma} + F_{j\alpha} F_{i\gamma} T_{\beta\gamma}],$$

$$\frac{\partial^2 I_8}{\partial F_{i\alpha} \partial F_{j\beta}} = \delta_{ij}(M_\alpha T_{\beta\gamma} + M_\beta T_{\alpha\gamma}) M_\gamma,$$

$$\frac{\partial^2 I_9}{\partial F_{i\alpha} \partial F_{j\beta}} = \delta_{ij}(T_{\alpha\gamma} M_\gamma C_{\beta\delta} + T_{\beta\gamma} M_\gamma C_{\alpha\delta} + C_{\alpha\gamma} T_{\gamma\delta} M_\beta + C_{\beta\gamma} T_{\gamma\delta} M_\alpha) M_\delta$$
$$+ \delta_{\alpha\beta}(F_{i\gamma} T_{\gamma\delta} F_{j\varepsilon} + F_{j\gamma} T_{\gamma\delta} F_{i\varepsilon}) M_\delta M_\varepsilon + B_{ij}(M_\alpha T_{\beta\gamma} + T_{\alpha\gamma} M_\beta) M_\gamma$$
$$+ (M_\alpha F_{i\beta} F_{j\gamma} + M_\beta F_{j\alpha} F_{i\gamma}) T_{\gamma\delta} M_\delta + (T_{\alpha\gamma} F_{i\beta} F_{j\delta} + T_{\beta\gamma} F_{j\alpha} F_{i\delta}) M_\gamma M_\delta.$$

When evaluated in the reference configuration these specialize to

$$\frac{\partial^2 I_1}{\partial F_{i\alpha} \partial F_{j\beta}} = 2\delta_{ij}\delta_{\alpha\beta}, \quad \frac{\partial^2 I_4}{\partial F_{i\alpha} \partial F_{j\beta}} = 2\delta_{ij} M_\alpha M_\beta, \quad \frac{\partial^2 I_6}{\partial F_{i\alpha} \partial F_{j\beta}} = 2\delta_{ij} T_{\alpha\beta},$$

$$\frac{\partial^2 I_5}{\partial F_{i\alpha} \partial F_{j\beta}} = 6\delta_{ij} M_\alpha M_\beta + 2(\delta_{\alpha\beta} M_i M_j + \delta_{i\beta} M_j M_\alpha + \delta_{j\alpha} M_i M_\beta),$$

$$\frac{\partial^2 I_7}{\partial F_{i\alpha} \partial F_{j\beta}} = 6\delta_{ij} T_{\alpha\beta} + 2(\delta_{\alpha\beta} T_{ij} + \delta_{i\beta} T_{j\alpha} + \delta_{j\alpha} T_{i\beta}),$$

$$\frac{\partial^2 I_8}{\partial F_{i\alpha} \partial F_{j\beta}} = \delta_{ij}(M_\alpha T_{\beta\gamma} + M_\beta T_{\alpha\gamma}) M_\gamma,$$

$$\frac{\partial^2 I_9}{\partial F_{i\alpha} \partial F_{j\beta}} = 3\delta_{ij}(M_\alpha T_{\beta\gamma} + M_\beta T_{\alpha\gamma})M_\gamma + \delta_{\alpha\beta}(M_i T_{j\gamma} + M_j T_{i\gamma})M_\gamma$$

$$+ \delta_{i\beta}(M_\alpha T_{j\gamma} + M_j T_{\alpha\gamma})M_\gamma + \delta_{j\alpha}(M_i T_{\beta\gamma} + M_\beta T_{i\gamma})M_\gamma.$$

Appendix B. Connections with Biot's equations

It is instructive to relate the formulation of the incremental equations adopted here to that of Biot (see, for example, [Biot 1965] for details). For a compressible material, which we deal with here, the nominal stress tensor $\boldsymbol{S}$ and Cauchy stress tensor $\boldsymbol{\sigma}$ are related by $\boldsymbol{S} = J\boldsymbol{F}^{-1}\boldsymbol{\sigma}$, where $\boldsymbol{F}$ is deformation gradient and $J = \det \boldsymbol{F}$. On taking the increment of this equation we obtain

$$\dot{\boldsymbol{S}} = J\boldsymbol{F}^{-1}\dot{\boldsymbol{\sigma}} + \dot{J}\boldsymbol{F}^{-1}\boldsymbol{\sigma} + J\overline{\dot{\boldsymbol{F}^{-1}}}\boldsymbol{\sigma} = J\boldsymbol{F}^{-1}[\dot{\boldsymbol{\sigma}} + (\operatorname{tr}\boldsymbol{L})\boldsymbol{\sigma} - \boldsymbol{L}\boldsymbol{\sigma}],$$

where $\boldsymbol{L} = \operatorname{grad}\boldsymbol{u}$ and $\boldsymbol{u}$ is the displacement vector. On updating the reference configuration from $\mathcal{B}_r$ to $\mathcal{B}_0$ we obtain

$$\dot{\boldsymbol{S}}_0 = \dot{\boldsymbol{\sigma}} + (\operatorname{tr}\boldsymbol{L})\boldsymbol{\sigma} - \boldsymbol{L}\boldsymbol{\sigma}.$$

Taking the divergence of both sides leads to

$$\operatorname{div}\dot{\boldsymbol{S}}_0 = \operatorname{div}\dot{\boldsymbol{\sigma}} + \operatorname{div}[(\operatorname{tr}\boldsymbol{L})\boldsymbol{\sigma} - \boldsymbol{L}\boldsymbol{\sigma}].$$

If the configuration $\mathcal{B}_0$ is uniform then we have

$$\operatorname{div}[(\operatorname{tr}\boldsymbol{L})\boldsymbol{\sigma} - \boldsymbol{L}\boldsymbol{\sigma}] = L_{kk,j}\sigma_{ji} - L_{jk,j}\sigma_{ki} = L_{kk,j}\sigma_{ji} - L_{jj,k}\sigma_{ki} = 0,$$

and the incremental equation of motion becomes

$$\operatorname{div}\dot{\boldsymbol{S}}_0 \equiv \operatorname{div}\dot{\boldsymbol{\sigma}} = \rho\boldsymbol{u}_{,tt}.$$

Now, Biot works in terms of components referred to different sets of axes, in particular a set of Cartesian axes and a set of axes obtained by rotation therefrom, with the rotation associated with the incremental deformation. Let $\boldsymbol{e}_i$, $i = 1, 2, 3$, be a Cartesian coordinate basis in $\mathcal{B}_0$, and let $\boldsymbol{e}_i'$, $i = 1, 2, 3$, be the axes obtained by the rotation. Then, $\boldsymbol{e}_i' \equiv \boldsymbol{e}_i + \boldsymbol{W}\boldsymbol{e}_i$, where the tensor $\boldsymbol{W}$ is the antisymmetric part of the displacement gradient $\boldsymbol{L}$. After the increment the total Cauchy stress is $\boldsymbol{\sigma} + \dot{\boldsymbol{\sigma}}$, which has components on the rotated axes given by, to the first order in incremental quantities,

$$[(\boldsymbol{\sigma} + \dot{\boldsymbol{\sigma}})\boldsymbol{e}_i'] \cdot \boldsymbol{e}_j' = [(\boldsymbol{\sigma} + \dot{\boldsymbol{\sigma}})(\boldsymbol{e}_i + \boldsymbol{W}\boldsymbol{e}_i)] \cdot (\boldsymbol{e}_j + \boldsymbol{W}\boldsymbol{e}_j) = \sigma_{ij} + \dot{\sigma}_{ij} + \sigma_{ik}W_{kj} - W_{ik}\sigma_{kj},$$

where the symmetry of $\boldsymbol{\sigma}$ and the antisymmetry of $\boldsymbol{W}$ have been used, and σ_{ij} and W_{ij} are their components referred to the original axes (note that W_{ij} here are not the same as the second derivatives of the strain-energy function with respect to the invariants introduced in (29)). The component form of the incremental equation of motion with uniform $\boldsymbol{\sigma}$ therefore becomes

$$\dot{\sigma}_{ji,j} + \sigma_{ik}W_{kj,j} - \sigma_{jk}W_{ik,j} = \rho u_{i,tt}.$$

This is translated into Biot's notation [1965] by setting $\sigma_{ij} = S_{ij}$, $\dot{\sigma}_{ij} = s_{ij}$ and $W_{ij} = -\omega_{ij}$, and then yields Biot's equation

$$s_{ij,j} + S_{jk}\omega_{ik,j} + S_{ik}\omega_{jk,j} = \rho u_{i,tt}$$

appropriate for homogeneous S_{ij} in the absence of body forces [Biot 1965, p. 264]. Biot's constitutive equation for incremental deformations may be written in the form $\dot{\sigma}_{ij} = \mathcal{B}_{ijkl} L_{kl}$, where $L_{kl} = u_{k,l}$ and

$$\mathcal{B}_{ijkl} = \mathcal{B}_{jikl} = \mathcal{B}_{ijlk}.$$

Since $W_{ij} = (L_{ij} - L_{ji})/2$, it follows that

$$\mathcal{A}_{ijkl} L_{lk} = \mathcal{B}_{ijkl} L_{lk} - \tfrac{1}{2}(L_{ik} + L_{ki})\sigma_{kj} - \tfrac{1}{2}\sigma_{ik}(L_{kj} - L_{jk}) + L_{kk}\sigma_{ij},$$

and since this holds for arbitrary L_{ij} we deduce that

$$\mathcal{A}_{ijkl} = \mathcal{B}_{ijkl} - \tfrac{1}{2}\delta_{il}\sigma_{jk} - \tfrac{1}{2}\delta_{ik}\sigma_{jl} - \tfrac{1}{2}\delta_{jk}\sigma_{il} + \tfrac{1}{2}\delta_{jl}\sigma_{ik} + \delta_{kl}\sigma_{ij}.$$

This identity holds in the configuration $\mathcal{B}_0$ and when $\mathcal{B}_0$ is taken to coincide with $\mathcal{B}_r$ the Cauchy stress $\boldsymbol{\sigma}$ becomes the initial stress $\boldsymbol{T}$. Using the symmetry $\mathcal{A}_{ijkl} = \mathcal{A}_{klij}$, it follows that

$$\mathcal{B}_{ijkl} - \mathcal{B}_{klij} = \delta_{ij}\sigma_{kl} - \delta_{kl}\sigma_{ij},$$

which recovers a formula of Biot [1965, p. 71], albeit in different notation.

References

[Adkins 1960] J. E. Adkins, "Symmetry relations for orthotropic and transversely isotropic materials", *Arch. Ration. Mech. An.* **4**:1 (1960), 193–213.

[Akbarov and Guz 2004] S. D. Akbarov and A. N. Guz, "Axisymmetric longitudinal wave propagation in pre-stressed compound circular cylinders", *Int. J. Eng. Sci.* **42**:8-9 (2004), 769–791.

[Anderson 1961] D. L. Anderson, "Elastic wave propagation in layered anisotropic media", *J. Geophys. Res.* **66**:9 (1961), 2953–2963.

[Barnett and Lothe 1985] D. M. Barnett and J. Lothe, "Free surface (Rayleigh) waves in anisotropic elastic half-spaces: the surface impedance method", *Proc. R. Soc. Lond. A* **402**:1822 (1985), 135–152.

[Biot 1939] M. A. Biot, "Non-linear theory of elasticity and the linearized case for a body under initial stress", *Philos. Mag.* **27**:183 (1939), 468–489.

[Biot 1940] M. A. Biot, "The influence of initial stress on elastic waves", *J. Appl. Phys.* **11**:8 (1940), 522–530.

[Biot 1965] M. A. Biot, *Mechanics of incremental deformations*, Wiley, New York, 1965.

[Chadwick 1997] P. Chadwick, "The application of the Stroh formalism to prestressed elastic media", *Math. Mech. Solids* **2**:4 (1997), 379–403.

[Chadwick and Jarvis 1979] P. Chadwick and D. A. Jarvis, "Surface waves in a pre-stressed elastic body", *Proc. R. Soc. Lond. A* **366**:1727 (1979), 517–536.

[Chadwick and Smith 1977] P. Chadwick and G. D. Smith, "Foundations of the theory of surface waves in anisotropic elastic materials", *Adv. Appl. Mech.* **17** (1977), 303–376.

[Coleman and Noll 1964] B. D. Coleman and W. Noll, "Material symmetry and thermostatic inequalities in finite elastic deformations", *Arch. Ration. Mech. An.* **15**:2 (1964), 87–111.

[Destrade 2001a] M. Destrade, "The explicit secular equation for surface acoustic waves in monoclinic elastic crystals", *J. Acoust. Soc. Am.* **109**:4 (2001), 1398–1402.

[Destrade 2001b] M. Destrade, "Surface waves in orthotropic incompressible materials", *J. Acoust. Soc. Am.* **110**:2 (2001), 837–840.

[Destrade 2003] M. Destrade, "Rayleigh waves in symmetry planes of crystals: explicit secular equations and some explicit wave speeds", *Mech. Mater.* **35**:9 (2003), 931–939.

[Destrade et al. 2002] M. Destrade, P. A. Martin, and T. C. T. Ting, "The incompressible limit in linear anisotropic elasticity, with applications to surface waves and elastostatics", *J. Mech. Phys. Solids* **50**:7 (2002), 1453–1468.

[Destrade et al. 2005] M. Destrade, M. Ottenio, A. V. Pichugin, and G. A. Rogerson, "Non-principal surface waves in deformed incompressible materials", *Int. J. Eng. Sci.* **43**:13-14 (2005), 1092–1106.

[Dey and De 1999] S. Dey and P. K. De, "Velocity of shear waves in an initially stressed incompressible anisotropic medium", *Sadhana* **24**:3 (1999), 215–220.

[Dowaikh and Ogden 1990] M. A. Dowaikh and R. W. Ogden, "On surface waves and deformations in a pre-stressed incompressible elastic solid", *IMA J. Appl. Math.* **44**:3 (1990), 261–284.

[Dowaikh and Ogden 1991] M. A. Dowaikh and R. W. Ogden, "On surface waves and deformations in a compressible elastic half-space", *Stab. Appl. Anal. Cont. Media* **1**:1 (1991), 27–45.

[Edmondson and Fu 2009] R. T. Edmondson and Y. B. Fu, "Stroh formulation for a generally constrained and pre-stressed elastic material", *Int. J. Non-Linear Mech.* **44**:5 (2009), 530–537.

[Flavin 1963] J. N. Flavin, "Surface waves in pre-stressed Mooney material", *Q. J. Mech. Appl. Math.* **16**:4 (1963), 441–449.

[Guz 2002] A. N. Guz, "Elastic waves in bodies with initial (residual) stresses", *Int. Appl. Mech.* **38**:1 (2002), 23–59.

[Hayes and Rivlin 1961] M. Hayes and R. S. Rivlin, "Surface waves in deformed elastic materials", *Arch. Ration. Mech. An.* **8**:1 (1961), 358–380.

[Hoger 1985] A. Hoger, "On the residual stress possible in an elastic body with material symmetry", *Arch. Ration. Mech. An.* **88**:3 (1985), 271–289.

[Hoger 1986] A. Hoger, "On the determination of residual stress in an elastic body", *J. Elasticity* **16**:3 (1986), 303–324.

[Hoger 1993a] A. Hoger, "The constitutive equation for finite deformations of a transversely isotropic hyperelastic material with residual stress", *J. Elasticity* **33**:2 (1993), 107–118.

[Hoger 1993b] A. Hoger, "The elasticity tensors of a residually stressed material", *J. Elasticity* **31**:3 (1993), 219–237.

[Hoger 1996] A. Hoger, "The elasticity tensor of a transversely isotropic hyperelastic material with residual stress", *J. Elasticity* **42**:2 (1996), 115–132.

[Ieşan 1989] D. Ieşan, *Prestressed bodies*, Pitman research notes in mathematics **195**, Longman Scientific, Harlow, 1989.

[Johnson and Hoger 1993] B. E. Johnson and A. Hoger, "The dependence of the elasticity tensor on residual stress", *J. Elasticity* **33**:2 (1993), 145–165.

[Man 1998] C.-S. Man, "Hartig's law and linear elasticity with initial stress", *Inverse Problems* **14**:2 (1998), 313–319.

[Man and Lu 1987] C.-S. Man and W. Y. Lu, "Towards an acoustoelastic theory for measurement of residual stress", *J. Elasticity* **17**:2 (1987), 159–182.

[Merodio and Ogden 2002] J. Merodio and R. W. Ogden, "Material instabilities in fiber-reinforced nonlinearly elastic solids under plane deformation", *Arch. Mech.* **54** (2002), 525–552.

[Merodio and Ogden 2003] J. Merodio and R. W. Ogden, "Instabilities and loss of ellipticity in fiber-reinforced compressible non-linearly elastic solids under plane deformation", *Int. J. Solids Struct.* **40**:18 (2003), 4707–4727.

[Merodio and Ogden 2005] J. Merodio and R. W. Ogden, "Mechanical response of fiber-reinforced incompressible non-linearly elastic solids", *Int. J. Non-Linear Mech.* **40**:2-3 (2005), 213–227.

[Mozhaev 1995] V. G. Mozhaev, "Some new ideas in the theory of surface acoustic waves in anisotropic media", pp. 455–462 in *IUTAM Symposium on Anisotropy, Inhomogeneity and Nonlinearity in Solid Mechanics: proceedings of the IUTAM-ISIMM symposium* (Nottingham, UK, 1994), edited by D. F. Parker and A. H. England, Kluwer, Dordrecht, 1995.

[Musgrave 1959] M. J. P. Musgrave, "The propagation of elastic waves in crystals and other anisotropic media", *Rept. Progr. Phys.* **22** (1959), 74–96.

[Nair and Sotiropoulos 1997] S. Nair and D. A. Sotiropoulos, "Elastic waves in orthotropic incompressible materials and reflection from an interface", *J. Acoust. Soc. Am.* **102**:1 (1997), 102–109.

[Nair and Sotiropoulos 1999] S. Nair and D. A. Sotiropoulos, "Interfacial waves in incompressible monoclinic materials with an interlayer", *Mech. Mater.* **31**:3 (1999), 225–233.

[Norris 1983] A. N. Norris, "Propagation of plane waves in a pre-stressed elastic medium", *J. Acoust. Soc. Am.* **74**:5 (1983), 1642–1643.

[Norris and Sinha 1995] A. N. Norris and B. Sinha, "The speed of a wave along a fluid/solid interface in the presence of anisotropy and prestress", *J. Acoust. Soc. Am.* **98**:2 (1995), 1147–1154.

[Ogden 1984] R. W. Ogden, *Non-linear elastic deformations*, Ellis Horwood, Chichester, 1984.

[Ogden 2003] R. W. Ogden, "Nonlinear elasticity, anisotropy and residual stresses in soft tissue", pp. 65–108 in *Biomechanics of soft tissue in cardiovascular systems*, edited by G. A. Holzapfel and R. W. Ogden, CISM courses and lectures, 441, Springer, Wien, 2003.

[Ogden 2007] R. W. Ogden, "Incremental statics and dynamics of pre-stressed elastic materials", pp. 1–26 in *Waves in nonlinear pre-stressed materials*, edited by M. Destrade and G. Saccomandi, CISM courses and lectures, 495, Springer, Wien, 2007.

[Ogden and Vinh 2004] R. W. Ogden and P. C. Vinh, "On Rayleigh waves in incompressible orthotropic elastic solids", *J. Acoust. Soc. Am.* **115**:2 (2004), 530–533.

[Prikazchikov and Rogerson 2004] D. A. Prikazchikov and G. A. Rogerson, "On surface wave propagation in incompressible, transversely isotropic, pre-stressed elastic half-spaces", *Int. J. Eng. Sci.* **42**:10 (2004), 967–986.

[Rayleigh 1885] L. Rayleigh, "On waves propagated along the plane surface of an elastic solid", *Proc. Lond. Math. Soc.* **17** (1885), 4–11.

[Royer and Dieulesaint 1984] D. Royer and E. Dieulesaint, "Rayleigh wave velocity and displacement in orthorhombic, tetragonal, hexagonal and cubic crystals", *J. Acoust. Soc. Am.* **76**:5 (1984), 1438–1444.

[Saravanan 2008] U. Saravanan, "Representation for stress from a stressed reference configuration", *Int. J. Eng. Sci.* **46**:11 (2008), 1063–1076.

[Scott and Hayes 1985] N. H. Scott and M. Hayes, "A note on wave propagation in internally constrained hyperelastic materials", *Wave Motion* **7**:6 (1985), 601–605.

[Sharma and Garg 2006] M. D. Sharma and N. Garg, "Wave velocities in a pre-stressed anisotropic elastic medium", *J. Earth Syst. Sci.* **115**:2 (2006), 257–265.

[Song and Fu 2007] Y. Q. Song and Y. B. Fu, "A note on perturbation formulae for the surface-wave speed due to perturbations in material properties", *J. Elasticity* **88**:3 (2007), 187–192.

[Spencer 1971] A. J. M. Spencer, *Theory of invariants*, vol. 1, edited by A. C. Eringen, Academic Press, New York, 1971.

[Stoneley 1963] R. Stoneley, "The propagation of surface waves in an elastic medium with orthorhombic symmetry", *Geophys. J. R. Astron. Soc.* **8**:2 (1963), 176–186.

[Tanuma and Man 2008] K. Tanuma and C.-S. Man, "Perturbation formulas for polarization ratio and phase shift of Rayleigh waves in prestressed anisotropic media", *J. Elasticity* **92**:1 (2008), 1–33.

[Ting 2002a] T. C. T. Ting, "An explicit secular equation for surface waves in an elastic material of general anisotropy", *Q. J. Mech. Appl. Math.* **55**:2 (2002), 297–311.

[Ting 2002b] T. C. T. Ting, "Explicit secular equations for surface waves in monoclinic materials with the symmetry plane $x_1 = 0$, $x_2 = 0$ or $x_3 = 0$", *Proc. R. Soc. Lond. A* **458**:2021 (2002), 1017–1031.

[Ting 2002c] T. C. T. Ting, "A unified formalism for elastostatics or steady state motion of compressible or incompressible anisotropic elastic materials", *Int. J. Solids Struct.* **39**:21-22 (2002), 5427–5445.

[Tolstoy 1982] I. Tolstoy, "On elastic waves in prestressed solids", *J. Geophys. Res.* **87** (1982), 6823–6827.

[Zheng 1994] Q.-S. Zheng, "Theory of representations for tensor functions: a unified invariant approach to constitutive equations", *Appl. Mech. Rev.* **47**:11 (1994), 545–587.

RAY W. OGDEN: r.ogden@abdn.ac.uk
Department of Mathematics, University of Glasgow, University Gardens, Glasgow, G12 8QW, United Kingdom
Current address: *School of Engineering, University of Aberdeen, King's College, Aberdeen, AB24 3UE, United Kingdom*

BALJEET SINGH: dr_baljeet@hotmail.com
Department of Mathematics, Postgraduate Government College, Sector 11, Chandigarh 160011, India

∎
∎∎msp

VARIABLE-ORDER FINITE ELEMENTS FOR NONLINEAR, FULLY INTRINSIC BEAM EQUATIONS

MAYURESH J. PATIL AND DEWEY H. HODGES

Fully intrinsic equations and boundary conditions involve only force, moment, velocity, and angular velocity variables, but no displacement or rotation variables. This paper presents variable-order finite elements for the geometrically exact, nonlinear, fully intrinsic equations for both nonrotating and rotating beams. The finite element technique allows for *hp*-adaptivity. Results show that these finite elements lead to very accurate solutions for the static equilibrium state as well as for modes and frequencies for infinitesimal motions about that state. For the same number of variables, the accuracy of the finite elements increases with the order of the finite element. The results based on the Galerkin approximation (which is a special case of the present approach) are the most accurate but require evaluation of complex integrals. Cubic elements are shown to provide a near optimal combination of accuracy and complexity.

1. Introduction

Beam-like structures are used in a wide range of applications, from rotor blades and high-aspect-ratio wings to nanosensors. The displacement formulation of the geometrically exact equations of motion for beams was presented in [Borri and Mantegazza 1985; Simo and Vu-Quoc 1988; Bauchau and Kang 1993]. A mixed formulation was presented in [Hodges 1990], which gave the partial differential equations of motion in so-called intrinsic form, following the terminology of [Green and Laws 1966; Reissner 1973] for beams and [Danielson 1970] for shells. Intrinsic equations are independent of the way displacement and rotation are parametrized. In the mixed formulation of [Hodges 1990], the kinematical partial differential equations relate the chosen displacement and rotation variables to the generalized strain and velocity measures. It was later discovered that one can have the complete beam formulation of [Hodges 1990] without carrying displacement or rotation variables as unknowns. This approach may be inferred from [Hegemier and Nair 1977] though exclusion of rotation variables was not mentioned therein; see [Hodges 2009] for additional details of the pertinent history. As a result, the fully intrinsic, geometrically exact partial differential equations of motion and kinematical partial differential equations were presented in [Hodges 2003]. These fully intrinsic equations and boundary conditions involve only force, moment, velocity and, angular velocity variables, with no displacement or rotation variables; see (1)–(6) below.

The goal of this paper is to present a variable-order finite element solution to the fully intrinsic equations. The research assumes that a suitable cross-sectional analysis is available for beams of arbitrary geometry and material distribution. These cross-sectional properties can be calculated using an analytical thin-walled theory [Volovoi and Hodges 2000; Patil and Johnson 2005] or computational FEM analysis

[Giavotto et al. 1983; Cesnik and Hodges 1997; Yu et al. 2002] for general configurations. Although the work of [Volovoi and Hodges 2000] addresses both closed- and open-section beams, the present work pertains only to closed-section beams. The beam formulation of [Hodges 1990; 2003] is not appropriate for beams with open cross-section, which require a separate warping displacement variable and associated boundary conditions [Simo and Vu-Quoc 1991; Yu et al. 2005]. The book [Hodges 2006] gives a more comprehensive treatment for beams of all types.

Beam dynamics analysis can be broken up into five parts:

- Cross-sectional analysis to determine elastic constants and stress/strain recovery relations;

- Beam partial differential equations;

- Techniques for solving those equations (including discretization);

- Nonlinear model-order reduction schemes;

- Application to coupled-field problems such as aeroelasticity, active blades, and control design.

The focus of the present work is on the third item, that is, discretization and solution of the geometrically exact, intrinsic beam equations.

Beam equations. There are a number of geometrically exact formulations for the nonlinear dynamics of beams that can be used for analysis and design [Borri and Mantegazza 1985; Simo and Vu-Quoc 1988; Bauchau and Kang 1993]. The present work is based on the "intrinsic" formulation developed in [Hodges 1990], which can be written in a simple matrix form with only second-degree nonlinearities. To say that these equations are intrinsic is to say that they are independent of displacement and rotation measures. They were used along with generalized strain- and velocity-displacement measures to solve for beam dynamics, yielding excellent agreement with experimental results as given in [Hodges and Patil 2004]. Recently, a set of generalized strain-velocity compatibility relations were derived in [Hodges 2003] which, along with the equations of motion, make up a complete set of intrinsic equations that can be solved without using displacement/rotation measures for certain loading and boundary conditions. Although they incorporate all the nonlinearities and anisotropic couplings, these equations are very simple.

Discrete equations of motion. The equations described above are partial differential equations in space and time for 12 variables (force, moment, velocity, and angular velocity vectors). The solution of these equations requires discretization in space to convert the equations to ordinary differential equations in time. For example, one may use finite elements or a series of assumed functions. It is also possible to use a combination of the two, where the order of the polynomials within the elements as well as the number of elements both vary, leading to variable-order (or *hp*) finite elements [Gui and Babuška 1986].

A simple nonlinear finite element representation of the beam equations was presented in [Hodges 2003]. The FEM equations were used successfully to conduct nonlinear dynamic analysis and control design of integrally actuated helicopter blades [Traugott et al. 2006] as well as the nonlinear aeroelastic analysis of high-aspect-ratio flying wing configurations [Patil and Hodges 2006]. A new nonlinear, energy-consistent, Galerkin approach has been developed recently and can be found in [Patil and Althoff 2006]. The Galerkin approach leads to a highly accurate solution with a low computational cost. The application of the nonlinear Galerkin approach is feasible because of the simplicity of the intrinsic equations. Since the highest degree nonlinearity is quadratic, the Galerkin integrals can be evaluated

exactly without resorting to numerical quadrature when the properties are uniform along the span. This contributes to the method's accuracy and efficiency.

The Galerkin approach leads to the exact nonlinear solution (to five decimal places) with as few as 10 assumed functions per variable. Also it is clear [Patil and Althoff 2006] from the slow convergence of the low-order FEM approach, that one would require very large number of finite element nodes to generate results with the same order of accuracy as those generated by the Galerkin approach using 10 functions per variable.

The nonlinear Galerkin approach leads to an approximate solution that is more accurate for a given number of unknowns (or that has fewer unknowns for a given level of accuracy). This is especially important if one needs to conduct numerous optimizations, simulations, and control scenarios. The limitation of the Galerkin approach is the handling of discontinuities. Thus, when one has a complex beam, such as a helicopter rotor blade with multiple changes in properties along the span, then it will be more accurate to break the beam into multiple finite elements, to each of which one can apply the Galerkin approach. This would lead to variable-order finite elements. It is the focus of this research to develop a variable-order finite element scheme to optimally model future structures, including helicopter blades and aircraft wings.

2. Nonlinear, intrinsic beam equations

The nonlinear, fully intrinsic governing equations for the dynamics of a nonuniform, initially curved and twisted, anisotropic beam undergoing large deflections and rotations are given as

$$F' + (\tilde{k} + \tilde{\kappa})F + f = \dot{P} + \widetilde{\Omega}P, \tag{1}$$

$$M' + (\tilde{k} + \tilde{\kappa})M + (\tilde{e}_1 + \tilde{\gamma})F + m = \dot{H} + \widetilde{\Omega}H + \widetilde{V}P, \tag{2}$$

$$V' + (\tilde{k} + \tilde{\kappa})V + (\tilde{e}_1 + \tilde{\gamma})\Omega = \dot{\gamma}, \tag{3}$$

$$\Omega' + (\tilde{k} + \tilde{\kappa})\Omega = \dot{\kappa}, \tag{4}$$

where $(\ )'$ denotes the partial derivative with respect to the axial coordinate of the undeformed beam, $\dot{(\ )}$ denotes the partial derivative with respect to time, $\tilde{z}$ is the skew-symmetric cross-product operator matrix corresponding to the column matrix z, $F(x, t)$ and $M(x, t)$ are the measure numbers of the internal force and moment vectors (cross-section stress resultants), $P(x, t)$ and $H(x, t)$ are the measure numbers of the linear and angular momentum vector (generalized momenta), $\gamma(x, t)$ and $\kappa(x, t)$ are the beam strains and curvatures (generalized strains), $V(x, t)$ and $\Omega(x, t)$ are velocity and angular velocity measures (generalized velocities), and $f(x, t)$ and $m(x, t)$ are the external force and moment measures. Measure numbers of all variables except for k are calculated in the deformed beam cross-sectional frame. The initial twist and curvature of the beam are represented by $k(x) = \lfloor k_1(x) \ \ k_2(x) \ \ k_3(x) \rfloor$, the measure numbers of which are in the undeformed beam cross-sectional frame. Finally, $e_1 = \lfloor 1 \ \ 0 \ \ 0 \rfloor^T$. The first two equations in the above set are the equations of motion [Hodges 1990] while the last two are the intrinsic kinematical equations [Hodges 2003] derived from the generalized strain- and velocity-displacement equations.

The cross-sectional stress resultants of the beam are related to the generalized strains via the cross-sectional stiffnesses or flexibilities. These cross-sectional properties can be calculated using an analytical thin-walled theory [Volovoi and Hodges 2000; Patil and Johnson 2005] or computational FEM analysis

[Cesnik and Hodges 1997; Yu et al. 2002] for a general configuration. Such an analysis gives the linear constitutive law

$$\begin{Bmatrix} \gamma \\ \kappa \end{Bmatrix} = \begin{bmatrix} \mathrm{R} & \mathrm{S} \\ \mathrm{S}^T & \mathrm{T} \end{bmatrix} \begin{Bmatrix} F \\ M \end{Bmatrix},$$

(5)

where $\mathrm{R}(x)$, $\mathrm{S}(x)$, and $\mathrm{T}(x)$, are the cross-sectional flexibilities of the beam. This linear constitutive law is valid only for small strain, but the beam deflections and rotations due to deformation still may be large.

The generalized momenta are related to the generalized velocities via the cross-sectional inertia matrix

$$\begin{Bmatrix} P \\ H \end{Bmatrix} = \begin{bmatrix} \mu\Delta & -\mu\tilde{\xi} \\ \mu\tilde{\xi} & I \end{bmatrix} \begin{Bmatrix} V \\ \Omega \end{Bmatrix} = \begin{bmatrix} \mathrm{G} & \mathrm{K} \\ \mathrm{K}^T & I \end{bmatrix} \begin{Bmatrix} V \\ \Omega \end{Bmatrix},$$

(6)

where $\mu(x)$, $\xi(x)$, $I(x)$ are, respectively, the mass per unit length, the mass center offset (a vector in the cross-section from the beam reference axis to the cross-sectional mass center), and the cross-sectional inertia matrix consisting of mass moments of inertia per unit length on the diagonals, with $I_{11} = I_{22} + I_{33}$, plus the mass product of inertia per unit length I_{23}.

Usually, the constitutive laws are used to replace some variables in terms of others. Here it was decided to express the generalized strains in terms of the cross-section stress resultants, allowing easy specification of zero flexibility, and the generalized momenta in terms of generalized velocities, allowing easy specification of zero inertia. Thus, the primary variables of interest are F, M, V, and Ω.

Finally the boundary conditions need to be specified. For the given beam of length L, there will be two boundary conditions at each end. In this paper we consider only primitive boundary conditions of the form

$$V(0, t) = \mathrm{V}^0 \quad \text{or} \quad F(0, t) = \mathrm{F}^0, \qquad \Omega(0, t) = \Omega^0 \quad \text{or} \quad M(0, t) = \mathrm{M}^0,$$
$$V(\mathrm{L}, t) = \mathrm{V}^\mathrm{L} \quad \text{or} \quad F(\mathrm{L}, t) = \mathrm{F}^\mathrm{L}, \qquad \Omega(\mathrm{L}, t) = \Omega^\mathrm{L} \quad \text{or} \quad M(\mathrm{L}, t) = \mathrm{M}^\mathrm{L},$$

(7)

Boundary conditions involving attached springs or bodies at the ends may be formulated in terms of fully intrinsic variables as well.

For ease of presentation, we consider a beam clamped at its root. It should be noted that the formulation as well as the conclusions presented are general enough to be applicable to all possible boundary conditions [Sotoudeh and Hodges 2011]. Thus, the assumed boundary conditions are

$$V(0, t) = \mathrm{V}^0, \quad \Omega(0, t) = \Omega^0, \quad F(\mathrm{L}, t) = \mathrm{F}^\mathrm{L}, \quad M(\mathrm{L}, t) = \mathrm{M}^\mathrm{L}.$$

(8)

These equations are intrinsic and do not contain displacement and rotation variables. To calculate the displacements and rotations, the following equations, which relate the strains and curvatures to displacements and rotations, are used:

$$(r + u)' = C^{\mathrm{def}}(\gamma + e_1), \quad C^{\mathrm{def}'} = -(\tilde{\kappa} + \tilde{k})C^{\mathrm{def}},$$

(9)

where r is the position vector of the beam axis from the origin of the reference frame, u is the displacement, and C^{def} is the rotation matrix of the deformed beam cross-sectional frame relative to the undeformed beam cross-sectional frame.

3. Energy-consistent weighting

Let us assume that the beam is discretized into n elements as shown in the figure. To create a finite element model we need to choose trial functions as well as weighting functions.
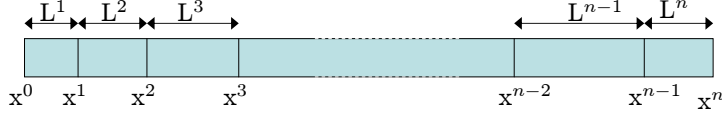
Let the solution in the i-th element be given by V^i, Ω^i, F^i, and M^i. We require that it satisfy (approximately) the equations of motion, the kinematic equations, and the boundary conditions given above. We also require that continuity equations be satisfied (approximately) between adjacent elements:

$$V^i(\mathrm{L}^i, t) = V^{i+1}(0, t), \quad \Omega^i(\mathrm{L}^i, t) = \Omega^{i+1}(0, t), \quad F^i(\mathrm{L}^i, t) = F^{i+1}(0, t), \quad M^i(\mathrm{L}^i, t) = M^{i+1}(0, t).$$

These continuity conditions will be modified for any node at which there is a concentrated mass, a rigid body, a nodal force or moment, or a kink in the axis.

Now consider the following weighting of the equations of motion, kinematical equations, continuity conditions, and boundary conditions:

$$
\sum_{i=1}^{n} \left\langle \int_0^{\mathrm{L}^i} \left\{ V^{i^T}[\dot{P}^i + \widetilde{\Omega}^i P^i - F^{i'} - (\tilde{\mathrm{k}}^i + \tilde{\kappa}^i)F^i - \mathrm{f}^i] + \Omega^{i^T}[\dot{H}^i + \widetilde{\Omega}^i H^i + \widetilde{V}^i P^i - M^{i'} - (\tilde{\mathrm{k}}^i + \tilde{\kappa}^i)M^i \right. \right.
$$
$$
\left. -(\tilde{\mathrm{e}}_1 + \tilde{\gamma}^i)F^i - \mathrm{m}^i] + F^{i^T}[\dot{\gamma}^i - V^{i'} - (\tilde{\mathrm{k}}^i + \tilde{\kappa}^i)V^i - (\tilde{\mathrm{e}}_1 + \tilde{\gamma}^i)\Omega^i] + M^{i^T}[\dot{\kappa}^i - \Omega^{i'} - (\tilde{\mathrm{k}}^i + \tilde{\kappa}^i)\Omega^i] \right\} dx^i \right\rangle
$$
$$
+ \sum_{i=1}^{n-1} \left\langle F^{i+1}(0, t)[V^i(\mathrm{L}^i, t) - V^{i+1}(0, t)] + M^{i+1}(0, t)[\Omega^i(\mathrm{L}^i, t) - \Omega^{i+1}(0, t)] \right.
$$
$$
\left. + V^i(\mathrm{L}^i, t)[F^i(\mathrm{L}^i, t) - F^{i+1}(0, t)] + \Omega^i(\mathrm{L}^i, t)[M^i(\mathrm{L}^i, t) - M^{i+1}(0, t)] \right\rangle - F^1(0, t)^T[V^1(0, t) - \mathrm{V}^0]
$$
$$
- M^1(0, t)^T[\Omega^1(0, t) - \Omega^0] + V^n(\mathrm{L}^n, t)^T[F^n(\mathrm{L}^n, t) - \mathrm{F}^\mathrm{L}] + \Omega^n(\mathrm{L}^n, t)^T[M^n(\mathrm{L}^n, t) - \mathrm{M}^\mathrm{L}] = 0.
$$

Note that the constitutive equations are not included, as these equations are satisfied exactly. Since the continuity conditions and the boundary conditions are satisfied weakly, we will not get exact satisfaction of the boundary conditions and the variables will not be continuous at the nodes.

We are conducting a weighted residual above and thus there are no guarantees in terms of energy conservation. Furthermore, we are adding the weighted residuals of various types of equations including equations describing kinetics, kinematics, and continuity. But, if we appropriately choose the weighting function (for each equation over space and relative to the other equations), energy conservation can result. We know that the exact solution satisfies energy conservation, but by the correct choice of the weighting functions we can maintain energy conservation in the reduced space. To prove energy conservation, we integrate these weighted residual equation by parts and simplify to give

$$
\sum_{i=1}^{n} \int_0^{\mathrm{L}^i} (V^{i^T}\dot{P}^i + \Omega^{i^T}\dot{H}^i)\, dx + \sum_{i=1}^{n} \int_0^{\mathrm{L}^i} (F^{i^T}\dot{\gamma}^i + M^{i^T}\dot{\kappa}^i)\, dx
$$
$$
= \sum_{i=1}^{n} \int_0^{\mathrm{L}^i} (V^{i^T}\mathrm{f}^i + \Omega^{i^T}\mathrm{m}^i)\, dx + [V^n(\mathrm{L}^n, t)^T\mathrm{F}^\mathrm{L} + \Omega^n(\mathrm{L}^n, t)^T\mathrm{M}^\mathrm{L} - F^1(0, t)^T\mathrm{V}^0 - M^1(0, t)^T\Omega^0]. \quad (10)
$$

The first term above is the rate of change of kinetic energy, and the second is the rate of change of potential energy. The third is the rate of work done (power) due to applied forces in the interior of the beam, and the fourth is the power due to applied forces at the boundaries. The equation states that the rate of change of the energy of the beam is equal to the rate of work done on the beam. Thus, the above weighting of all the equations leads to an energy balance, on the basis of which we derive the FEM equations. It should be noted that energy conservation is not an approximation but is satisfied exactly.

4. Variable-order FEM

The independent trial functions used are the shifted Legendre polynomials [Abramowitz and Stegun 1964], denoted by $\mathcal{P}^j(\bar{x})$, which constitute a complete set of polynomials that are orthogonal over the shifted interval $0 \le \bar{x} \le 1$, so that

$$\int_0^1 \mathcal{P}^j(\bar{x})\mathcal{P}^k(\bar{x})\,d\bar{x} = \frac{\delta_{jk}}{2i+1}. \tag{11}$$

These polynomials can be obtained from the recursive relations

$$\mathcal{P}^0(\bar{x}) = 1, \quad \mathcal{P}^1(\bar{x}) = 2\bar{x} - 1, \quad \mathcal{P}^{i+1}(\bar{x}) = \frac{(2i+1)(2\bar{x}-1)\mathcal{P}^i(\bar{x}) - i\,\mathcal{P}^{i-1}(\bar{x})}{i+1}. \tag{12}$$

The use of orthogonality makes certain of the linear coefficient matrices diagonal.

Expanding all twelve variables in terms of these polynomials, one finds that the unknowns can be written as

$$V^i(x^i, t) = \sum_{j=0}^m \mathcal{P}^j(\bar{x}^i)v^{j,i}(t), \qquad \Omega^i(x^i, t) = \sum_{j=0}^m \mathcal{P}^j(\bar{x}^i)\omega^{j,i}(t),$$

$$F^i(x^i, t) = \sum_{j=0}^m \mathcal{P}^j(\bar{x}^i)f^{j,i}(t), \qquad M^i(x^i, t) = \sum_{j=0}^m \mathcal{P}^j(\bar{x}^i)m^{j,i}(t), \tag{13}$$

where $\bar{x}^i = x^i/L^i$, and $v^{j,i}$, $\omega^{j,i}$, $f^{j,i}$, and $f^{j,i}$ are column matrices of the unknowns of the formulation, corresponding to the i-th element and j-th order. With $i = 1, 2, \ldots, n$ and $j = 0, 1, \ldots, m$, we have a total of $12(m+1)n$ variables.

The FEM equations for the i-th element can be derived based on the energy-conserving integral equation (10) as

$$\int_0^{L^i} \mathcal{P}^k[(G^i\mathcal{P}^j\dot{v}^{j,i} + K^i\mathcal{P}^j\dot{\omega}^{j,i}) + \widetilde{\mathcal{P}^l\omega^{l,i}}(G^i\mathcal{P}^j v^{j,i} + K^i\mathcal{P}^j\omega^{j,i}) - \mathcal{P}^{j'}f^{j,i}$$
$$- (\tilde{k}^i + S^{i^T}\widetilde{\mathcal{P}^l f^{l,i}} + T^i\widetilde{\mathcal{P}^l m^{l,i}})\mathcal{P}^j f^{j,i} - f^i]dx^i + \mathcal{P}^k(1)[\mathcal{P}^j(1)f^{j,i} - \mathcal{P}^j(0)f^{j,i+1}] = 0, \tag{14}$$

$$\int_0^{L^i} \mathcal{P}^k[(K^{i^T}\mathcal{P}^j\dot{v}^{j,i} + I^i\mathcal{P}^j\dot{\omega}^{j,i}) + \widetilde{\mathcal{P}^l\omega^{l,i}}(K^{i^T}\mathcal{P}^j v^{j,i} + I^i\mathcal{P}^j\omega^{j,i}) + \widetilde{\mathcal{P}^l v^{l,i}}(G^i\mathcal{P}^j v^{j,i} + K^i\mathcal{P}^j\omega^{j,i}) - \mathcal{P}^{j'}m^{j,i}$$
$$- (\tilde{k}^i + S^{i^T}\widetilde{\mathcal{P}^l f^{l,i}} + T^i\widetilde{\mathcal{P}^l m^{l,i}})\mathcal{P}^j m^{j,i} - (\tilde{e}_1 + R^i\widetilde{\mathcal{P}^l f^{l,i}} + S^i\widetilde{\mathcal{P}^l m^{l,i}})\mathcal{P}^j f^{j,i} - m^i]dx^i$$
$$+ \mathcal{P}^k(1)[\mathcal{P}^j(1)m^{j,i} - \mathcal{P}^j(0)m^{j,i+1}] = 0, \tag{15}$$

$$\int_0^{\mathrm{L}^i} \mathscr{P}^k[(\mathrm{R}^i\mathscr{P}^j\dot{f}^{j,i}+\mathrm{S}^i\mathscr{P}^j\dot{m}^{j,i})-\mathscr{P}^{j\prime}v^{j,i}-(\tilde{\mathrm{k}}^i+\mathrm{S}^{i^T}\widetilde{\mathscr{P}^l f^{l,i}}+\mathrm{T}^i\widetilde{\mathscr{P}^l m^{l,i}})\mathscr{P}^j v^{j,i}$$
$$-(\tilde{\mathrm{e}}_1+\mathrm{R}^i\widetilde{\mathscr{P}^l f^{l,i}}+\mathrm{S}^i\widetilde{\mathscr{P}^l m^{l,i}})\mathscr{P}^j\omega^{j,i}]dx^i+\mathscr{P}^k(0)[\mathscr{P}^j(1)v^{j,i-1}-\mathscr{P}^j(0)v^{j,i}]=0,\quad(16)$$

$$\int_0^{\mathrm{L}^i}\mathscr{P}^k[(\mathrm{S}^{i^T}\mathscr{P}^j\dot{f}^{j,i}+\mathrm{T}^i\mathscr{P}^j\dot{m}^{j,i})-\mathscr{P}^{j\prime}\omega^{j,i}-(\tilde{\mathrm{k}}^i+\mathrm{S}^{i^T}\widetilde{\mathscr{P}^l f^{l,i}}+\mathrm{T}^i\widetilde{\mathscr{P}^l m^{l,i}})\mathscr{P}^j\omega^{j,i}]dx^i$$
$$+\mathscr{P}^k(0)[\mathscr{P}^j(1)\omega^{j,i-1}-\mathscr{P}^j(0)\omega^{j,i}]=0.\quad(17)$$

In these equations R^i, S^i, and T^i are the cross-sectional flexibility coefficients for the i-th element, G^i, K^i, I^i are the cross-sectional inertia coefficients for the i-th element, k^i is the initial curvature for the i-th element, and f^i and m^i define the loading for the i-th element. In the equations, summation is assumed over indices j and l. Thus we have a set of equation for each i (element) and k (order), giving us a total of $12(m+1)n$ equations for as many unknowns.

We need to calculate the above integrals so as to obtain the equations in a form suitable for solution. For demonstration, we assume that the cross-sectional properties, the initial twist and curvature, and the distributed loading are all constant within each element. With the above assumptions, the FEM equations for the i-th element can be derived and we obtain the discretized equations of motion as

$$\mathscr{A}^{kj}\mathrm{L}^i(\mathrm{G}^i\dot{v}^{j,i}+\mathrm{K}^i\dot{\omega}^{j,i})+\mathscr{C}^{kjl}\mathrm{L}^i\tilde{\omega}^{l,i}(\mathrm{G}^i v^{j,i}+\mathrm{K}^i\omega^{j,i})-\mathscr{B}^{kj}f^{j,i}$$
$$-\mathscr{A}^{kj}\mathrm{L}^i\tilde{\mathrm{k}}^i f^{j,i}-\mathscr{C}^{kjl}\mathrm{L}^i(\mathrm{S}^{i^T}\widetilde{f^{l,i}}+\widetilde{\mathrm{T}^i m^{l,i}})f^{j,i}-\mathscr{D}^k\mathrm{L}^i\mathrm{f}^i+\mathscr{P}_1^k\mathscr{P}_1^j f^{j,i}-\mathscr{P}_1^k\mathscr{P}_0^j f^{j,i+1}=0,\quad(18)$$

$$\mathscr{A}^{kj}\mathrm{L}^i(\mathrm{K}^{i^T}\dot{v}^{j,i}+I^i\dot{\omega}^{j,i})+\mathscr{C}^{kjl}\mathrm{L}^i\tilde{\omega}^{l,i}(\mathrm{K}^{i^T}v^{j,i}+I^i\omega^{j,i})+\mathscr{C}^{kjl}\mathrm{L}^i\tilde{v}^{l,i}(\mathrm{G}^i v^{j,i}+\mathrm{K}^i\omega^{j,i})-\mathscr{B}^{kj}m^{j,i}$$
$$-\mathscr{A}^{kj}\mathrm{L}^i\tilde{\mathrm{k}}^i m^{j,i}-\mathscr{C}^{kjl}\mathrm{L}^i(\mathrm{S}^{i^T}\widetilde{f^{l,i}}+\widetilde{\mathrm{T}^i m^{l,i}})m^{j,i}-\mathscr{A}^{kj}\mathrm{L}^i\tilde{\mathrm{e}}_1 f^{j,i}-\mathscr{C}^{kjl}\mathrm{L}^i(\widetilde{\mathrm{R}^i f^{l,i}}+\widetilde{\mathrm{S}^i m^{l,i}})f^{j,i}$$
$$-\mathscr{D}^k\mathrm{L}^i\mathrm{m}^i+\mathscr{P}_1^k\mathscr{P}_1^j m^{j,i}-\mathscr{P}_1^k\mathscr{P}_0^j m^{j,i+1}=0,\quad(19)$$

$$\mathscr{A}^{kj}\mathrm{L}^i(\mathrm{R}^i\dot{f}^{j,i}+\mathrm{S}^i\dot{m}^{j,i})-\mathscr{B}^{kj}v^{j,i}-\mathscr{A}^{kj}\mathrm{L}^i\tilde{\mathrm{k}}^i v^{j,i}-\mathscr{C}^{kjl}\mathrm{L}^i(\mathrm{S}^{i^T}\widetilde{f^{l,i}}+\widetilde{\mathrm{T}^i m^{l,i}})v^{j,i}$$
$$-\mathscr{A}^{kj}\mathrm{L}^i\tilde{\mathrm{e}}_1\omega^{j,i}-\mathscr{C}^{kjl}\mathrm{L}^i(\widetilde{\mathrm{R}^i f^{l,i}}+\widetilde{\mathrm{S}^i m^{l,i}})\omega^{j,i}-\mathscr{P}_0^k\mathscr{P}_0^j v^{j,i}+\mathscr{P}_0^k\mathscr{P}_1^j v^{j,i-1}=0,\quad(20)$$

$$\mathscr{A}^{kj}\mathrm{L}^i(\mathrm{S}^{i^T}\dot{f}^{j,i}+\mathrm{T}^i\dot{m}^{j,i})-\mathscr{B}^{kj}\omega^{j,i}-\mathscr{A}^{kj}\mathrm{L}^i\tilde{\mathrm{k}}^i\omega^{j,i}-\mathscr{C}^{kjl}\mathrm{L}^i(\mathrm{S}^{i^T}\widetilde{f^{l,i}}+\widetilde{\mathrm{T}^i m^{l,i}})\omega^{j,i}$$
$$-\mathscr{P}_0^k\mathscr{P}_0^j\omega^{j,i}+\mathscr{P}_0^k\mathscr{P}_1^j\omega^{j,i-1}=0.\quad(21)$$

In the above equations summation is implied over indices j and l, and $\mathscr{A}^{kj}$, $\mathscr{B}^{kj}$, $\mathscr{C}^{kjl}$, and $\mathscr{D}^k$ are dimensionless integrals, given by

$$\mathscr{A}^{kj}=\int_0^1\mathscr{P}^k(\bar{x})\mathscr{P}^j(\bar{x})d\bar{x},\qquad\qquad\mathscr{B}^{kj}=\int_0^1\mathscr{P}^k(\bar{x})\big(\mathscr{P}^j(\bar{x})\big)'d\bar{x},$$
$$\mathscr{C}^{kjl}=\int_0^1\mathscr{P}^k(\bar{x})\mathscr{P}^j(\bar{x})\mathscr{P}^l(\bar{x})d\bar{x},\qquad\mathscr{D}^k=\int_0^1\mathscr{P}^k(\bar{x})d\bar{x}.\tag{22}$$

Now, representing all the system unknowns of the i-th element as

$$
q^i(t) = \begin{Bmatrix} v^{0,i}(t) \\ \omega^{0,i}(t) \\ f^{0,i}(t) \\ m^{0,i}(t) \\ \vdots \\ v^{m,i}(t) \\ \omega^{m,i}(t) \\ f^{m,i}(t) \\ m^{m,i}(t) \end{Bmatrix} \quad \text{and} \quad q(t) = \begin{Bmatrix} q^1(t) \\ q^2(t) \\ \vdots \\ q^n(t) \end{Bmatrix}, \tag{23}
$$

the complete system consists of $N = 12n(m+1)$ equations and unknowns. The equations can then be written in the form

$$
A_{ji}\dot{q}_i + B_{ji}q_i + C_{jik}q_iq_k + D_j = 0, \tag{24}
$$

where summation is assumed over the indices i and k.

5. Results

The equations were solved using the variable-order FEM for a simple prismatic beam case with these data:

Span	16 m
Chord	1 m
Mass per unit length	0.75 kg/m
Mom. inertia (50% chord)	0.1 kg m
Spanwise elastic axis	50% chord
Center of gravity	50% chord
Bending rigidity	2×10^4 Nm^2
Torsional rigidity	1×10^4 Nm^2
Bending rigidity (chordwise)	4×10^6 Nm^2
Shear/extensional rigidity	∞

Three examples are presented: (1) a beam undergoing large deformation due to a tip follower force, (2) the natural frequencies and modeshapes of a beam, and (3) the natural frequencies and modeshapes of a rotating beam (this involves calculation of the nonlinear steady state and linearizing the system about this nonlinear steady state).

Let us first consider a beam loaded at the tip with a transverse follower force. As the force increases, the deformation increases; and due to the large deformation, the force direction as well as its position relative to the beam root changes. The correct nonlinear solution thus has a lower root bending moment as compared to the linear solution. The exact solution to the problem satisfies the transcendental equation

$$
\frac{1}{6}\bar{m}\sqrt{4 - \bar{m}^4\bar{p}^2} + \frac{2\sqrt{2}}{3\sqrt{\bar{p}}}F\left[\sin^{-1}\left(\frac{\bar{m}\sqrt{\bar{p}}}{\sqrt{2}}\right)\,\middle|\,-1\right] = 1 - \bar{x}, \tag{25}
$$

where $\bar{m}$ is the bending moment made dimensionless by PL, $\bar{p}$ is the tip force made dimensionless by EI/L^2, $\bar{x}$ is the axial coordinate made dimensionless by L, and $F(\phi|k)$ is the elliptic integral of the first kind with $k = -1$ and $\sin\phi = \bar{m}\sqrt{\bar{p}}/\sqrt{2}$.

Figure 1 shows the root bending moment calculated using variable-order finite elements compared to exact results for $\bar{p} = 3$. Figure 1a shows the convergence of the root bending moment as the order of the system increases. The red line corresponds to the *h*-version, and the blue line corresponds to the *p*-method. Finally, the green dots correspond to the *hp*-method. As expected, for this simple case, the Galerkin approximation (the *p*-version) is the best of the three. It should be noted that when the Galerkin approximation is applied to beams with properties varying along the span, it may not be possible to calculate the integrals exactly. Thus, the Galerkin approach may become computationally intensive for a general configuration. Figure 1b shows the convergence of the various methods. The Galerkin approximation is seen to reach the exact result with error of the order of machine precision using twelfth-order polynomials. The linear, first-order finite elements have about third-order convergence, that is, for every doubling of the number of finite elements, the error decreases by a factor of eight (that is, 2^3). Figure 1c shows the convergence of the error for finite elements of various orders. The quadratic finite elements exhibit fifth-order convergence, and the cubic finite elements have seventh-order convergence. Finally, the quartic finite elements show a whopping ninth-order convergence, that is, for every doubling of the
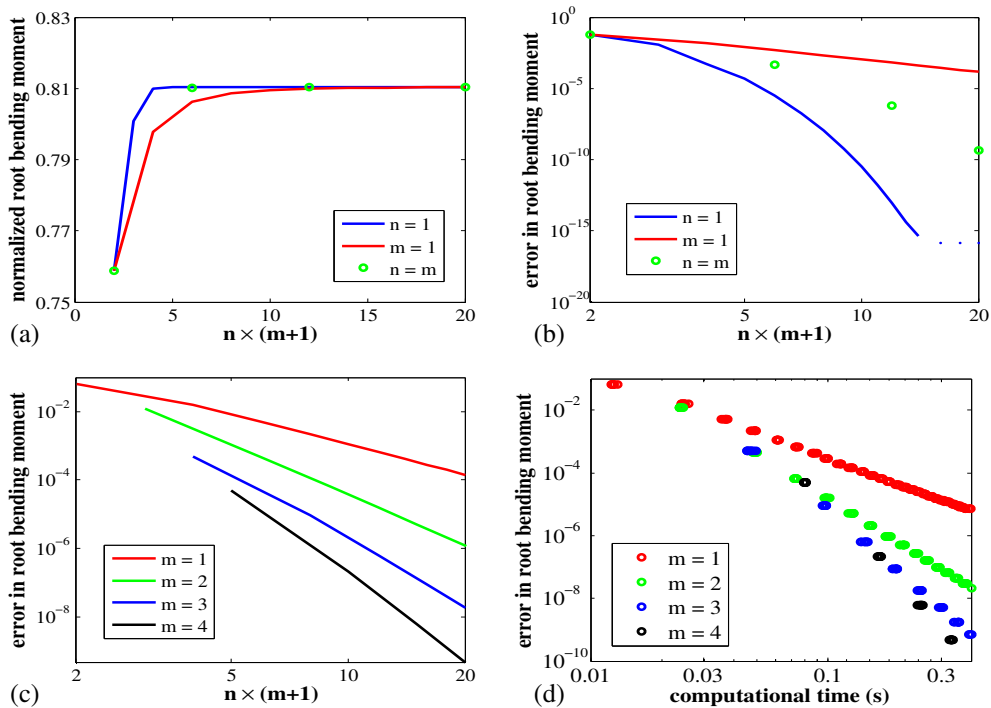
Figure 1. Results for the root bending moment due to follower force at the tip: (a) convergence of root bending moment; (b) error in root bending moment; (c) error in root bending moment for various model orders; (d) error in root bending moment relative to computational time.

number of finite elements there is a reduction in the error by a factor of 512. In other words, the element of n-th order has a convergence of order $2n + 1$, so that doubling of the number of elements decreases the error by a factor of 2^{2n+1}.

To further understand the effect of order on the accuracy and computational time, we have plotted the error in results relative to computational time for various order finite elements. It should be noted that the computational time used in this plot does not include time for calculation of the analytical integrals, which can be calculated once for constant properties and stored. For variable properties one will have to recalculate the integrals which can take considerable time. In the present case for a constant property beam, for the same amount of time (0.4 s) we can use 54×12 variables (27 elements) for $m = 1$, 45×12 variables (15 elements) for $m = 2$, 32×12 variables (8 elements) for $m = 3$, and between 20×12 and 25×12 variables (4–5 elements) for $m = 4$. One can see that in terms of computational time (even without recalculation of integrals), the various order finite elements are much closer together. The linear finite elements still have relatively slow convergence, but the quadratic finite elements have behavior more like cubic ones, and the cubic finite elements are still more like the quartic ones. The calculations were done using the Matlab linear equation solver for sparse matrices. So, we can see that sparsity matters and we can analyze more variables with lower order finite elements, but the order of accuracy of higher-order finite elements increases at a faster rate than the computational time and thus they are in general preferable. For beams with varying properties one may increase efficiency by using more elements so that the preference will be for lower-order finite elements. One can recalculate the integrals and matrices for higher-order elements analytically or by Gauss integrations by taking into account the variations, but it may increase substantially the computational effort. Thus, one should use very high-order finite elements or the Galerkin form only if model order reduction is important (for example, in control design) or if we are reusing the calculated integrals multiple times (for example, for time marching). Finally, for problems with multiple (stiffness or load) discontinuities, we would have no choice but to place nodes at the discontinuities, as the solution using higher-order elements (without nodes at discontinuities) will not give good results near the discontinuities. Again, hp finite elements are good in that one can pick and choose an appropriate order depending on the problem at hand.

Now consider the frequencies of nonrotating as well as rotating beams. Table 1 lists the calculated frequencies and compares them with exact results from [Wright et al. 1982]. The frequency predictions

	Cantilevered beam $\omega = 0, \ v = 0$				Rotating cantil. beam $\omega = 3.189 \, \text{rad/s}, \ v = 0$				Rot. cantil. beam with offset $\omega = 3.189 \, \text{rad/s}, \ v = 51.03 \, \text{m/s}$			
Mode	Exact	$n=9$ $m=1$	$n=3$ $m=3$	$n=1$ $m=9$	Exact	$n=9$ $m=1$	$n=3$ $m=3$	$n=1$ $m=9$	Exact	$n=9$ $m=1$	$n=3$ $m=3$	$n=1$ $m=9$
1st bending	2.243	2.243	2.243	2.243	4.114	4.114	4.114	4.114	5.703	5.703	5.703	5.703
2nd bending	14.06	14.03	14.06	14.06	16.23	16.21	16.23	16.23	18.72	18.69	18.72	18.72
3rd bending	39.36	39.22	39.38	39.36	41.59	41.44	41.62	41.59	44.50	44.33	44.53	44.50
1st torsion	31.05	31.05	31.05	31.05								
2nd torsion	93.14	93.17	93.14	93.14								

Table 1. Beam structural frequencies (rad/s).

for a nonrotating as well as a rotating beam using the present approach with $mn = 9$ are shown. For $m = 9$ and $n = 1$, that is, a single, high-order element, the frequencies are obtained to three significant digits for both the bending and torsion modes. This approach is equivalent to the Galerkin approach discussed in [Patil and Althoff 2006] and to formulations commonly referred to as p-version [Babuška et al. 1981]. On the other hand, for $m = 1$ and $n = 9$, the maximum number of the crudest possible elements for $mn = 9$, the solution is not as accurate, leading to errors greater than 0.3% for the third bending mode. This approach corresponds to the lowest-order FEM method, commonly referred to as h-version, consisting of linear shape functions. Finally, the case of $m = 3$ and $n = 3$ is a good balance between these two extreme approaches. Here cubic polynomials are used to represent the variables in each of three elements. The results are quite good with negligible errors.

The left column of Figure 2 presents the convergence of results for the first bending mode of a nonrotating beam. Figure 2a shows the convergence of frequencies with increase in the order of the system. As expected, again the Galerkin approximation is the best of the three. Figure 2b shows the convergence of the various methods. The Galerkin approximation is seen to reach the exact result with error of the order of machine precision using eighth-order polynomials. Figure 2c shows the convergence of the error in the first frequency for finite elements of various orders. Similar to the nonlinear solution results, for the linearized perturbation results the linear finite element shows third-order convergence, quadratic finite elements show a fifth-order convergence, cubic finite elements show a seventh-order convergence, and quartic finite elements show a ninth-order convergence. Finally, Figure 2d shows the convergence of mode shape for three implementations. The mode shape obtained by the Galerkin approach is the closest to the exact mode shape. The mode shape predicted by the lowest-order, linear finite elements shows deviation from the exact mode shape and is discontinuous as expected, because the continuity conditions are weakly satisfied. Similarly, the mode shape does not satisfy the boundary conditions exactly because of weak satisfaction of the boundary conditions.

The right column of Figure 2 shows the corresponding plots for the first torsion mode. For the torsional frequencies the Galerkin approximation again leads to the most accurate results while the accuracy increases with the increase in the order of the finite element. The order of the relative error is third-order, fifth-order, seventh-order and, ninth-order for the linear, quadratic, cubic, and quartic elements respectively.

Figure 3 shows the results for the first bending mode of a rotating beam with root velocity. For the rotating beam, the static steady-state solution is nontrivial. Thus, the accuracy of the frequencies obtained from linearizing about the nonlinear steady state is dependent on the accuracy of the steady-state solution and the accuracy of the linearized perturbation. We obtain the exact steady-state solution for finite elements of second and higher order. The errors are in general higher for the rotating beam as compared to the nonrotating beam. Furthermore, the rate of convergence for the rotating beam is slightly slower than that of the nonrotating beam.

We have addressed the impact of bandedness or sparsity of the matrices in the variable-order finite element formulation. For low-order elements, the coefficient matrices are very sparse and one can take advantage of this. In addition, the coefficient matrices for low-order elements can be calculated using lower-order Gauss integration thus further reducing the computational time. The higher-order element, though more accurate, may not be computationally as efficient as the low-order element. This aspect has not been addressed in the present paper. For example, a global Galerkin approximation requires
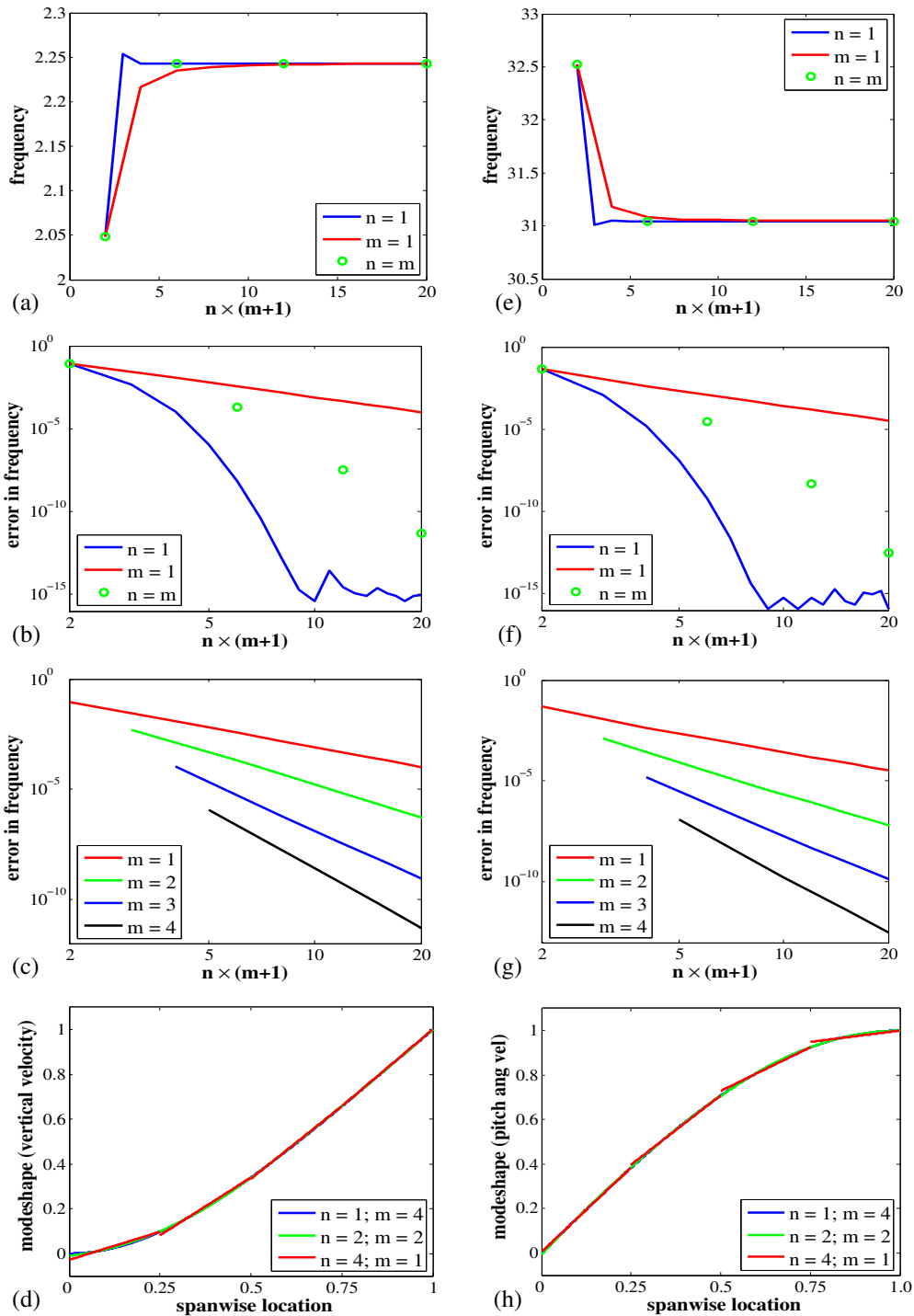
Figure 2. Results for the first bending mode (left column) and the first torsion mode (right column) of a nonrotating cantilevered beam: (a,e) convergence of frequency; (b,f) error in frequency; (c,g) error in frequency; (d,h) convergence of mode shape.
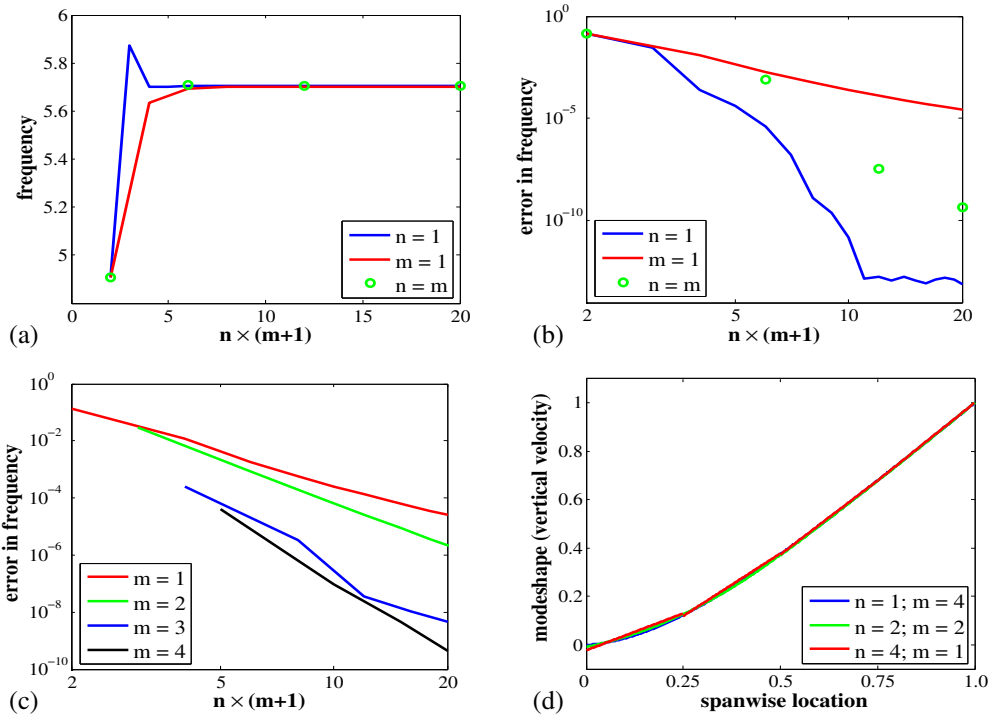
Figure 3. Results for the first bending mode of a rotating cantilevered beam with offset:
(a) convergence of first bending frequency; (b) error in first bending frequency; (c) error
in first bending frequency; (d) convergence of first bending mode shape.

calculation of complex integrals and will lead to fully populated matrices. But, if one is interested
in using nonlinear beam analysis as a part of a multidisciplinary analysis (aeroelasticity), preliminary
design (HALE aircraft, helicopter flight mechanics), and control synthesis, the reduced number of de-
grees of freedom is essential. We have presented a way to calculate the nonlinear dynamics of the
beam accurately using a low number of degrees of freedom that will lead to significant advancement in
nonlinear aeroelastic calculations (which lead to loss of sparsity anyway), flight dynamics simulations
(with complete geometrically nonlinear aeroelasticity), and design optimization. A detailed assessment
of the computational cost of the method vis-à-vis the accuracy for various types of beams and various
applications will be addressed in a later paper.

6. Conclusions

A variable-order finite element technique is presented and applied to beams with uniform properties
along the span. This technique is based on a geometrically exact, fully intrinsic formulation. The results
presented show that one can obtain approximately third-order, fifth-order, seventh-order, and ninth-order
convergence for the linear, quadratic, cubic, and quartic finite elements. It is recommended that one use
quadratic- or higher-order finite elements for a better approximation of mode shapes. The cubic finite
elements provide an especially good balance of accuracy, computational efficiency, and applicability to

general configurations. Additional work needs to be done to assess the method's accuracy when applied to beams with properties varying along the span.

References

[Abramowitz and Stegun 1964] M. Abramowitz and I. A. Stegun (editors), *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, National Bureau of Standards Applied Mathematics Series **55**, U.S. Government Printing Office, Washington, DC, 1964.

[Babuška et al. 1981] I. Babuška, B. A. Szabo, and I. N. Katz, "The *p*-version of the finite element method", *SIAM J. Numer. Anal.* **18**:3 (1981), 515–545.

[Bauchau and Kang 1993] O. A. Bauchau and N. K. Kang, "A multibody formulation for helicopter structural dynamic analysis", *J. Am. Helicopter Soc.* **38**:2 (1993), 3–14.

[Borri and Mantegazza 1985] M. Borri and P. Mantegazza, "Some contributions on structural and dynamic modeling of helicopter rotor blades", *Aerotec. Missili Spaz.* **64**:9 (1985), 143–154.

[Cesnik and Hodges 1997] C. E. S. Cesnik and D. H. Hodges, "VABS: a new concept for composite rotor blade cross-sectional modeling", *J. Am. Helicopter Soc.* **42**:1 (1997), 27–38.

[Danielson 1970] D. A. Danielson, "Simplified intrinsic equations for arbitrary elastic shells", *Int. J. Eng. Sci.* **8**:3 (1970), 251–259.

[Giavotto et al. 1983] V. Giavotto, M. Borri, P. Mantegazza, G. Ghiringhelli, V. Carmaschi, G. C. Maffioli, and F. Mussi, "Anisotropic beam theory and applications", *Comput. Struct.* **16**:1-4 (1983), 403–413.

[Green and Laws 1966] A. E. Green and N. Laws, "A general theory of rods", *Proc. R. Soc. Lond. A* **293**:1433 (1966), 145–155.

[Gui and Babuška 1986] W. Gui and I. Babuška, "The *h, p* and *h-p* versions of the finite element methods in 1 dimension, III: The adaptive *h-p* version", *Numer. Math.* **49**:6 (1986), 659–683.

[Hegemier and Nair 1977] G. A. Hegemier and S. Nair, "A nonlinear dynamical theory for heterogeneous, anisotropic, elastic rods", *AIAA J.* **15**:1 (1977), 8–15.

[Hodges 1990] D. H. Hodges, "A mixed variational formulation based on exact intrinsic equations for dynamics of moving beams", *Int. J. Solids Struct.* **26**:11 (1990), 1253–1273.

[Hodges 2003] D. H. Hodges, "Geometrically exact, intrinsic theory for dynamics of curved and twisted anisotropic beams", *AIAA J.* **41**:6 (2003), 1131–1137.

[Hodges 2006] D. H. Hodges, *Nonlinear composite beam theory*, Progress in Astronautics and Aeronautics **213**, AIAA, Reston, VA, 2006.

[Hodges 2009] D. H. Hodges, "Erratum: geometrically exact, intrinsic theory for dynamics of curved and twisted anisotropic beams", *AIAA J.* **47**:5 (2009), 1308–1309.

[Hodges and Patil 2004] D. H. Hodges and M. J. Patil, "Technical note: correlation of geometrically-exact beam theory with the Princeton data", *J. Am. Helicopter Soc.* **49**:3 (2004), 357–360.

[Patil and Althoff 2006] M. J. Patil and M. Althoff, "Energy-consistent, Galerkin approach for the nonlinear dynamics of beams using mixed, intrinsic equations", in *47th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference* (Newport, RI, 2006), AIAA, Reston, VA, 2006.

[Patil and Hodges 2006] M. J. Patil and D. H. Hodges, "Flight dynamics of highly flexible flying wings", *J. Aircraft* **43**:6 (2006), 1790–1799.

[Patil and Johnson 2005] M. J. Patil and E. R. Johnson, "Cross-sectional analysis of anisotropic, thin-walled, closed-section beams with embedded strain actuation", in *Proceedings of the 46th AIAA/ASME/ASCE/AHS/ASC structures, structural dynamics and materials conference* (Austin, TX, 2005), AIAA, Reston, VA, 2005.

[Reissner 1973] E. Reissner, "On one-dimensional large-displacement finite-strain beam theory", *Stud. Appl. Math.* **52**:2 (1973), 87–95.

[Simo and Vu-Quoc 1988] J. C. Simo and L. Vu-Quoc, "On the dynamics in space of rods undergoing large motions: a geometrically exact approach", *Comput. Methods Appl. Mech. Eng.* **66**:2 (1988), 125–161.

[Simo and Vu-Quoc 1991] J. C. Simo and L. Vu-Quoc, "A geometrically-exact rod model incorporating shear and torsion-warping deformation", *Int. J. Solids Struct.* **27**:3 (1991), 371–393.

[Sotoudeh and Hodges 2011] Z. Sotoudeh and D. H. Hodges, "Modeling beams with various boundary conditions using fully intrinsic equations", *J. Appl. Mech.* **78**:1 (2011), article 031010.

[Traugott et al. 2006] J. P. Traugott, M. J. Patil, and F. Holzapfel, "Nonlinear modeling of integrally actuated beams", *Aerosp. Sci. Technol.* **10**:6 (2006), 509–518.

[Volovoi and Hodges 2000] V. V. Volovoi and D. H. Hodges, "Theory of anisotropic thin-walled beams", *J. Appl. Mech.* (*ASME*) **67**:3 (2000), 453–459.

[Wright et al. 1982] A. D. Wright, C. E. Smith, R. W. Thresher, and J. L. C. Wang, "Vibration modes of centrifugally stiffened beams", *J. Appl. Mech.* (*ASME*) **49**:1 (1982), 197–202.

[Yu et al. 2002] W. Yu, D. H. Hodges, V. V. Volovoi, and C. E. S. Cesnik, "On Timoshenko-like modeling of initially curved and twisted composite beams", *Int. J. Solids Struct.* **39**:19 (2002), 5101–5121.

[Yu et al. 2005] W. Yu, D. H. Hodges, V. V. Volovoi, and E. D. Fuchs, "A generalized Vlasov theory for composite beams", *Thin-Walled Struct.* **43**:9 (2005), 1493–1511.

MAYURESH J. PATIL: mpatil@vt.edu
Department of Aerospace and Ocean Engineering, Virginia Polytechnic Institute and State University,
Blacksburg, VA 24061-0203, United States
http://www.aoe.vt.edu/people/faculty.php?fac_id=mpatil

DEWEY H. HODGES: dhodges@gatech.edu
Daniel Guggenheim School of Aerospace Engineering, Georgia Institute of Technology, 270 Ferst Drive,
Atlanta, GA 30332-0150, United States
http://www.ae.gatech.edu/~dhodges/

EFFECTS OF CONTACT SURFACE SHAPE ON LIFETIME OF CELLULAR FOCAL ADHESION

GREGORY J. RIZZA, JIN QIAN AND HUAJIAN GAO

As a means of better understanding cell-matrix adhesion, we consider the effects that the contact surface shape of a cellular focal adhesion has on its adhesion lifetime. An idealized model of focal adhesion is adopted in which two dissimilar elastic media are bonded together by an array of ligand/receptor bonds modeled as Hookean springs on a curved surface. The cluster of bonds is subjected to a constant applied tensile load F, and the distribution of traction forces on the individual bonds is assumed to obey the elasticity equations of continuous elastic media. The rupturing and rebinding of bonds are described by stochastic equations solved using the Monte Carlo method. The contact surface in the model is scaled by an optimal shape defined as the deformed surface profile of a planar elastic half-space that is subjected to a uniform pressure applied over the contact region with the total force equal to F. Our model shows that contact surface shape does have a substantial impact on adhesion lifetime. The model also shows that the adhesion lifetime becomes increasingly sensitive to variations in contact surface shape as the focal adhesion increases in size.

1. Introduction

Cell adhesion is becoming an important topic of study in all fields involving cellular research. It has been observed that a large amount of cell activities such as cell migration, cell differentiation, wound healing, etc. are dependent upon the ability of a cell to adhere to a variety of biological surfaces [Alberts et al. 2002; Chen and Gumbiner 2006]. An interesting feature of cell adhesion on substrates, which could be characterized as a cell adhering to the extracellular matrix (ECM) within an organism of study, is that the bonds formed between the cell surface and ECM via ligand/receptor connections are usually arranged in clusters of multiple individual bonds [Zamir et al. 2000; Zaidel-Bar et al. 2003]. These bond clusters are of a finite size with a maximum length of only a few microns depending on the stiffness of the system, and have been referred to as focal adhesions (FAs) due to their definite size and discreteness [Alberts et al. 2002; Chen and Gumbiner 2006; Zamir et al. 2000; Zaidel-Bar et al. 2003; Bershadsky et al. 2003]. There are many biological, chemical and mechanical processes which occur during the sequences involved in cellular adhesion. However, for the purposes of this paper we will only examine cellular focal adhesion on the basis of a mechanical model in order to develop a better understanding of the effects of contact surface shape on FA lifetime.

There have been both theoretical and experimental studies in developing an understanding that single ligand/receptor bonds only have finite adhesion lifetime. Even in the absence of any applied load, bond dissociations due to thermal fluctuation are expected to occur in a statistical fashion [Evans and Ritchie

1997; Evans 2001; Evans and Calderwood 2007]. This understanding of single molecular bonds is currently accepted and assumed to be representative of the behavior of a large number of ligand/receptor types in cell adhesion. However, the mechanics of multiple bond adhesions, such as those in focal adhesions, is not understood to the same degree of single bond adhesion due to its complexity.

The study of adhesions made up of many molecular bonds was pioneered by Bell [1978], who developed a kinetic theory to predict the thermodynamic competition between bonds forming and breaking under an applied load. The investigation of multiple bond adhesion was further progressed by the work of Erdmann and Schwartz [2004a; 2004b] who developed stochastic equations and performed Monte Carlo simulations to investigate the adhesion lifetime as a function of cluster size. They showed that, unlike single bond adhesion which only has a limited lifetime, focal adhesion clusters made up of multiple bonds could have more prolonged lifetimes. Their model showed that the larger the focal adhesion becomes the longer adhesion lifetime it can experience. Qian et al. [2008; 2009] were the first to model focal adhesions between cell and substrate in terms of clusters of ligand/receptor bonds between two dissimilar continuous elastic media, showing the effects of varying adhesion size and system stiffness on focal adhesion strength and lifetime. According to their results, a possible explanation for the characteristic micron-size scale of focal adhesions is that small adhesion size leads to single-molecule-like behavior of limited lifetime because of statistical effects whereas large adhesion also leads to single-molecule-like behavior because of the focusing effect of stress concentration that confines the bond dissociation events to a smaller number of bonds near the adhesion edges. Their model also showed that stiffer cell and substrate result in larger and more stable focal adhesions than softer ones. This finding is consistent with a range of experimental observations, including that focal adhesions are small or absent in cells cultured on sufficiently compliant substrates, and that cells tend to actively migrate towards stiffer region when cultured on an elastically nonhomogeneous substrate [Pelham and Wang 1997; Lo et al. 2000; Tan et al. 2003].

So far we know that clustering multiple bonds into focal contacts greatly enhances the lifetime and stability of cell-matrix adhesion [Erdmann and Schwarz 2004a; 2004b]. However, due to the elasticity effects of the system, there is an upper limit in size of a focal adhesion in order for it to avoid stress concentration effect [Qian et al. 2008; 2009]. It has also been shown that stiffness of cell and substrate is a factor in controlling the stability of focal adhesions, and cells have a preference towards stiffer substrates potentially due to the increased focal adhesion lifetime [Qian et al. 2008; 2009].

The present paper explores how contact surface shape can be used as a factor to control focal adhesion lifetime. Here, the contact surface shape is defined as the initial gap between a pair of contacting surfaces in the absence of any external loads [Gao and Yao 2004; Yao and Gao 2006]. In the special case when one of the contact surfaces is a flat plane, the contact surface shape then corresponds to the actual shape of the opposing partner surface. Optimization of contact surface shape is not an uncommon strategy when trying to increase adhesion strength in contact mechanics problems. In biology, optimized surface shapes can be seen in species that rely on nonspecific adhesion for survival such as gecko and many types of insects, and the optimal surface shape for adhesion between two elastic media via van der Waals interaction has been defined as the profile of the contact surface that distributes an applied load uniformly within the adhesion region at the critical moment of pull-off [Gao and Yao 2004; Yao and Gao 2006; Shi et al. 2006]. Following a similar line of thought, here we define the optimal contact surface shape of focal adhesions via specific molecular bonds as the shape that gives rise to a uniform traction distribution

among individual bonds within the adhesion domain whenever the closed bonds are evenly spaced. This even distribution of load avoids stress concentrations along the contact edges, leading to substantially increased lifetime of focal adhesions.

The goal of this article is to explore the effects of contact surface shape on focal adhesion lifetime. Based on previous studies on the lifetime of multiple bond adhesion under constant tensile loads, we plan to demonstrate the effects that altering and optimizing the geometric shape of the adhesion surface have on focal adhesion lifetime. To do this, a range of surface shapes which are related to the optimal shape through a numerical shape factor are studied under various focal adhesion parameters. From this model, we plan to show:

(1) Optimizing surface shape can greatly increase focal adhesion lifetime.

(2) As focal adhesions increase in size, deviations from the optimal shape result in substantial decrease in adhesion lifetime when compared to the cases of focal adhesions with fewer bonds.

(3) Varying system stiffness and applied tensile load on the focal adhesion have interesting effects with respects to parts (1) and (2).

2. Model

The theoretical model under consideration is illustrated in Figure 1, in which two elastic media are seen bonding to each other via ligand/receptor bonds. The upper elastic medium represents a biological cell whereas the lower elastic medium is representative of the extracellular matrix. The Young's moduli and
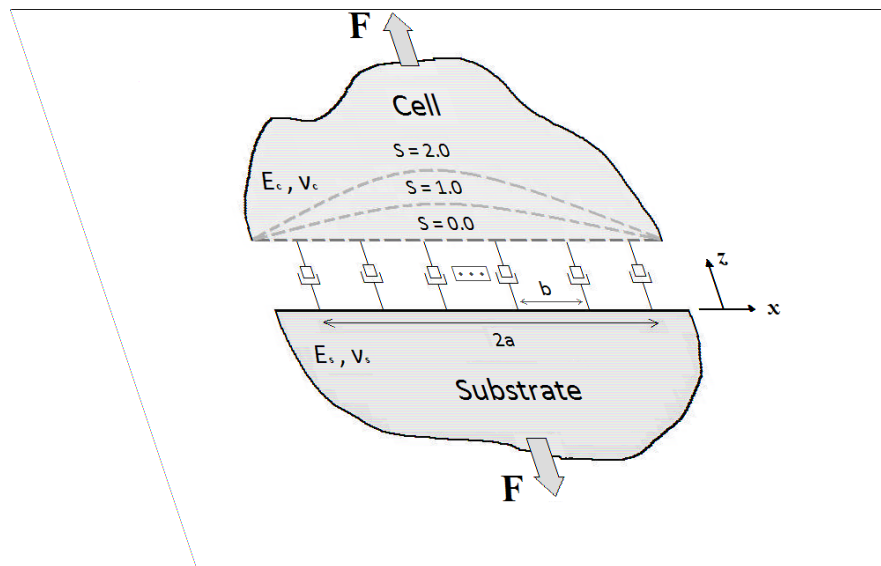
Figure 1. Idealized model of a single focal adhesion between a cell and substrate undergoing tensile loading. The cell and substrate are bonded together by an array of ligand/receptor bonds. The contact surface shapes of the focal adhesion are determined by a shape factor S which ranges in $[0, 2]$, where $S = 1$ corresponds to the optimal contact shape for an initial uniformly loaded bond cluster.

Poisson's ratios of these media are defined as E_c, ν_c for the cell, and E_s, ν_s for the substrate. It is usually convenient to define a combined elastic modulus E^* as

$$\frac{1}{E^*} = \frac{1 - \nu_c^2}{E_c} + \frac{1 - \nu_s^2}{E_s} \tag{2-1}$$

when analyzing contact problems involving two different elastic media [Johnson 1985]. The cell/substrate system is subject to a constant tensile load F such that the load distribution among individual bonds is governed by elasticity equations. The two media are bonded together by ligand/receptor bonds idealized as Hookean springs with a spring coefficient k_{LR}. It is assumed in the model that the bonds do not have any interaction with each other. It is also assumed that when rupture occurs in a bond, the energy required to rupture the ligand/receptor bond is much lower than the energy required to rupture the connection of the ligand or receptor from the cell or substrate surface [Bell 1978]. What this means is when the ligand/receptor bond is undergoing tensile loading, rupturing of the bond will most likely occur at the ligand/receptor interface due to the less energy required for it to rupture. The bonds are evenly spaced at distance b from each other in the lateral x-direction (Figure 1), corresponding to a bond density of $\rho_{LR} = 1/b^2$. As can be seen in Figure 2, the initial location of each bond in the vertical z-direction is determined by a shape factor S. The shape factor value of $S = 0$ corresponds to the surfaces of the cell and substrate being modeled as horizontal planes, $S = 1$ corresponds to an optimized curved surface profile that gives rise to uniform traction distribution among individual ligand/receptor bonds within the adhesion domain under a tensile load (see details in next section), and $S = 2$ corresponds to an overly curved surface in comparison to the optimal surface ($S = 1$). The adhesion being analyzed in this model is defined as a single focal adhesion cluster which is made up of an array of bonds aligned axially with
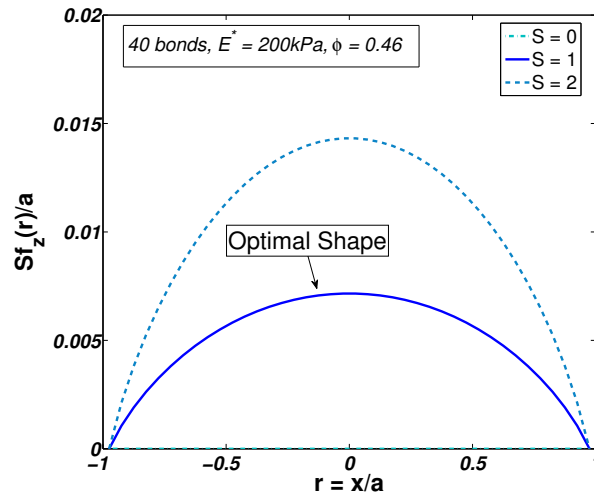
Figure 2. Schematic of a focal adhesion surface containing 40 bonds, having a system stiffness of $E^* = 200\,\text{kPa}$, and a tensile load factor of $\varphi = 0.46$. The surface shapes represented are for the following shape factors: $S = 0$ corresponding to adhesion between conforming surfaces, $S = 1$ corresponding to the optimal shape, and $S = 2$ corresponding to an overly curved focal adhesion surface in comparison to the optimal case.

an out of plane thickness b. The size of the adhesion is defined as $2a$ where a is the half-width of the focal adhesion, and the total number of involved bonds is $N = 2a/b$.

2.1. *Optimal shape and shape factor.*

Following a similar concept introduced in reference [Gao and Yao 2004], an optimal shape of contact surface is defined as the deformed surface profile of a planar half-space that is subjected to a uniform pressure applied over the contact region with the total force equal to F on the bond cluster. For a cluster of evenly spaced ligand/receptor bonds, the optimal contact surface shape that gives rise to a uniform traction distribution among individual bonds within the adhesion domain is

$$f_z(x) = -\frac{\varphi \cdot F_b}{\pi E^* b^2} \left((a+x) \ln \left(\frac{a+x}{a}\right)^2 + (a-x) \ln \left(\frac{a-x}{a}\right)^2 + C_x \right), \qquad (2\text{-}2)$$

where C_x is a length constant to satisfy the condition that the optimal shape gives $z = 0$ at a reference point x_∞, F_b is a force scale related to the dissociation of single bonds (see (2-9) below), and $\varphi = F/(F_b N)$ is a load factor representing the nominal force sustained by individual bonds (normalized by the force scale F_b).

As can be seen in Figure 2, a shape factor S has been applied to the elasticity part of modeling in order to study the effects of surface shape on adhesion lifetime. The shape factor S is defined in such a way that $S = 0$ corresponds to two bonded surfaces in perfect conformity with each other in the absence of external loads, $S = 1$ corresponds to the optimal contact surface shape defined in (2-2), and $S = 2$ corresponds to an overly curved surface when compared to the optimal surface shape.

2.2. *Elasticity model.*

Let us define the axial direction running parallel to the surfaces of the cell and substrate to be x, and assume that the origin $x = 0$ corresponds to the center of the focal adhesion. Also, let us define $u(x)$ to be the difference between the stretched length of the ligand/receptor bond located at position x and its rest length l_b, and $w(x)$ be the surface displacement discontinuity of the two elastic media at x. The relationship between $u(x)$ and $w(x)$ is governed by the compatibility condition at the cell-substrate interface, to ensure that the bond length after deformation, $u(x) + l_b$, is equal to the displacement and shape discontinuities, $-w(x)$ and $S \cdot f_z(x)$, plus a constant separation h, i.e.,

$$u(x) + w(x) + l_b = h + S \cdot f_z(x). \qquad (2\text{-}3)$$

In our model, x_i is constricted to the domain of $[-a, a]$ and is representative of the location of a bond in the x direction within a focal adhesion of size $2a$. The elastic Green's function that defines the relative displacement of the combined elastic media at position x_i induced by a load being applied on a bond at x_j $(i \neq j)$ is given by the equation [Johnson 1985]

$$w_{ij} = \frac{1}{\pi E^* b} \cdot 2F_j \left(\ln |x_\infty - x_j| - \ln |x_i - x_j| \right), \qquad (2\text{-}4)$$

where F_j is the individual force acting on the bond at location x_j, and x_∞ is an arbitrary reference point chosen such that $w_{ij} = 0$ at x_∞ for a load F_j. To avoid singularity, the relative displacement of the elastic media at position x_i caused by the on-site load F_i acting on the bond is given by the equation [Johnson 1985]

$$w_{ii} = -\frac{1}{\pi E^*} \cdot \frac{F_i}{2a_0 b} (2a_0 \ln 4 + C_i), \qquad (2\text{-}5)$$

where a_0 denotes the radius of an individual bond and is typically only a few nanometers [Arnold et al. 2004], and C_i is a length constant chosen such that $w_{ii} = 0$ at x_∞.

Since we have chosen to model ligand/receptor bonds as Hookean springs, the relationship between the extension of a bond at a position x_i and the load F_i applied on that bond is

$$F_i = k_{LR} u_i. \tag{2-6}$$

Combining Equations (2-3)–(2-6) gives the elasticity equation

$$\sum_{j=1, j\neq i}^{n} \left\{ \frac{1}{\pi E^* b} 2F_j \left(\ln |x_\infty - x_j| - \ln |x_i - x_j| \right) \right\} - \frac{1}{\pi E^*} * \frac{F_i}{2a_0 b}(2a_0 \ln 4 + C_i) + \frac{F_i}{k_{LR}} + l_b$$
$$= h + S \cdot f_z(x_i), \quad (2\text{-}7)$$

where n is the number of closed bonds in the focal adhesion. This relationship in (2-7) represents n equations with $n+1$ unknowns $(F_1, \ldots, F_n, h)$. The elastic part of the model is completed with an additional equation coming from the global force balance

$$\sum_{i=1}^{n} F_i = F = \varphi F_b N, \tag{2-8}$$

where F is the constant tensile load being applied to the system.

2.3. Stochastic model. The elasticity equations shown above govern how the applied load F is distributed among all closed bonds. Each bond site within the focal adhesion in the present model has the ability to rupture from attachment as well as rebind after rupture. These rupturing and rebinding behaviors are represented as rate equations and are defined using the following stochastic equations.

Bond rupturing is assumed to occur according to [Bell 1978]

$$k_{\text{off}} = k_0 e^{F_i/F_b}, \tag{2-9}$$

where the rate of bond rupture increases exponentially as the load acting on the individual bonds F_i increases. The rate of bond rupture as defined by Bell involves the parameters F_b, which is a force scale typically in the piconewton range [Bell 1978], and k_0, which is the spontaneous dissociation rate of an individual bond in the absence of any type of load. The spontaneous dissociation rate can range anywhere from 0.1 to 100 seconds depending on the bond type and environmental conditions [Evans and Calderwood 2007].

Bond rebinding behavior is assumed to depend on the separation distance between the open receptor and free hanging ligand as described by Erdmann and Schwarz [2007; 2006], which the reader can consult for the next few equations. The probability of a ligand falling in close proximity to a receptor within a reacting radius l_{bind} is

$$p = \frac{l_{\text{bind}}}{Z} e^{-k_{LR}(\delta - l_b)^2/(2k_B T)}, \tag{2-10}$$

where l_{bind} is defined as the length scale within which the receptor and ligand will react to form a bond, δ is the surface separation between cell and substrate where the open receptor and ligand are anchored,

$k_B T$ is the Boltzmann constant multiplied by the absolute temperature, and Z is a partition function defined as

$$Z = \sqrt{\frac{\pi k_B T}{2 k_{\text{LR}}}} \left\{ \text{erf}\left((\delta - l_b) \sqrt{\frac{k_{\text{LR}}}{2 k_B T}} \right) + \text{erf}\left(l_b \sqrt{\frac{k_{\text{LR}}}{2 k_B T}} \right) \right\}. \tag{2-11}$$

From these equations, the rebinding rate can be defined as

$$k_{\text{on}} = k_{\text{on}}^0 \frac{l_{\text{bind}}}{Z} e^{-k_{\text{LR}}(\delta - l_b)^2 / (2 k_B T)}, \tag{2-12}$$

where k_{on}^0 is a spontaneous reaction rate of open ligand/receptor pairs within the reacting radius l_{bind}. It is beneficial to normalize both the rupture and rebinding rates in our model in order to define certain analytical factors. Both the rupture and rebinding rates are defined by

$$r_i = e^{F_i / F_b}, \tag{2-13}$$

$$g_i = 2\gamma \sqrt{\frac{\beta}{\pi}} \frac{e^{-\beta(\Delta_i - L_b)^2}}{\text{erf}\left(\sqrt{\beta}(\Delta_i - L_b) \right) + \text{erf}\left(L_b \sqrt{\beta} \right)}, \tag{2-14}$$

where $\Delta_i = \delta_i / b$ is the normalized bond separation, $L_b = l_b / b$ is the normalized bond rest length, and

$$\beta = \frac{k_{\text{LR}} b^2}{2 k_B T}. \tag{2-15}$$

In the bond rebinding rate equation, γ is a rebinding factor defined as

$$\gamma = \frac{k_{\text{on}}^0}{k_0} \frac{l_{\text{bind}}}{b}, \tag{2-16}$$

which is a measure of propensity for rebinding to occur. Once the separation length δ has reached a value of roughly $2l_b$, rebinding of the ligand/receptor becomes almost impossible no matter how great the rebinding factor is [Qian et al. 2008].

In our model we consider both stochastic rebinding and rupturing of all bonds within a focal adhesion. The binding evolution of the bond cluster is characterized by the master equation [van Kampen 1992]

$$\frac{dP_n(\tau)}{d\tau} = \hat{g}_{n-1} P_{n-1}(\tau) + \hat{r}_{n+1}(\tau) P_{n+1}(\tau) - \left[\hat{r}_n(\tau) + \hat{g}_n(\tau) \right] P_n(\tau), \tag{2-17}$$

where $\tau = k_0 t$ is normalized time, n is the number of closed bonds in the adhesion, $\hat{r}_n$ and $\hat{g}_n$ correspond to the normalized total bond cluster rupture and rebinding rates when n bonds are closed (these rates differ from r_i and g_i which are rupture and rebinding rates for a single bond i), and $P_n(\tau)$ is the probability that n bonds are closed at time τ. Using a Monte Carlo method, which will be discussed in the next section, our model initially starts with n equal to the total number of bonds N within the focal adhesion. As time passes, stochastic binding and breaking occur until n equals zero, corresponding to failure of the adhesion as all the bonds within the focal adhesion become open.

3. Numerical method

To analyze (2-17) we use a Monte Carlo method based on Gillespie's algorithm [1976; 1977]. Initially all bonds within the focal adhesion are assumed to be closed. The evolution of the adhesion is characterized

total number of bonds, N	6–90
spacing between neighboring bonds, b	32 nm
focal adhesion size, $2a$	~ 0.2–$3\,\mu$m
bond rest length, l_b	11 nm
radius of bond, a_0	5 nm
single bond stiffness, k_{LR}	0.25 pN/nm
combined elastic modulus, E^*	2–400 kPa
rebinding rate factor, γ	100
load factor, φ	0.44–0.52
force scale in bond dissociation, F_b	4 pN
shape factor, S	0–2
Boltzmann's constant $\times$ temperature, $k_B T$	4.1 pN·nm

Table 1. Parameters used in the model of a focal adhesion between a cell and a substrate.

by cycles in which bond rupturing or rebinding occurs. Whether a closed bond will rupture or an open bond will rebind is determined by the probability of these actions occurring. The idea is to determine using randomized numbers and probability the likely location of a particular bond rupturing or rebinding for each cycle during the focal adhesion's lifetime until every bond in the focal adhesion is open. The likelihood of bond rupture/rebinding to occur at a specific location becomes dependent on the elastic solutions of the model at that current cycle time, and is globally based on the parameters used to define the specific focal adhesion under study, which are listed for our model in Table 1. It is necessary to note that because of the stochastic behaviors of both bond rupturing and rebinding for the focal adhesion, many simulations of each particular parameter set of focal adhesion must be performed and their output results averaged in order to formulate a reasonable assessment representative of the behavior of the adhesion.

In the present study, we analyze our model using the "first reaction method" of Gillespie's algorithm, following [Erdmann and Schwarz 2004a; 2004b]. In the first reaction method, a series of independent numbers are randomly generated in which they are uniformly distributed from 0 to 1. Each random number ξ_i is assigned to a specific bond site such that $\xi = (\xi_1, \ldots, \xi_N)$ where N is the total number of bonds in the focal adhesion. The normalized reaction time for each specific bond site is calculated as [Gillespie 1976; 1977].

$$d\tau_i = -\frac{\ln \xi_i}{q_i}, \tag{3-1}$$

where q_i is either the normalized rebinding rate if the bond classified by index i is open or the normalized rupture rate if the bond is closed. The normalized reaction time for the particular cycle is chosen to be the minimum component of $d\tau_i$ such that,

$$d\tau = \min(d\tau_i), \tag{3-2}$$

where the location of actual reaction taking place is identified as where the shortest reaction time is chosen. This means that if the chosen bond at the reaction location is currently closed, the bond will rupture, and if the chosen bond is currently open, it will rebind, and the normalized time between the

reaction just prior to this one and the reaction just after this one is $d\tau$. Each time this reaction occurs constitutes one cycle and the number of cycles continues to grow until all bonds in the focal adhesion are open, which we have defined as the point when adhesion fails. The normalized lifetime of the focal adhesion is defined as the sum of all the reaction times for each cycle,

$$\tau = \sum_{k=1}^{\mu} d\tau_k,$$

(3-3)

where μ is the total number of cycles leading up to adhesion failure. This process is repeated a prescribed number of times and a normalized mean lifetime is determined.

4. Results

The first class of focal adhesions we chose to model were three types of focal adhesions with varying combined elastic modulus E^* values of 200 kPa, 100 kPa, and 20 kPa. For the model we chose a rebinding factor $\gamma = 100$ which is representative of highly favorable rebinding conditions [Qian et al. 2008; 2009], and a load factor $\varphi = 0.46$ which corresponds to the applied tensile load on the system being $\rho = 1.78$ kPa where $\rho = F/2ab$ and $F = \varphi F_b N$. The model is defined to represent a focal adhesion being experimented on at room temperature, corresponding to a value of $k_B T = 4.1$ pN · nm. Each ligand/receptor bond of the adhesion was modeled as a Hookean spring with a spring coefficient of $k_{\mathrm{LR}} = 0.25$ pN/nm, which has been estimated as an appropriate value when modeling fibronectin/integrin bonds [Erdmann and Schwarz 2006]. The force factor defined for bond rupturing was chosen to be a constant value $F_b = 4$ pN [Erdmann and Schwarz 2004a; 2004b]. The ligand/receptor dimensions were defined to have a rest length $l_b = 11$ nm [Zuckerman and Bruinsma 1998] and an individual radius of $a_0 = 5$ nm [Arnold et al. 2004]. Each ligand/receptor bond is evenly spaced a distance $b = 32$ nm from each other, with the length of the entire focal adhesion defined as $2a$ and the total number of bonds in the adhesion as $N = 2a/b$. As shown in the left column of Figure 3, for each value of E^*, we varied the number of bonds making up the focal adhesion from 6 to 90, and for each focal adhesion size we varied the shape factor S within $[0,2]$ in order to analyze each focal adhesion's mean lifetime. This range in number of bonds corresponds to a focal adhesion size of roughly 200 nm - 3 μm.

Figure 3, top left, shows a clear increase in adhesion lifetime as the contact surface shape progresses towards the optimal shape, especially as the focal adhesion increases in size. However, as the adhesion increases in size, the optimal shape remains to have the maximum adhesion lifetime, but any discrepancies from this optimal shape result in a much lower adhesion lifetime. This is most clearly seen in Figure 3, top right, in which the adhesion lifetime is plotted against the number of bonds for shape factor values $S = 0, 0.8, 1$. The optimal shape represented in the graph by $S = 1$ clearly has the longest adhesion lifetime. However, as the adhesion grows in size, subtle variance from the optimal surface shape results in a drastic change in adhesion lifetime, which can be seen by comparing the lifetimes of $S = 0.8$ and $S = 1$. The difference between the two focal adhesion shapes is geometrically subtle, but the effect it has on adhesion lifetime becomes very drastic as the focal adhesion grows in size. These shape effects have also been studied in other areas of biology with regards to gecko adhesion, and similar results pertaining to adhesion strength were seen [Gao and Yao 2004]. A significant difference in adhesion lifetime appears to occur once the focal adhesion expands to 50 bonds when comparing the two surface shapes. This
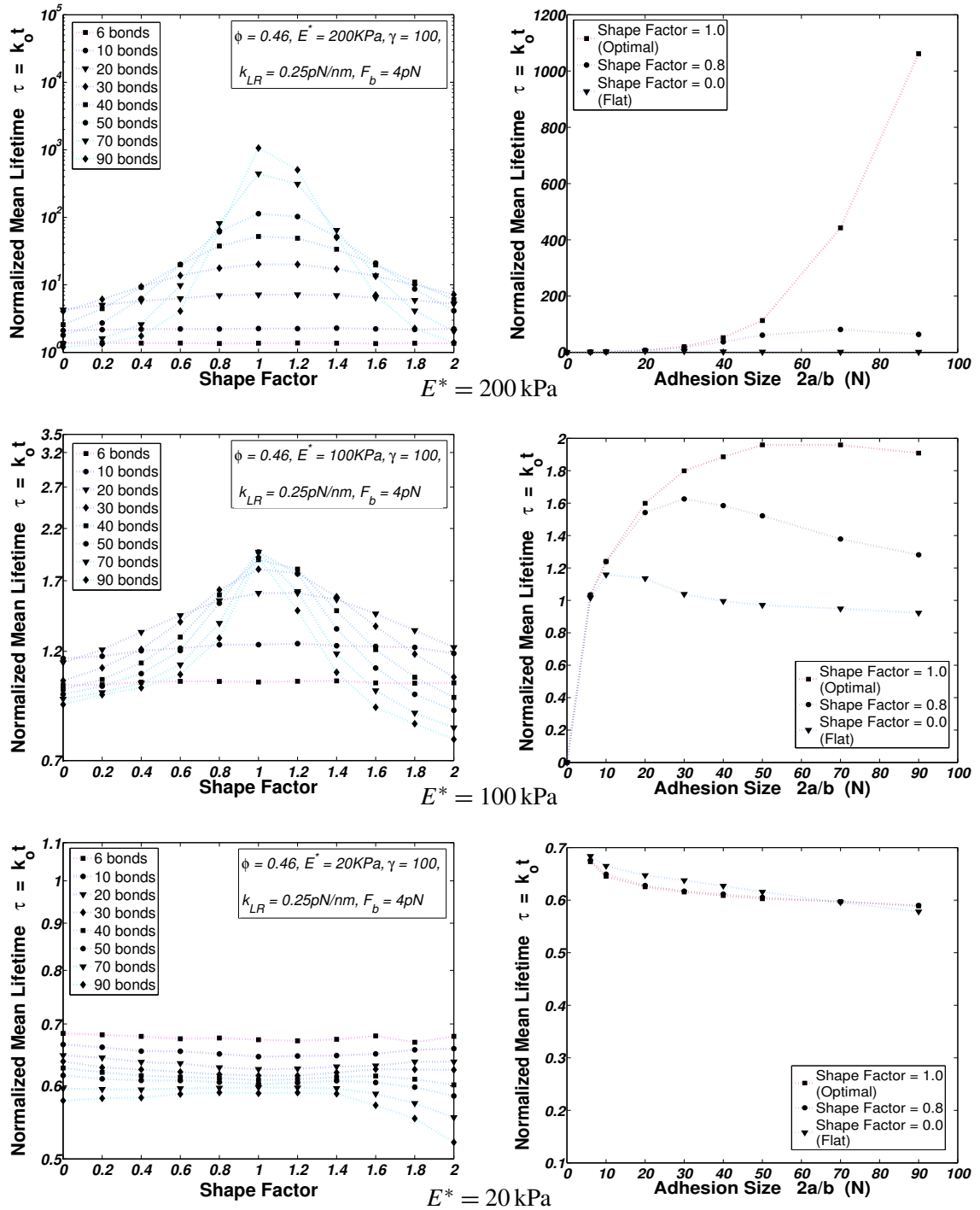
Figure 3. Effects of varying surface shape as defined by numerical shape factor values ranging within [0,2] on focal adhesion lifetime for different focal adhesion sizes ranging from 6 to 90 individual bonds. From top to bottom, $E^* = 200\,\text{kPa}$, $100\,\text{kPa}$, $20\,\text{kPa}$.

difference in lifetime increases more drastically as the adhesion continues to grow in size. This drastic reduction in adhesion lifetime for larger bonds with varying adhesion surface geometry from the optimal shape can be understood by the fact that generally stress concentrations are more severe as the focal adhesion increases in size [Qian et al. 2008; 2009]. What the optimal shape does for adhesion lifetime is to remove the preexisting stress concentrations from the bond cluster at its initial state. Therefore, the optimal contact shape can improve the lifetime of a cluster of bonds more significantly in larger adhesions.

The next type of focal adhesion modeled can be seen in Figure 3, middle left, in which the number of bonds within the adhesion varies from 6 to 90, the shape factor's observed range in [0, 2], and the combined elastic modulus $E^* = 100$ kPa. Comparing the focal adhesion lifetimes between $E^* = 200$ kPa and $E^* = 100$ kPa shows that the adhesion lifetimes are much smaller for the softer system. This is concurrent to the results shown in other studies of system stiffness and adhesion lifetime [Qian et al. 2008; 2009]. Besides this difference in magnitude of adhesion lifetime, the behavior of the data in the case of $E^* = 100$ kPa seems to differ when compared to the results shown in Figure 3, top left. The focal adhesion still has the longest lifetime on the optimal contact surface shape, but the largest focal adhesion size does not correspond to the longest adhesion lifetime. This is more clearly seen in Figure 3, middle right, where the maximum adhesion lifetime is achieved for the optimal shape $S = 1$, when the adhesion contains about 50 bonds. After this critical size, the lifetime of the focal adhesion appears to shorten as the number of bonds increases. Similar behavior occurs for $S = 0.8$ where the critical number of bonds is 30, and for the flat surface $S = 0$ where the critical number of bonds is 10. A possible reason for this behavior, which has been previously studied for flat surface adhesion [Qian et al. 2008; 2009] is that as focal adhesions increase in size, they develop large stress concentrations at edges of the focal adhesion, and these stress concentrations result in crack-like failure of the focal adhesion, thereby decreasing its lifetime. Since the optimal shape was defined as the cell surface shape which results in the initially even distribution of load among all closed bonds, as the focal adhesion continues to cycle in which bonds are rupturing and rebinding, we believe that crack-like singularities begin to form within the adhesion domain and eventually dominate adhesion lifetime. This behavior of decreased adhesion lifetime with increasing adhesion size is also seen in the top row of Figure 3 for $E^* = 200$ kPa for the majority of nonoptimal cases, and we believe if the adhesion size continues to increase in bond number this behavior would then be seen even in the optimal case, as has been displayed for $E^* = 100$ kPa.

The bottom row of Figure 3 shows the results for a focal adhesion with combined elastic modulus $E^* = 20$ kPa. The data acquired for this focal adhesion differs greatly in comparison to the $E^* = 200$ kPa, 100 kPa cases mainly because the longest adhesion lifetimes seem to occur when the adhesion is very small. Also, the longest adhesion lifetime for most of the adhesion sizes seems to occur when the geometry of the adhesive surface is flat. Not until the adhesion reaches a size of roughly 70 bonds does the behavior of the focal adhesion lifetime seem to resemble that of the behavior shown in the above two cases. However, even in the case of optimal shape involving 90 bonds where the adhesion lifetime is longest, it still is shorter when compared to the adhesion lifetimes for any adhesion size and shape factor analyzed. From this data it appears that once the system reaches a critical stiffness threshold between soft matter and very soft matter, the effects of bond size on the adhesion outweigh the potential strengthening effects of shape on the adhesion. It also appears that optimizing the shape seems to decrease adhesion lifetime slightly, which differs in behavior from the previous two cases. However, the reasoning for these
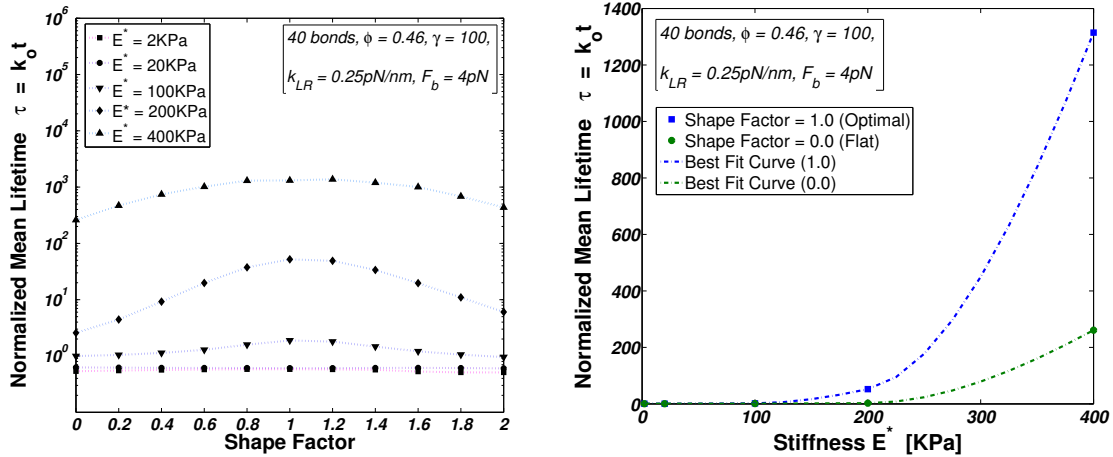
Figure 4. The effects of varying system stiffness as defined by the combined elastic modulus E^* on focal adhesion lifetime for an adhesion made up of 40 bonds with a load factor of $\varphi = 0.46$. Left: focal adhesion lifetime as a function of the surface shape factor for values of $E^* = 2, 20, 100, 200, 400$ kPa. Right: focal adhesion lifetime as a function of system stiffness for the flat surface case ($S = 0$) and the optimal surface case ($S = 1$).

unexpected results could be because the load applied to the adhesion defined by a load factor of $\varphi = 0.46$ is too great. The range in mean normalized lifetimes of all the adhesion sizes for the $E^* = 20$ kPa case only extends a normalized time frame of 0.1 which, when compared to the cases of $E^* = 100$ kPa and 200 kPa, is roughly 10 and 10,000 times smaller. In order to get a more thorough analysis of the effects of shape on a system of very soft stiffness, the load factor would have to be reduced in such a way that from the data shown here, a normalized time range of at least 1 would have to be represented. This would allow for a better representation of the effects of shape on adhesion systems with very soft cell and substrate.

For further understanding on the effects of contact surface shape on adhesion lifetime, an adhesion made up of 40 bonds was analyzed under varying stiffness values and varying load factor values. In Figure 4, a 40-bond focal adhesion was analyzed with a load factor of $\varphi = 0.46$, a rebinding factor of $\gamma = 100$, and the combined elastic stiffness of the system ranged $E^* = 2$ kPa, 20 kPa, 100 kPa, 200 kPa, 400 kPa. From the left part of the figure, we see that the adhesion with optimal surface shape has the longest adhesion lifetime for the cases of $E^* = 100$ kPa and greater, and it appears that the adhesion lifetime gets exponentially longer as the stiffness of the system doubles in magnitude. Also as expected, due to the medium focal adhesion size, the difference between slight deviations from the optimal shape does not seem to have the drastic effect it had for larger bond sizes. However, it can be seen in Figure 4, right, that there is a large difference between adhesion lifetime between the optimal surface shape and the flat surface shape. This difference is large enough such that any attempt of optimizing the shape of the adhesion would have a substantial increase in the lengthening of adhesion lifetime, especially in cases where the system stiffness is large.
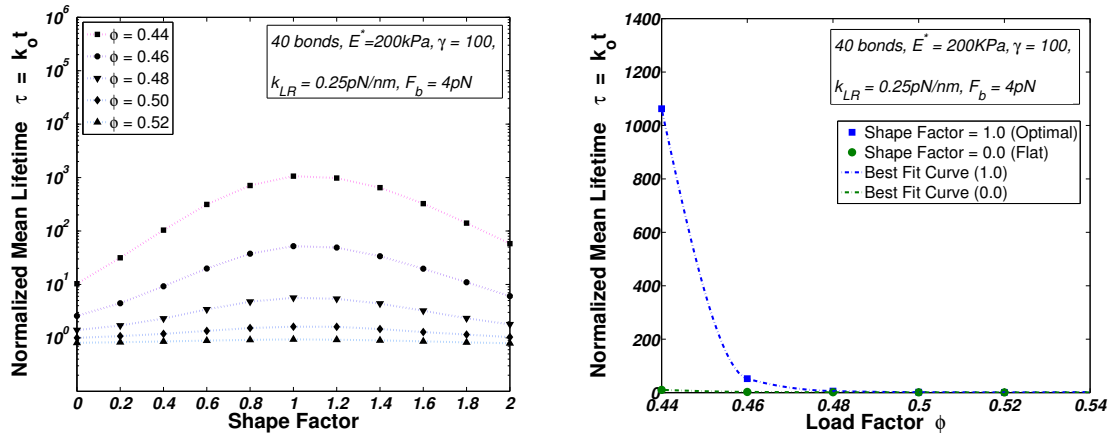
Figure 5. The effects of varying applied tensile load acting on the cell/substrate as defined by the load factor φ on focal adhesion lifetime for an adhesion made up of 40 bonds with a combined elastic modulus of $E^* = 200$ kPa. Left: focal adhesion lifetime as a function of the surface shape factor for values of $\varphi = 0.52, 0.50, 0.48, 0.46, 0.44$. Right: focal adhesion lifetime as a function of the applied load for the flat surface case $(S = 0)$ and the optimal surface case $(S = 1)$.

In Figure 5, a similar 40-bond focal adhesion is considered with a combined elastic modulus of $E^* = 200$ kPa, a rebinding factor of $\gamma = 100$, and the load factor for the adhesion is varied such that $\varphi = 0.52, 0.50, 0.48, 0.46, 0.44$. As with the previous model involving varying system stiffness, the focal adhesion with the optimal surface shape has the longest adhesion lifetime and the difference in lifetime appears to increase exponentially as the load decreases, which is to be expected. The effects the shape of the adhesion surface has on adhesion lifetime seem to be more substantially affected by decreasing the load factor than varying the stiffness, which is shown by the greater normalized lifetime difference between the optimal and flat surface cases displayed in Figure 5, right. However there appears to be substantially more sensitivity to deviations from the optimal shape as the load decreases in magnitude in comparison to the effects of deviation from the optimal shape for the increasing stiffness case. This can be taken to mean that decreasing the applied tensile load on an adhesion of large size with some geometrical variance from the optimal shape would be less effective in increasing the focal adhesion lifetime than would be done by increasing the system stiffness.

5. Conclusion

In this paper we have examined an idealized model of a cell/substrate focal adhesion bonded together by ligand/receptor bonds evenly spaced along curved surfaces. The focal adhesion was modeled as a bond cluster between continuous elastic media undergoing an applied tensile load which is distributed as traction forces on each bond. Initially all bonds in the adhesion were considered to be closed and the magnitude of the distributed traction forces on individual bonds was determined by the geometrical shape of the contact surface. The contact surface shape was defined by a numeric shape factor S in proportion to

an optimal shape that renders the applied tensile load being uniformly distributed among all closed bonds whenever these bonds are evenly spaced within the adhesion domain. The model accounted for both bond rupturing and rebinding via stochastic equations, which were analyzed numerically using Monte Carlo methods. A series of simulations were performed using a coupling of elastic and stochastic equations in order to model several sets of focal adhesions in which the effects of contact surface shape were observed with respect to the adhesion lifetime. In our model we have chosen to define adhesion lifetime as the time taken for all bonds within the adhesion to go from their initial state to failure. The calculated adhesion lifetimes were normalized using k_0, the spontaneous dissociation rate of an individual bond in the absence of any type of load. The following main conclusions can be drawn from our analysis.

(1) As the contact surface shape begins to change from flat towards the optimal shape for adhesion, the lifetime of the focal adhesion increases to a maximum value. The optimal surface shape is defined to have a shape factor of $S = 1$. Depending on the size of the focal adhesion and the number of ligand/receptor bonds within the adhesion, the advantage of having an optimal surface shape in increasing adhesion lifetime becomes more significant as the adhesion increases in size.

(2) The larger the adhesion size becomes, the more drastic the effect slight deviations from the optimal contact surface shape have on the focal adhesion lifetime. Similar behavior has been demonstrated for adhesion of elastic media via van der Waals forces [Gao and Yao 2004; Yao and Gao 2006], as well as vesicles adhering on curved substrates [Shi et al. 2006].

(3) As the size of a focal adhesion increases, the effect of the increased adhesion lifetime due to optimal surface shape begins to be overtaken by crack-like singularity behaviors which result in earlier failure of the focal adhesion. This transition of adhesion behavior leads to the observation of a critical adhesion size that corresponds to a maximum adhesion lifetime for the prescribed focal adhesion parameters. Even though the optimal surface shape almost always corresponds to the maximum adhesion lifetime (except in the case of very soft matter), as the focal adhesion increases in size, the adhesion lifetime begins to decrease after extending past a critical focal adhesion size. This is consistent with the behavior that occurs with flat surface focal adhesion as it increases in size [Qian et al. 2008; 2009].

(4) In comparing the cases of increasing system stiffness or decreasing the magnitude of the applied tensile load in order to increase focal adhesion lifetime, it seems that both cases amplify the effects of optimal shape by a greater magnitude in comparison to the nonoptimized flat adhesion surface orientation. However, decreasing the applied load seems to be less effective in mitigating the effects of reduced adhesion lifetime due to deviations from the optimal surface shape in comparison to increasing system stiffness for a single focal adhesion. It is worthwhile to note that the sensitivity of focal adhesions to cell/substrate stiffness cannot be alleviated by removing stress concentration through optimal shape design. Even for molecular clusters under initially uniform pulling forces, the cell/substrate elasticity can destabilize focal adhesions by suppressing rebinding of open bonds [Qian and Gao 2010].

Experiments have shown that nanostructured surfaces have great impacts on cellular behavior and tissue engineering. The concepts developed in this paper should be understood as a theoretical model displaying the effects of surface geometry on adhesion lifetime for molecular adhesion based on specific ligand/receptor bonds. The present study shows from Monte Carlo simulations of discrete bond evolution that the contact surface shape could greatly influence the stability of cell-matrix adhesion. The analysis

by Wang and Gao [2010] based on a continuum formulation of cell-substrate interaction led to similar shape-dependent adhesion. This principle of optimal adhesion shape has already been quantitatively demonstrated by experiments in the case of nonspecific adhesion [Waters et al. 2011] and is now being analyzed and studied for specific adhesion as well. Once proven, we could contemplate further controlled experiments to explore using the micro- and nanoscale surface geometry to probe cell-matrix interactions.

References

[Alberts et al. 2002] B. Alberts, A. Johnson, J. Lewis, M. Raff, R. Keith, and P. Walter, *Molecular biology of the cell*, 4th ed., Garland Science, New York, 2002.

[Arnold et al. 2004] M. Arnold, E. A. Cavalcanti-Adam, R. Glass, J. Blümmel, W. Eck, M. Kantlehner, H. Kessler, and J. P. Spatz, "Activation of integrin function by nanopatterned adhesive interfaces", *ChemPhysChem* **5**:3 (2004), 383–388.

[Bell 1978] G. I. Bell, "Models for the specific adhesion of cells to cells", *Science* **200**:4342 (1978), 618–627.

[Bershadsky et al. 2003] A. D. Bershadsky, N. Q. Balaban, and B. Geiger, "Adhesion-dependent cell mechanosensitivity", *Annu. Rev. Cell Dev. Biol.* **19** (2003), 677–695.

[Chen and Gumbiner 2006] X. Chen and B. M. Gumbiner, "Crosstalk between different adhesion molecules", *Curr. Opin. Cell Biol.* **18**:5 (2006), 572–578.

[Erdmann and Schwarz 2004a] T. Erdmann and U. S. Schwarz, "Stability of adhesion clusters under constant force", *Phys. Rev. Lett.* **92**:10 (2004), 108102.

[Erdmann and Schwarz 2004b] T. Erdmann and U. S. Schwarz, "Stochastic dynamics of adhesion clusters under shared constant force and with rebinding", *J. Chem. Phys.* **121**:18 (2004), 8997–9017.

[Erdmann and Schwarz 2006] T. Erdmann and U. S. Schwarz, "Bistability of cell-matrix adhesions resulting from nonlinear receptor-ligand dynamics", *Biophys. J.* **91**:6 (2006), L60–L62.

[Erdmann and Schwarz 2007] T. Erdmann and U. S. Schwarz, "Impact of receptor-ligand distance on adhesion cluster stability", *Eur. Phys. J. E* **22**:2 (2007), 123–137.

[Evans 2001] E. A. Evans, "Probing the relation between force—lifetime—and chemistry in single molecular bonds", *Annu. Rev. Biophys. Biomol. Struct.* **30** (2001), 105–128.

[Evans and Calderwood 2007] E. A. Evans and D. A. Calderwood, "Forces and bond dynamics in cell adhesion", *Science* **316**:5828 (2007), 1148–1153.

[Evans and Ritchie 1997] E. A. Evans and K. Ritchie, "Dynamic strength of molecular adhesion bonds", *Biophys. J.* **72**:4 (1997), 1541–1555.

[Gao and Yao 2004] H. Gao and H. Yao, "Shape insensitive optimal adhesion of nanoscale fibrillar structures", *Proc. Nat. Acad. Sci. USA* **101**:21 (2004), 7851–7856.

[Gillespie 1976] D. T. Gillespie, "A general method for numerically simulating the stochastic time evolution of coupled chemical reactions", *J. Comput. Phys.* **22**:4 (1976), 403–434.

[Gillespie 1977] D. T. Gillespie, "Exact stochastic simulation of coupled chemical reactions", *J. Phys. Chem.* **81**:25 (1977), 2340–2361.

[Johnson 1985] K. L. Johnson (editor), *Contact mechanics*, Cambridge University Press, New York, 1985.

[van Kampen 1992] N. G. van Kampen, *Stochastic processes in physics and chemistry*, North-Holland, New York, 1992.

[Lo et al. 2000] C.-M. Lo, H.-B. Wang, M. Dembo, and Y.-L. Wang, "Cell movement is guided by the rigidity of the substrate", *Biophys. J.* **79**:1 (2000), 144–152.

[Pelham and Wang 1997] R. J. Pelham, Jr. and Y.-L. Wang, "Cell locomotion and focal adhesions are regulated by substrate flexibility", *Proc. Nat. Acad. Sci. USA* **94**:25 (1997), 13661–13665.

[Qian and Gao 2010] J. Qian and H. Gao, "Soft matrices suppress cooperative behaviors among receptor-ligand bonds in cell adhesion", *PLoS ONE* **5**:8 (2010), e12342.

[Qian et al. 2008] J. Qian, J. Wang, and H. Gao, "Lifetime and strength of adhesive molecular bond clusters between elastic media", *Langmuir* **24**:4 (2008), 1262–1270.

[Qian et al. 2009] J. Qian, J. Wang, Y. Lin, and H. Gao, "Lifetime and strength of periodic bond clusters between elastic media under inclined loading", *Biophys. J.* **97**:9 (2009), 2438–2445.

[Shi et al. 2006] W. Shi, X. Q. Feng, and H. Gao, "Two-dimensional model of vesicle adhesion on curved substrates", *Acta Mech. Sinica* **22**:6 (2006), 529–535.

[Tan et al. 2003] J. L. Tan, J. Tien, D. M. Pirone, D. S. Gray, K. Bhadriraju, and C. S. Chen, "Cells lying on a bed of microneedles: an approach to isolate mechanical force", *Proc. Nat. Acad. Sci. USA* **100**:4 (2003), 1484–1489.

[Wang and Gao 2010] J. Wang and H. Gao, "Size and shape dependent steady-state pull-off force in molecular adhesion between soft elastic materials", *Int. J. Fract.* **166**:1–2 (2010), 13–19.

[Waters et al. 2011] J. F. Waters, H. J. Gao, and P. R. Guduru, "On adhesion enhancement due to concave surface geometries", *J. Adhes.* **87**:3 (2011), 194–213.

[Yao and Gao 2006] H. Yao and H. Gao, "Optimal shapes for adhesive binding between two elastic bodies", *J. Colloid Interface Sci.* **298**:2 (2006), 564–572.

[Zaidel-Bar et al. 2003] R. Zaidel-Bar, C. Ballestrem, Z. Kam, and B. Geiger, "Early molecular events in the assembly of matrix adhesions at the leading edge of migrating cells", *J. Cell Sci.* **116**:22 (2003), 4605–4613.

[Zamir et al. 2000] E. Zamir, M. Katz, Y. Posen, N. Erez, K. M. Yamada, B.-Z. Katz, S. Lin, D. C. Lin, A. Bershadsky, Z. Kam, and B. Geiger, "Dynamics and segregation of cell-matrix adhesions in cultured fibroblasts", *Nat. Cell Biol.* **2**:4 (2000), 191–196.

[Zuckerman and Bruinsma 1998] D. M. Zuckerman and R. F. Bruinsma, "Vesicle-vesicle adhesion by mobile lock-and-key molecules: Debye–Hückel theory and Monte Carlo simulation", *Phys. Rev. E* **57**:1 (1998), 964–977.

GREGORY J. RIZZA: gregory_rizza@brown.edu
Solid Mechanics Group, School of Engineering, Brown University, Providence, RI 02912, United States

JIN QIAN: jqian9@mail.gatech.edu
Coulter Department of Biomedical Engineering, Georgia Institute of Technology, Atlanta, GA 30332, United States

HUAJIAN GAO: huajian_gao@brown.edu
Solid Mechanics Group, School of Engineering, Brown University, Providence, RI 02912, United States

LONG WAVELENGTH BIFURCATIONS AND MULTIPLE NEUTRAL AXES OF ELASTIC LAYERED STRUCTURES SUBJECT TO FINITE BENDING

SARA ROCCABIANCA, DAVIDE BIGONI AND MASSIMILIANO GEI

Dedicated to Marie-Louise and Charles Steele

Geometries and rigidities involving the presence of more than one neutral axis during finite (plane-strain) bending of a multilayered elastic (incompressible) block make numerically stiff the differential equations governing the incremental problem necessary to investigate diffuse-mode instabilities. We have developed a compound matrix method to solve these cases, so that we have shown that the presence of two neutral axes occurs within sets of parameters where the elastic system may display long-wavelength bifurcation modes. Following the predictions of the theory, we have designed and realized qualitative experiments in which these modes become visible.

1. Introduction

Recently Roccabianca et al. [2010] have addressed, solved and experimented the problem of bifurcation of a layered incompressible elastic block, subject to finite bending in plane strain. They were able to show that a layered system exhibits bifurcation loads and angles of bending completely different from those occurring in a uniform elastic block. They found that for several geometries and stiffness contrasts the first ("critical") bifurcation load corresponds to a long-wavelength mode, which results to be very close to the bifurcation load associated with the surface instability limit of vanishing wavelength,[1] a feature also common to the behavior of a uniform elastic block. This feature explains the experimental observation (on uniform blocks [Gent and Cho 1999; Gent 2005] and bilayers [Roccabianca et al. 2010]) that short-wavelength modes become visible, instead of the long-wavelength modes that are predicted to occur before. Therefore, the question was left open whether or not wavelength modes longer than the short-wavelength modes available at surface instability and visible in the experiments can be experimentally displayed with a layered system in which an appropriate selection is made of stiffness and thickness contrasts between layers. We provide a positive answer to this problem in the present article, so that our calculations, based now on the compound matrix method [Backus and Gilbert 1967; Ng and Reid 1979a;

[1] Surface instability occurs in a uniformly strained half space as a bifurcated mode of arbitrary wavelength, corresponding to a Rayleigh wave of vanishing speed. In the limit of vanishing wavelength, surface instability can be viewed as a bifurcation mode adaptable to every boundary and state of stress of a strained body, so that it becomes a local instability mode (also called "failure of the complementing condition" in [Benallal et al. 1993]).
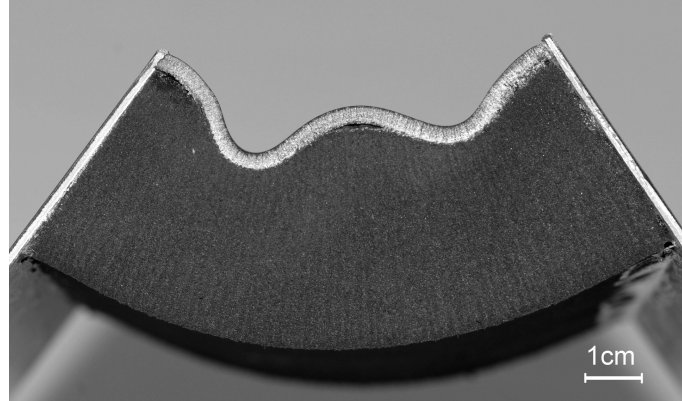
Figure 1. Bifurcation of a two-layer rubber block under finite bending evidencing long-wavelength bifurcation modes. Stiffness and thickness ratios between the layers are $(2.687 \, \text{N/mm}^2)/(0.095 \, \text{N/mm}^2)$ and $(3 \, \text{mm})/(40 \, \text{mm})$. The stiff layer ($86 \, \text{mm} \times 3 \, \text{mm} \times 150 \, \text{mm}$, made up of natural rubber, marked with a white pencil on the sample) is at the compressive side and coats a neoprene layer ($86 \, \text{mm} \times 40 \, \text{mm} \times 150 \, \text{mm}$).

1979b; 1985], allow us to conclude that there are situations in which the long-wavelength modes are well-separated from the surface instability, so that systems exhibiting bifurcation modes of long wavelength can be designed. These systems have been realized by us and qualitatively tested, showing that the theory predictions are generally followed, Figure 1.

It turns out that the hunt for long-wavelength modes can be related[2] to the presence of more than one neutral axis (namely, the line of null normal stresses) during bending, an issue that was noted in [Roccabianca et al. 2010], but not investigated in detail.

In a bilayer, two neutral axes typically occur when a stiff layer is placed at the compressive side of the system, a case in which the differential equations become numerically stiff, so that we have employed an ad hoc version of the compound matrix method, which is shown to allow systematic investigation of the situations in which more than one neutral axis occurs. In these cases we find a sort of inversion of the sequence of bifurcation modes with the aspect ratio of the system, so that high-wavenumber modes are relevant for lower slender ratios than small-wavenumber modes.

2. Finite bending of an elastic multilayer and incremental bifurcations

We review the theory developed in [Roccabianca et al. 2010] for finite bending of an elastic incompressible multilayered block and for incremental bifurcations, both addressed here only in the special case of a Mooney–Rivlin material,[3] a choice which affects results, but simplifies the treatment.

We denote a generic layer with a superscript (s) $(s = 1, \dots, N)$, so that, in the reference stress-free configuration, a Cartesian coordinate system $O_0^{(s)} x_1^{0(s)} x_2^{0(s)} x_3^{0(s)}$ is introduced for each layer, centered at

---

[2] The relation is that within the parameter range in which long-wavelength bifurcations occur well-separated from surface instability, two neutral axes are often found.

[3] Bending of an elastic homogeneous block has been solved in [Rivlin 1949], while incremental bifurcations have been analyzed in [Triantafyllidis 1980; Dryburgh and Ogden 1999; Coman and Destrade 2008; Destrade et al. 2009; 2010].

its centroid, with basis vectors e_i^0 ($i = 1, 2, 3$), $x_1^{0(s)} \in [-h_0^{(s)}/2, h_0^{(s)}/2]$, $x_2^{0(s)} \in [-l_0/2, l_0/2]$, and with $x_3^{0(s)}$ denoting the out-of-plane coordinate.

The deformed configuration of each layer is a sector of a cylindrical tube of half-angle $\bar{\theta}$, so that it becomes convenient to introduce a cylindrical coordinate system $O^{(s)}r^{(s)}\theta^{(s)}z^{(s)}$, with basis vectors e_r, e_θ and e_z, $r^{(s)} \in [r_i^{(s)}, r_i^{(s)} + h^{(s)}]$, $\theta^{(s)} \in [-\bar{\theta}, +\bar{\theta}]$, and with out-of-plane coordinate $z^{(s)}$.

The imposition of incompressibility (conservation of volume) yields

$$r_i^{(s)} = \frac{l_0 h_0^{(s)}}{2\bar{\theta} h^{(s)}} - \frac{h^{(s)}}{2}, \tag{1}$$

where $h^{(s)}$ is the current thickness of the s–th layer. Following [Roccabianca et al. 2010], the imposition of the incompressibility constraint in terms of principal stretches (λ_r, λ_θ and λ_z) and of the boundary conditions at $x_2^{0(s)} = \pm l_0/2$ and $x_1^{0(s)} = -h_0^{(s)}/2$ provides

$$\lambda_r^{(s)} = \frac{1}{\alpha r^{(s)}}, \quad \lambda_\theta^{(s)} = \alpha r^{(s)}, \quad \lambda_z^{(s)} = 1, \tag{2}$$

where $\alpha = 2\bar{\theta}/l_0$; note that α is independent of index s.

The N layers forming the multilaminated structure are assumed to be perfectly bonded to each other, so that

$$r_i^{(s)} = r_i^{(s-1)} + h^{(s-1)} \quad (s = 2, \dots, N), \tag{3}$$

where $r_i^{(s)}$ is given by (1). Condition (3) provides all thicknesses $h^{(s)}$ ($s = 2, \dots, N$) as a function of the thickness of the first layer $h^{(1)}$, remaining the sole kinematical unknown of the problem, to be determined as the solution of the boundary-value problem.

Since the layers are assumed to be perfectly bonded, all radial coordinates $r^{(s)}$ can be referred to the same origin O of a cylindrical coordinate system $Or\theta z$, common to all deformed layers; therefore, the index s will be omitted in the following in all coordinates and the deformed configuration will be described in terms of the global system $Or\theta z$.

Finally, the kinematics provides all stretches in the multilayer which can be represented as

$$\lambda_r = \frac{1}{\lambda} = \frac{l_0}{2\bar{\theta} r}, \quad \lambda_\theta = \lambda = \frac{2\bar{\theta} r}{l_0}, \quad \lambda_z = 1, \tag{4}$$

and the current thickness of the s–th layer, $h^{(s)}$, as a function of $h^{(s-1)}$, namely

$$h^{(s)} = -\frac{l_0 h_0^{(s-1)}}{2\bar{\theta} h^{(s-1)}} - \frac{h^{(s-1)}}{2} + \sqrt{\left(\frac{l_0 h_0^{(s-1)}}{2\bar{\theta} h^{(s-1)}} + \frac{h^{(s-1)}}{2}\right)^2 + \frac{l_0 h_0^{(s)}}{\bar{\theta}}} \quad (s = 2, \dots, N), \tag{5}$$

determining all current thicknesses as functions of the thickness of the first layer, $h^{(1)}$ (in [Roccabianca et al. 2010, Equation (21)], corresponding to our (5), a minus sign appears by mistake instead of a plus in front of the last term under the square root).

A constitutive prescription between the principal components of the Cauchy stress T_i $(i = 1, 2, 3)$ and the principal stretches λ_i $(i = 1, 2, 3)$ can be written as

$$T_i = -\pi + \lambda_i \frac{\partial W(\lambda_1, \lambda_2, \lambda_3)}{\partial \lambda_i}, \quad \lambda_1 \lambda_2 \lambda_3 = 1, \tag{6}$$

where the index i is not summed and π is an arbitrary Lagrangian multiplier representing the undetermined hydrostatic pressure.

For a generic layer s, the Cauchy stress can be represented in polar coordinates as

$$\boldsymbol{T}^{(s)} = T_r^{(s)} \boldsymbol{e}_r \otimes \boldsymbol{e}_r + T_\theta^{(s)} \boldsymbol{e}_\theta \otimes \boldsymbol{e}_\theta + T_z^{(s)} \boldsymbol{e}_z \otimes \boldsymbol{e}_z, \tag{7}$$

and calculated from the constitutive equations (6). Here, we adopt the Mooney–Rivlin strain-energy function defined in terms of the moduli c_1 and c_2 and of the left Cauchy–Green deformation tensor $\boldsymbol{B}$ as (for layer s)

$$W^{(s)} = \frac{c_1^{(s)}}{2}(\operatorname{tr} \boldsymbol{B} - 3) + \frac{c_2^{(s)}}{2}(\operatorname{tr} \boldsymbol{B}^{-1} - 3). \tag{8}$$

For a plane-strain deformation, defining the shear modulus μ_0 as

$$\mu_0^{(s)} = c_1^{(s)} + c_2^{(s)}, \tag{9}$$

the integration of the equilibrium equation ($\operatorname{div} \boldsymbol{T} = \boldsymbol{0}$) within each layer yields, for the components of Cauchy stress,

$$
\begin{aligned}
T_r^{(s)} &= \hat{W}^{(s)} + \gamma^{(s)} = \frac{\mu_0^{(s)}}{2}\left(\lambda^2 + \frac{1}{\lambda^2}\right) + \gamma^{(s)}, \\
T_\theta^{(s)} &= \left(\lambda \hat{W}^{(s)}\right)' + \gamma^{(s)} = \frac{\mu_0^{(s)}}{2}\left(3\lambda^2 - \frac{1}{\lambda^2}\right) + \gamma^{(s)},
\end{aligned}
\tag{10}
$$

where $\hat{W}^{(s)}(\lambda) = W^{(s)}(1/\lambda, \lambda, 1)$, $\gamma^{(s)}$ is an integration constant, and $()'$ denotes differentiation with respect to the stretch λ. The component $T_z^{(s)}$ can be inferred from (6). Note that in plane strain, the definition of W in (8) specializes to

$$W^{(s)} = \frac{c_1^{(s)} + c_2^{(s)}}{2}\left(\lambda^2 + \frac{1}{\lambda^2} - 2\right) = \frac{\mu_0^{(s)}}{2}\left(\lambda^2 + \frac{1}{\lambda^2} - 2\right). \tag{11}$$

The N constants $\gamma^{(s)}$ can be determined by imposing the boundary conditions at the external sides of the layered system, namely

$$T_r^{(1)}(r_i^{(1)}) = 0, \quad T_r^{(N)}(r_e^{(N)}) = 0, \tag{12}$$

and the interfacial continuity conditions

$$T_r^{(s-1)}(r_e^{(s-1)}) = T_r^{(s)}(r_i^{(s)}) \quad (s = 2, \ldots, N), \tag{13}$$

where $r_e^{(s-1)} = r_i^{(s-1)} + h^{(s-1)}$ $(s = 2, \ldots, N + 1)$. In particular, the constant $\gamma^{(N)}$ can be calculated from $(12)_2$ using $(10)_1$ as

$$\gamma^{(N)} = -\frac{\mu_0^{(N)}}{2}\left[(\alpha r_e^{(N)})^2 + \frac{1}{(\alpha r_e^{(N)})^2}\right], \tag{14}$$

while from (13) the following recursion rule, useful in computing all the remaining constants $\gamma^{(s)}$, can be inferred

$$\gamma^{(s-1)} = \frac{\mu_0^{(s)} - \mu_0^{(s-1)}}{2} \left[(\alpha r_e^{(s-1)})^2 + \frac{1}{(\alpha r_e^{(s-1)})^2} \right] + \gamma^{(s)} \quad (s = 2, \ldots, N), \tag{15}$$

where $r_i^{(s)} = r_e^{(s-1)}$. In the special case $N = 2$, (15) and (12)$_1$ provide

$$\frac{\mu_0^{(1)}}{2} \left[(\alpha r_i^{(1)})^2 + \frac{1}{(\alpha r_i^{(1)})^2} \right] + \frac{\mu_0^{(2)} - \mu_0^{(1)}}{2} \left[(\alpha r_e^{(1)})^2 + \frac{1}{(\alpha r_e^{(1)})^2} \right] + \gamma^{(2)} = 0, \tag{16}$$

in which $r_i^{(1)}$, $r_e^{(1)}$ and $\gamma^{(2)}$ are all functions of $h^{(1)}$, through (1), (5), and (15). Therefore, (16) can be numerically solved to obtain the current thickness of the first layer, $h^{(1)}$, which determines the solution of a bilayer subject to finite bending.

Now that the bending solution is known, this can be used as the fundamental solution in an analysis of incremental bifurcations. The incremental equilibrium is expressed in terms of the updated incremental first Piola–Kirchhoff stress $\boldsymbol{\Sigma}$ by

$$\text{div } \boldsymbol{\Sigma} = \boldsymbol{0}, \tag{17}$$

where

$$\boldsymbol{\Sigma} = \dot{\boldsymbol{S}} \boldsymbol{F}^T, \quad \dot{\boldsymbol{S}} = \dot{\boldsymbol{T}} \boldsymbol{F}^{-T} - \boldsymbol{T} \boldsymbol{L}^T \boldsymbol{F}^{-T}, \tag{18}$$

$\boldsymbol{S} = \boldsymbol{T} \boldsymbol{F}^{-T}$ and a superposed dot is used to denote a first-order increment. $\boldsymbol{L}$ is the gradient of incremental displacements with cylindrical components

$$\boldsymbol{L} = u_{r,r} \boldsymbol{e}_r \otimes \boldsymbol{e}_r + \frac{u_{r,\theta} - u_\theta}{r} \boldsymbol{e}_r \otimes \boldsymbol{e}_\theta + u_{\theta,r} \boldsymbol{e}_\theta \otimes \boldsymbol{e}_r + \frac{u_r + u_{\theta,\theta}}{r} \boldsymbol{e}_\theta \otimes \boldsymbol{e}_\theta, \tag{19}$$

subject to the constraint tr $\boldsymbol{L} = 0$ (incremental incompressibility), namely,

$$r u_{r,r} + u_r + u_{\theta,\theta} = 0. \tag{20}$$

The linearized constitutive equation is

$$\boldsymbol{\Sigma} = \mathbb{C} \boldsymbol{L} - \dot{\pi} \boldsymbol{I}, \tag{21}$$

where $\mathbb{C}$ is the fourth-order tensor of instantaneous elastic moduli and, since $\boldsymbol{\Sigma} = \dot{\boldsymbol{T}} - \boldsymbol{T} \boldsymbol{L}^T$ [see (18)], the balance of rotational momentum yields

$$\Sigma_{12} - \Sigma_{21} = T_2 L_{12} - T_1 L_{21}, \tag{22}$$

so that

$$\mathbb{C}_{ijji} + T_i = \mathbb{C}_{jiji} \quad (i \neq j), \tag{23}$$

where the indices i and j are not summed.

For an incompressible isotropic elastic material, the components of $\mathbb{C}$ can be written as functions of two incremental moduli, denoted by μ and μ_*, that depend on the current deformation. In cylindrical

coordinates, the nonvanishing components of $\mathbb{C}$ are (see [Hill and Hutchinson 1975; Gei and Ogden 2002])

$$
\begin{aligned}
C_{rrrr} = C_{\theta\theta\theta\theta} = 2\mu_* + p, \quad & C_{\theta r\theta r} = \mu - \Gamma, \\
C_{r\theta r\theta} = \mu + \Gamma, \quad & C_{r\theta\theta r} = C_{\theta rr\theta} = \mu + p,
\end{aligned}
\tag{24}
$$

where

$$
\Gamma = \frac{T_\theta - T_r}{2} \quad \text{and} \quad p = -\frac{T_\theta + T_r}{2},
\tag{25}
$$

describe the state of prestress. Hence, the incremental constitutive equations (21) take the explicit form

$$
\begin{aligned}
\Sigma_{rr} = -\dot{\pi} + (2\mu_* + p)u_{r,r}, \quad & \Sigma_{\theta\theta} = -\dot{\pi} + (2\mu_* + p)\frac{u_r + u_{\theta,\theta}}{r}, \\
\Sigma_{r\theta} = (\mu + \Gamma)\frac{u_{r,\theta} - u_\theta}{r} + (\mu + p)u_{\theta,r}, \quad & \Sigma_{\theta r} = (\mu + p)\frac{u_{r,\theta} - u_\theta}{r} + (\mu - \Gamma)u_{\theta,r}.
\end{aligned}
\tag{26}
$$

For a Mooney–Rivlin material deformed in plane strain, the two moduli μ and μ_* coincide and are equal to [see (8)]

$$
\mu = \mu_* = \frac{\mu_0}{2}\frac{(\lambda^4 + 1)}{\lambda^2}.
\tag{27}
$$

We seek bifurcation represented by an incremental displacement field in the form

$$
\begin{cases}
u_r(r, \theta) = f(r)\cos n\theta, \\
u_\theta(r, \theta) = g(r)\sin n\theta, \\
\dot{\pi}(r, \theta) = k(r)\cos n\theta,
\end{cases}
\tag{28}
$$

so that (20) can be reformulated as

$$
g = -\frac{(f + rf')}{n},
\tag{29}
$$

and the incremental equilibrium equations as

$$
\begin{aligned}
k' &= Df'' + \left(C_{,r} + D_{,r} + \frac{C + 2D}{r}\right)f' + \frac{E(1 - n^2)}{r^2}f, \\
k &= \frac{r^2 C}{n^2}f''' + \frac{F + 3C}{n^2}rf'' + \left(\frac{F}{n^2} - D\right)f' - \frac{1 - n^2}{n^2}\frac{F}{r}f,
\end{aligned}
\tag{30}
$$

where coefficients C, D, E and F can be expressed (for a Mooney–Rivlin material) as

$$
\begin{aligned}
C &= \mu - \Gamma = \frac{\mu_0}{\lambda^2}, \quad & D &= 2\mu_* - \mu = \frac{\mu_0}{2}\frac{\lambda^4 + 1}{\lambda^2}, \\
E &= \mu + \Gamma = \mu_0\lambda^2, \quad & F &= rC_{,r} + C = -\frac{\mu_0}{\lambda^2}.
\end{aligned}
\tag{31}
$$

By differentiating $(30)_2$ with respect to r and substituting the result into $(30)_1$, a single differential equation in terms of $f(r)$ is obtained:

$$
r^4 f'''' + 2r^3 f''' - \left(3 + n^2(\lambda^4 + 1)\right)r^2 f'' + \left(3 + n^2(1 - 3\lambda^4)\right)rf' + (n^2 - 1)(3 + n^2\lambda^4)f = 0,
\tag{32}
$$

defining the function $f(r)$ within a generic layer. Once $f(r)$ is known for each layer, the other functions, $g(r)$ and $k(r)$, can be calculated by employing (29) and (30)$_2$, respectively, so that function $f(r)$ becomes the primary unknown.

The differential equation (32) for the functions $f^{(s)}(r)$ ($s = 1, \dots, N$) is complemented by the following boundary conditions:

- continuity of incremental tractions and displacements across interfaces:

$$f^{(s)}\big|_{r=r_e^{(s)}} = f^{(s+1)}\big|_{r=r_i^{(s+1)}},$$

$$\{f + rf'\}^{(s)}\big|_{r=r_e^{(s)}} = \{f + rf'\}^{(s+1)}\big|_{r=r_i^{(s+1)}},$$

$$\left\{\mu_0\left(r^3 f''' + 2r^2 f'' - \left(1 + n^2\left(\tfrac{1}{2}(3+\lambda^4) - \bar{\gamma}\right)\right)rf' + (1-n^2)f\right)\right\}^{(s)}\bigg|_{r=r_e^{(s)}} =$$

$$\left\{\mu_0\left(r^3 f''' + 2r^2 f'' - \left(1 + n^2\left(\tfrac{1}{2}(3+\lambda^4) - \bar{\gamma}\right)\right)rf' + (1-n^2)f\right)\right\}^{(s+1)}\bigg|_{r=r_i^{(s+1)}},$$

(33)

$$\left\{\mu_0\left(r^2 f'' + \left(\tfrac{1}{2}(3+\lambda^4) + \bar{\gamma}\right)rf' + (1-n^2)\left(\tfrac{1}{2}(\lambda^4 - 1) + \bar{\gamma}\right)f\right)\right\}^{(s)}\bigg|_{r=r_e^{(s)}} =$$

$$\left\{\mu_0\left(r^2 f'' + \left(\tfrac{1}{2}(3+\lambda^4) + \bar{\gamma}\right)rf' + (1-n^2)\left(\tfrac{1}{2}(\lambda^4 - 1) + \bar{\gamma}\right)f\right)\right\}^{(s+1)}\bigg|_{r=r_i^{(s+1)}},$$

where $\bar{\gamma}^{(s)} = \gamma^{(s)}\lambda^2/\mu_0^{(s)}$;

- null tractions (for dead loading) at the external surfaces ($r = r_i^{(1)}$ and $r = r_e^{(N)}$):

$$\left\{r^3 f''' + 2r^2 f'' - \left(1 + n^2(2 + \lambda^4)\right)rf' + (1-n^2)f\right\}^{(1),(N)}\bigg|_{r=r_i^{(1)}, r_e^{(N)}} = 0,$$

(34)

$$\left\{r^2 f'' + rf' + (n^2 - 1)f\right\}^{(1),(N)}\bigg|_{r=r_i^{(1)}, r_e^{(N)}} = 0;$$

- null incremental normal displacement and nominal shear stress (at $\theta = \pm\bar{\theta}$), which yields

$$n = \frac{2m\pi}{\alpha l_0} \quad (m = 1, 2, 3, \dots). \tag{35}$$

3. The compound matrix method for a bilayer

The compound matrix method was initially proposed in [Backus and Gilbert 1967] and applied to problems of fluid mechanics [Backus and Gilbert 1967; Ng and Reid 1979a; 1979b; 1985, Anturkar et al. 1992; Yiantsios and Higgins 1988] and solid mechanics [Lindsay 1992; Lindsay and Rooney 1992]. Haughton and Orr [1995] used the method in incremental elasticity, while Haughton [1999], Dryburgh and Ogden [1999] and Destrade et al. [2009; 2010] employed it to investigate instabilities of a homogeneous block subjected to finite flexure. Our aim is to show the application to elastic multilayers subject to finite bending, in the simple case of a bilayer.

The differential equation (32) can be rewritten as a linear system of first-order ODEs, that in the case of two elastic layers can be cast in the standard form

$$\boldsymbol{y}' = \boldsymbol{A}\boldsymbol{y}, \quad \boldsymbol{z}' = \boldsymbol{B}\boldsymbol{z}, \tag{36}$$

where the vectors $\boldsymbol{y}$ and $\boldsymbol{z}$ are defined as

$$
\begin{aligned}
\boldsymbol{y}(r) &= \left[f^{(1)}(r) \ \ f^{(1)\prime}(r) \ \ f^{(1)\prime\prime}(r) \ \ f^{(1)\prime\prime\prime}(r) \right]^T, \\
\boldsymbol{z}(r) &= \left[f^{(2)}(r) \ \ f^{(2)\prime}(r) \ \ f^{(2)\prime\prime}(r) \ \ f^{(2)\prime\prime\prime}(r) \right]^T,
\end{aligned}
\tag{37}
$$

and the matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, which depend on the radial coordinate r, as

$$
\boldsymbol{A}(r) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ A_{41} & A_{42} & A_{43} & A_{44} \end{bmatrix}, \quad
\boldsymbol{B}(r) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ B_{41} & B_{42} & B_{43} & B_{44} \end{bmatrix}.
\tag{38}
$$

The components of $\boldsymbol{A}$ and $\boldsymbol{B}$, as well as those of other matrices and vectors introduced in this Section are listed in Appendix A.

The boundary conditions at the two external surfaces of the layer, equations (34), are equivalent to

$$\boldsymbol{C}\boldsymbol{y}(r_i) = \boldsymbol{0}, \quad \boldsymbol{D}\boldsymbol{z}(r_e) = \boldsymbol{0}, \tag{39}$$

where $r_i = r_i^{(1)}$, $r_e = r_i^{(2)} + h^{(2)}$ and matrices $\boldsymbol{C}$ and $\boldsymbol{D}$ are

$$
\boldsymbol{C} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} \\ C_{21} & C_{22} & C_{23} & 0 \end{bmatrix}, \quad
\boldsymbol{D} = \begin{bmatrix} D_{11} & D_{12} & D_{13} & D_{14} \\ D_{21} & D_{22} & D_{23} & 0 \end{bmatrix}.
\tag{40}
$$

The continuity conditions (33) between the two layers, can be written as

$$\boldsymbol{G}\boldsymbol{y}(r_m) + \boldsymbol{H}\boldsymbol{z}(r_m) = \boldsymbol{0}, \tag{41}$$

where $r_m = r_i^{(1)} + h^{(1)}$ and matrices $\boldsymbol{G}$ and $\boldsymbol{H}$ are defined as

$$
\boldsymbol{G} = \begin{bmatrix} G_{11} & G_{12} & G_{13} & G_{14} \\ G_{21} & G_{22} & G_{23} & 0 \\ G_{31} & 0 & 0 & 0 \\ G_{41} & G_{42} & 0 & 0 \end{bmatrix}, \quad
\boldsymbol{H} = \begin{bmatrix} H_{11} & H_{12} & H_{13} & H_{14} \\ H_{21} & H_{22} & H_{23} & 0 \\ H_{31} & 0 & 0 & 0 \\ H_{41} & H_{42} & 0 & 0 \end{bmatrix}.
\tag{42}
$$

It is now convenient to rearrange the four solutions of (36): two for the first layer, $\boldsymbol{y}^I$, $\boldsymbol{y}^{II}$, and two for the second layer, $\boldsymbol{z}^I$, $\boldsymbol{z}^{II}$, [these solutions already satisfy the boundary conditions (39), but still not the interface conditions (42)] into two matrices sharing the common structure

$$
\begin{bmatrix} *_1^I & *_1^{II} \\ *_2^I & *_2^{II} \\ *_3^I & *_3^{II} \\ *_4^I & *_4^{II} \end{bmatrix}
\tag{43}
$$

(where the symbol $*$ stands for either y or z) and defining the so-called compound matrices. Moreover, we introduce the vectors ϕ_i^y ($i = 1, \ldots, 6$) and ϕ_i^z ($i = 1, \ldots, 6$) collecting the components of the minors of matrices (43) as

$$\phi_1^* = *_1^I *_2^{II} - *_2^I *_1^{II}, \quad \phi_4^* = *_2^I *_3^{II} - *_3^I *_2^{II},$$

$$\phi_2^* = *_1^I *_3^{II} - *_3^I *_1^{II}, \quad \phi_5^* = *_2^I *_4^{II} - *_4^I *_2^{II}, \qquad (44)$$

$$\phi_3^* = *_1^I *_4^{II} - *_4^I *_1^{II}, \quad \phi_6^* = *_3^I *_4^{II} - *_4^I *_3^{II}.$$

With the definitions (43) and (44), the differential problem (36) can be shown [Ng and Reid 1979a] to be equivalent to the new problem

$$\left(\boldsymbol{\phi}^y\right)' = \boldsymbol{P}^A \boldsymbol{\phi}^y, \quad \left(\boldsymbol{\phi}^z\right)' = \boldsymbol{P}^B \boldsymbol{\phi}^z, \qquad (45)$$

where, introducing the symbol $\square$, equal to A (to B) for $\boldsymbol{\phi}^y$ (for $\boldsymbol{\phi}^z$), we define

$$\boldsymbol{P}^\square = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ \square_{42} & \square_{43} & \square_{44} & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ -\square_{41} & 0 & 0 & \square_{43} & \square_{44} & 1 \\ 0 & -\square_{41} & 0 & -\square_{42} & 0 & \square_{44} \end{bmatrix}. \qquad (46)$$

The system of differential equations (45) has to be solved using a Runge–Kutta (4,5) numerical method (we have used Matlab version 7.9) to determine the compound matrices $\boldsymbol{\phi}^y$ and $\boldsymbol{\phi}^z$.

The solution of the bifurcation problem can be written as a linear combination of the solutions $\boldsymbol{y}^I$, $\boldsymbol{y}^{II}$, $\boldsymbol{z}^I$, and $\boldsymbol{z}^{II}$,

$$\boldsymbol{y} = \xi_1 \boldsymbol{y}^I + \xi_2 \boldsymbol{y}^{II}, \quad \boldsymbol{z} = \xi_3 \boldsymbol{z}^I + \xi_4 \boldsymbol{z}^{II}, \qquad (47)$$

where the arbitrary coefficients ξ_i ($i = 1, \ldots, 4$), which set the amplitude of the bifurcation mode, remain undefined in a linearized analysis. The conditions at the internal interface (41) can be recast as

$$\boldsymbol{M}\boldsymbol{\xi} = \boldsymbol{0}, \quad \text{with } [\boldsymbol{\xi}] = [\xi_1 \ \xi_2 \ \xi_3 \ \xi_4]^T, \qquad (48)$$

where

$$\boldsymbol{M} = \begin{bmatrix} \left(\boldsymbol{G}\boldsymbol{y}^I\right)_1 & \left(\boldsymbol{G}\boldsymbol{y}^{II}\right)_1 & \left(\boldsymbol{H}\boldsymbol{z}^I\right)_1 & \left(\boldsymbol{H}\boldsymbol{z}^{II}\right)_1 \\ \left(\boldsymbol{G}\boldsymbol{y}^I\right)_2 & \left(\boldsymbol{G}\boldsymbol{y}^{II}\right)_2 & \left(\boldsymbol{H}\boldsymbol{z}^I\right)_2 & \left(\boldsymbol{H}\boldsymbol{z}^{II}\right)_2 \\ \left(\boldsymbol{G}\boldsymbol{y}^I\right)_3 & \left(\boldsymbol{G}\boldsymbol{y}^{II}\right)_3 & \left(\boldsymbol{H}\boldsymbol{z}^I\right)_3 & \left(\boldsymbol{H}\boldsymbol{z}^{II}\right)_3 \\ \left(\boldsymbol{G}\boldsymbol{y}^I\right)_4 & \left(\boldsymbol{G}\boldsymbol{y}^{II}\right)_4 & \left(\boldsymbol{H}\boldsymbol{z}^I\right)_4 & \left(\boldsymbol{H}\boldsymbol{z}^{II}\right)_4 \end{bmatrix}, \qquad (49)$$

so that the bifurcation condition, depending on the bending half-angle $\bar{\theta}$, the undeformed aspect ratios l_0/h_0 and $h_0^{(1)}/h_0^{(2)}$, and the stiffness ratio $\mu_0^{(1)}/\mu_0^{(2)}$, becomes

$$\det(\boldsymbol{M}) = 0. \qquad (50)$$

Condition (50) can be rewritten as the sum of 2×2-determinants as

$$\sum_{i=0}^{1}(-1)^i \left\{ \begin{vmatrix} M_{1+i,1} & M_{1+i,2} \\ M_{41} & M_{42} \end{vmatrix} \begin{vmatrix} M_{2-i,3} & M_{2-i,4} \\ M_{33} & M_{34} \end{vmatrix} - \begin{vmatrix} M_{2+i,1} & M_{2+i,2} \\ M_{11} & M_{12} \end{vmatrix} \begin{vmatrix} M_{3-i,3} & M_{3-i,4} \\ M_{43} & M_{44} \end{vmatrix} \right.$$
$$\left. - \begin{vmatrix} M_{2+2i,1} & M_{2+2i,2} \\ M_{31} & M_{32} \end{vmatrix} \begin{vmatrix} M_{4-2i,3} & M_{4-2i,4} \\ M_{13} & M_{14} \end{vmatrix} \right\} = 0. \quad (51)$$

The determinants can be expressed in terms of the compound matrices $\boldsymbol{\phi}^y$ and $\boldsymbol{\phi}^z$ as

$$\begin{vmatrix} M_{k1} & M_{k2} \\ M_{l1} & M_{l2} \end{vmatrix} = (G_{k1}G_{l2} - G_{k2}G_{l1})\phi_1^y + (G_{k1}G_{l3} - G_{k3}G_{l1})\phi_2^y + (G_{k1}G_{l4} - G_{k4}G_{l1})\phi_3^y$$
$$+ (G_{k2}G_{l3} - G_{k3}G_{l2})\phi_4^y + (G_{k2}G_{l4} - G_{k4}G_{l2})\phi_5^y + (G_{k3}G_{l4} - G_{k4}G_{l3})\phi_6^y \quad (52)$$

and

$$\begin{vmatrix} M_{k3} & M_{k4} \\ M_{l3} & M_{l4} \end{vmatrix} = (H_{k1}H_{l2} - H_{k2}H_{l1})\phi_1^z + (H_{k1}H_{l3} - H_{k3}H_{l1})\phi_2^z + (H_{k1}H_{l4} - H_{k4}H_{l1})\phi_3^z$$
$$+ (H_{k2}H_{l3} - H_{k3}H_{l2})\phi_4^z + (H_{k2}H_{l4} - H_{k4}H_{l2})\phi_5^z + (H_{k3}H_{l4} - H_{k4}H_{l3})\phi_6^z, \quad (53)$$

where the indices k and l take the values corresponding to the representation (51).

Once the undeformed aspect ratios l_0/h_0 and $h_0^{(1)}/h_0^{(2)}$ and the stiffness ratio $\mu_0^{(1)}/\mu_0^{(2)}$ have been fixed, the bifurcation condition (50), through representation (51), becomes a function of the bending half-angle $\bar{\theta}$ only, to be solved numerically (we have used the function "fzero" of Matlab). An example of the advantage related to the use of the compound matrix method over the usual determinantal method is reported in Figure 2, where $\det(\boldsymbol{M})$ is plotted as a function of $\bar{\theta}$ for a stiff case, in which the superiority of the former approach is evident (note the spurious oscillations of the determinantal method). In this
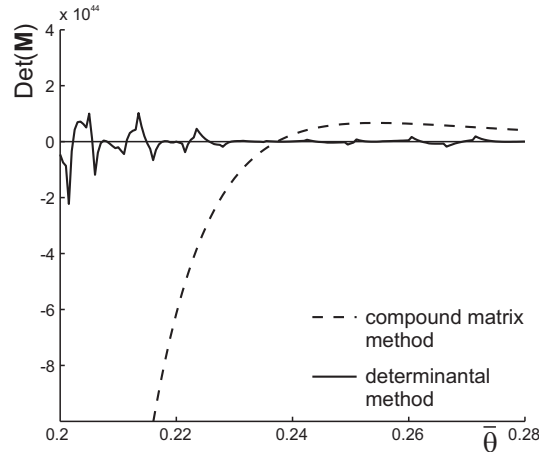
Figure 2. The compound matrix method (dashed line) against the determinantal method (solid line): $\det(\boldsymbol{M})$ is evaluated at different angles $\bar{\theta}$, for $l_0/h_0 = 0.1$, $h_0^{(1)}/h_0^{(2)} = (1\,\text{mm})/(5\,\text{mm})$ and $\mu_0^{(1)}/\mu_0^{(2)} = (7\,\text{N/mm}^2)/(1\,\text{N/mm}^2)$. Bifurcation corresponds to the vanishing of $\det(\boldsymbol{M})$; note the spurious oscillations of the latter method.

particular case, the 2-norm condition number of the matrix $\boldsymbol{M}$ is equal to 9.37×10^{27}, a value confirming that the matrix is bad conditioned.

4. Bifurcations of a two-layer system

Results for bifurcation of bent configurations for bilayers are shown in Figure 3, in terms of critical half-angle $\bar{\theta}_{cr}$ (upper graphs) and critical stretch $\lambda_{cr}(r_i^{(1)})$ at the compressed side of the specimen
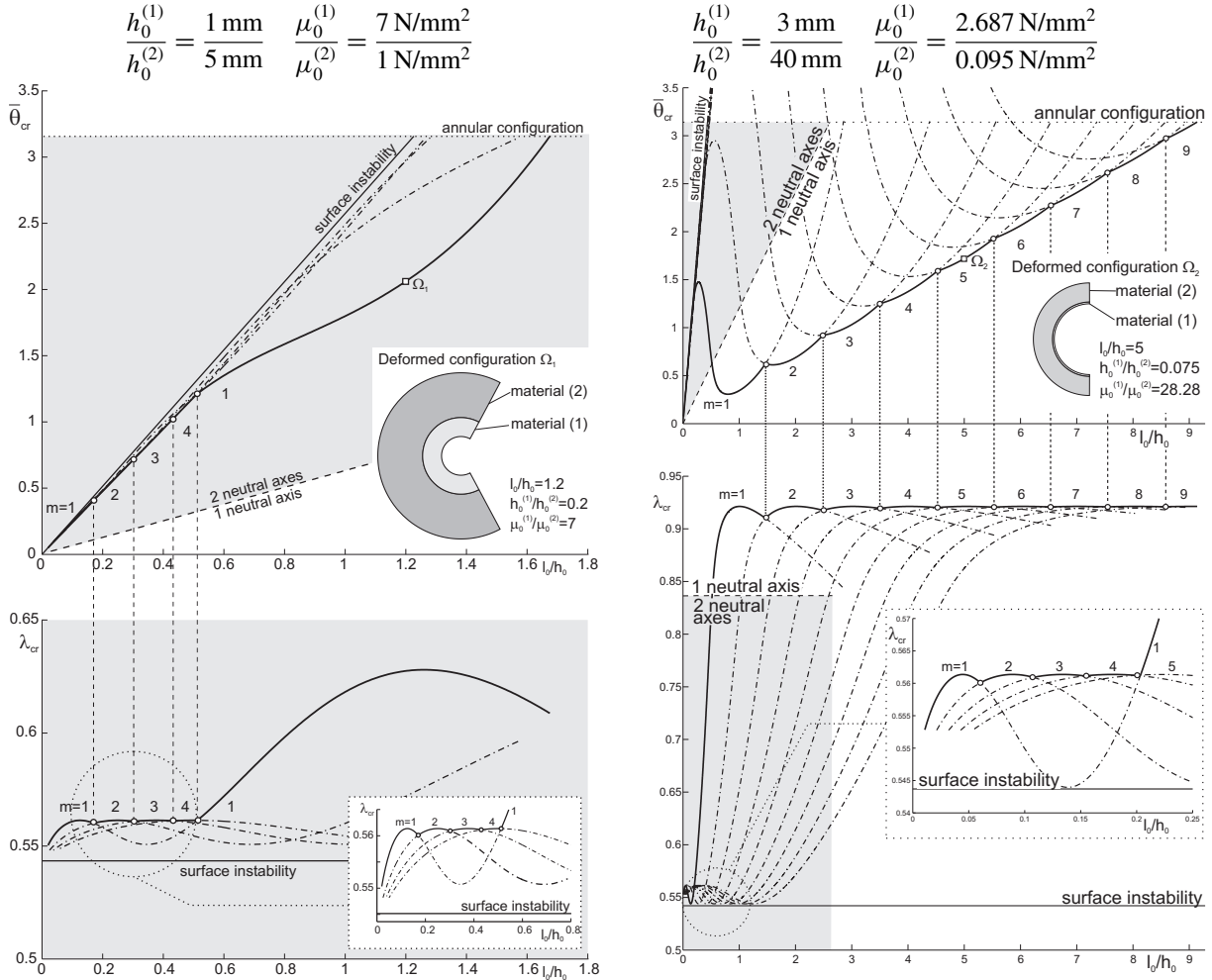
Figure 3. Critical angle $\bar{\theta}_{cr}$ and critical stretch λ_{cr} (evaluated at the internal boundary, $r = r_i^{(1)}$) versus aspect ratio l_0/h_0 for a Mooney–Rivlin bilayer coated with a stiff layer and subject to bending with $h_0^{(1)}/h_0^{(2)}$ and $\mu_0^{(1)}/\mu_0^{(2)}$ as indicated above the graphs. The stiff layer is located at the side in compression. In each plot, a small circle denotes a transition between two integer values of m (the parameter which sets the circumferential wavenumber). In the lower plots, the insert contains a magnification of the region where bifurcations occur at low l_0/h_0. Two neutral axes occur in the region marked gray.

(lower graphs) as functions of the global aspect ratio, that is, the initial length divided by the initial total thickness. The ratios between the thicknesses and the shear coefficients μ_0 of the layers are $(1\,\text{mm})/(5\,\text{mm})$ and $(7\,\text{N/mm}^2)/(1\,\text{N/mm}^2)$ for the left half of Figure 3, and $(3\,\text{mm})/(40\,\text{mm})$ and $(2.687\,\text{N/mm}^2)/(0.095\,\text{N/mm}^2)$ for the right half. The various curves reported in the figure represent solutions corresponding to different bifurcation modes, singled out by the circumferential wavenumber m. The mode visible in an experiment is that corresponding to the lower value of the critical half-angle, $\bar{\theta}_{cr}$, or to the higher value of critical stretch at the compressed side, λ_{cr}. Note that the gray zone represents the range of aspect ratios and bending half-angle for which two neutral axes occur.

Within the set of aspect ratios and stiffness contrast analyzed in the left part of Figure 3, a bifurcation only appears when two neutral axes have been formed, while it may occur when two or one neutral axes are present in the right half of the figure. In all cases analyzed, including Figure 3, we have found that the gray zone in the $\bar{\theta}_{cr}$–l_0/h_0 graphs is bounded by a straight line, becoming a horizontal line in the λ_{cr}–l_0/h_0 representation. An explanation for this fact is given in Appendix C.

The special feature emerging from Figure 3, left (and not found within the set of parameters and geometries investigated in [Roccabianca et al. 2010]) is that the mode $m = 1$ of bifurcation becomes the critical mode for a sufficiently high slenderness, so that here long-wavelength bifurcations (corresponding to small m) become well-separated from surface modes (corresponding to high m) and thus fully visible. This feature is also present in Figure 3, right, which has been produced with values of parameters corresponding to commercially available rubbers (and tested by us; see Appendix B). In this way it has been possible to produce the two samples shown in Figures 1 and 4, differing only in the aspect ratio (taken equal to 2 for the sample shown in Figure 1 and 1.5 for that shown in Figure 4) and evidencing long-wavelength bifurcation modes.

The progression of bending is shown in Figure 4, referred to a $64.5\,\text{mm} \times 3\,\text{mm} \times 150\,\text{mm}$ rubber strip glued to a $64.5\,\text{mm} \times 40\,\text{mm} \times 150\,\text{mm}$ neoprene plate (in which the larger dimension is that out-of-plane, taken sufficiently large, 150 mm, to simulate the plane strain condition). At a certain stage of finite bending, namely at a certain bending half-angle $\bar{\theta}_{cr}$, a long-wavelength mode can be detected to appear on the surface of the sample (Figure 4, right), which *qualitatively* confirms our findings.

Figure 4. Progressive bending of a two-layer rubber block (undeformed configuration is shown on the left, a bent configuration in the center), evidencing bifurcation with long-wavelength bifurcation modes (shown on the right). Stiffness and thickness ratios between layers are $(2.687\,\text{N/mm}^2)/(0.095\,\text{N/mm}^2)$ and $(3\,\text{mm})/(40\,\text{mm})$, respectively. The stiff layer ($64.5\,\text{mm} \times 3\,\text{mm} \times 150\,\text{mm}$, made up of natural rubber, marked with a white pencil on the sample) is at the compressive side and coats a neoprene layer ($64.5\,\text{mm} \times 40\,\text{mm} \times 150\,\text{mm}$).

From a quantitative point of view, the critical half-angle for bifurcation results from modeling to be equal to 39.40° for the sample shown in Figure 1 and 35.49° for that shown in Figure 4, values that are definitely higher than those found experimentally (30.00° for the former sample and 21.00° for the latter). The fact that the theoretical predictions correspond to bifurcation angles larger than those observed experimentally is also common to all previous experiments [Gent and Cho 1999; Roccabianca et al. 2010] and can be explained as the usual effect of imperfections (so that for instance the bending mode associated with the Euler buckling is always experimentally observed to become visible before the achievement of the critical load). The fact that the discrepancy between theoretical and experimental values is larger in the cases reported in the present article can be motivated in terms of the effect of the different *sensitivity* to imperfections. In fact, short-wavelength undulations introduced in the reference configuration begin to amplify and to become visible much closer to the bifurcation threshold than long-wavelength imperfections, a feature demonstrated through finite element numerical simulations [Roccabianca 2011].

5. Conclusions

Incremental bifurcations emanating from a finitely bent configuration of an elastic incompressible multi-layered block have been reconsidered after the article [Roccabianca et al. 2010]. It has been shown that the presence of two neutral axes is linked to the possibility of finding long-wavelength bifurcation modes well-separated from (short-wavelength) surface modes, a circumstance which was pointed out without proof in that article and is now definitely demonstrated. Since the equations governing the bifurcation become stiff (specially when more than one neutral axis is present), the treatment of the computational problem has been possible only through an use of the compound matrix method. Our findings have been substantiated by qualitative experiments, giving full evidence to long-wavelength modes.

Appendix A. Components of the compound matrices in (38), (40), (42), and (46)

$$A_{41}(r) = \frac{(n^2 - 1)(F^{(1)} - rF_{,r}^{(1)} - n^2 E^{(1)})}{C^{(1)}r^4}, \quad A_{42}(r) = \frac{2F^{(1)} - (rF^{(1)} + 2rn^2 D^{(1)})_{,r}}{C^{(1)}r^3},$$

$$A_{43}(r) = \frac{2n^2 D^{(1)} - (rF^{(1)})_{,r} - 4F^{(1)}}{C^{(1)}r^2}, \qquad A_{44}(r) = -2\frac{F^{(1)} + 2C^{(1)}}{C^{(1)}r}.$$

$$B_{41}(r) = \frac{(n^2 - 1)(F^{(2)} - rF_{,r}^{(2)} - n^2 E^{(2)})}{C^{(2)}r^4}, \quad B_{42}(r) = \frac{2F^{(2)} - (rF^{(2)} + 2rn^2 D^{(2)})_{,r}}{C^{(2)}r^3},$$

$$B_{43}(r) = \frac{2n^2 D^{(2)} - (rF^{(2)})_{,r} - 4F^{(2)}}{C^{(2)}r^2}, \qquad B_{44}(r) = -2\frac{F^{(2)} + 2C^{(2)}}{C^{(2)}r}.$$

$$C_{11} = F^{(1)}(n^2 - 1), \qquad C_{12} = r_i[F^{(1)} - n^2(2D^{(1)} + C^{(1)})], \qquad C_{13} = r_i^2(F^{(1)} + 3C^{(1)}),$$

$$D_{11} = F^{(2)}(n^2 - 1), \qquad D_{12} = r_e[F^{(2)} - n^2(2D^{(2)} + C^{(2)})], \qquad D_{13} = r_e^2(F^{(2)} + 3C^{(2)}),$$

$$C_{14} = r_i^3 C^{(1)}, \qquad C_{21} = n^2 - 1, \qquad C_{22} = r_i, \qquad C_{31} = r_i^2,$$

$$D_{14} = r_e^3 C^{(2)}, \qquad D_{21} = n^2 - 1, \qquad D_{22} = r_e, \qquad D_{23} = r_e^2.$$

$$G_{11} = F^{(1)}(n^2 - 1), \qquad G_{12} = r_m[F^{(1)} - n^2(2D^{(1)} + C^{(1)} - T_r^{(1)})],$$
$$H_{11} = F^{(2)}(1 - n^2), \qquad H_{12} = r_m[n^2(2D^{(2)} + C^{(2)} - T_r^{(2)}) - F^{(2)}],$$

$$G_{13} = r_m^2(F^{(1)} + 3C^{(1)}), \qquad G_{14} = r_m^3 C^{(1)}, \qquad G_{21} = (n^2 - 1)(C^{(1)} - T_r^{(1)}),$$
$$H_{13} = -r_m^2(F^{(2)} + 3C^{(2)}), \quad H_{14} = -r_m^3 C^{(2)}, \quad H_{21} = (1 - n^2)(C^{(2)} - T_r^{(2)}),$$

$$G_{22} = r_m(C^{(1)} + T_r^{(1)}), \qquad G_{23} = r_m^2 C^{(1)}, \qquad G_{31} = 1, \qquad G_{41} = 1, \qquad G_{42} = r_m,$$
$$H_{22} = -r_m(C^{(2)} + T_r^{(2)}), \qquad H_{23} = -r_m^2 C^{(2)}, \quad H_{31} = -1, \quad H_{41} = -1, \quad H_{42} = -r_m.$$

$$P_{31}^A = \frac{2F^{(1)} - (rF^{(1)} + 2rn^2 D^{(1)})_{,r}}{C^{(1)} r^3}, \qquad P_{32}^A = \frac{2n^2 D^{(1)} - (rF^{(1)})_{,r} - 4F^{(1)}}{C^{(1)} r^2}, \quad P_{33}^A = \frac{-2(F^{(1)} + 2C^{(1)})}{C^{(1)} r},$$

$$P_{51}^A = \frac{(1 - n^2)(F^{(1)} - rF_{,r}^{(1)} - n^2 E^{(1)})}{C^{(1)} r^4}, \quad P_{54}^A = \frac{2n^2 D^{(1)} - (rF^{(1)})_{,r} - 4F^{(1)}}{C^{(1)} r^2}, \quad P_{55}^A = \frac{-2(F^{(1)} + 2C^{(1)})}{C^{(1)} r},$$

$$P_{62}^A = \frac{(1 - n^2)(F^{(1)} - rF_{,r}^{(1)} - n^2 E^{(1)})}{C^{(1)} r^4}, \quad P_{64}^A = \frac{(rF^{(1)} + 2rn^2 D^{(1)})_{,r} - 2F^{(1)}}{C^{(1)} r^3}, \quad P_{66}^A = \frac{-2(F^{(1)} + 2C^{(1)})}{C^{(1)} r}.$$

$$P_{31}^B = \frac{2F^{(2)} - (rF^{(2)} + 2rn^2 D^{(2)})_{,r}}{C^{(2)} r^3}, \qquad P_{32}^B = \frac{2n^2 D^{(2)} - (rF^{(2)})_{,r} - 4F^{(2)}}{C^{(2)} r^2}, \quad P_{33}^B = \frac{-2(F^{(2)} + 2C^{(2)})}{C^{(2)} r},$$

$$P_{51}^B = \frac{(1 - n^2)(F^{(2)} - rF_{,r}^{(2)} - n^2 E^{(2)})}{C^{(2)} r^4}, \quad P_{54}^B = \frac{2n^2 D^{(2)} - (rF^{(2)})_{,r} - 4F^{(2)}}{C^{(2)} r^2}, \quad P_{55}^B = \frac{-2(F^{(2)} + 2C^{(2)})}{C^{(2)} r},$$

$$P_{62}^B = \frac{(1 - n^2)(F^{(2)} - rF_{,r}^{(2)} - n^2 E^{(2)})}{C^{(2)} r^4}, \quad P_{64}^B = \frac{(rF^{(2)} + 2rn^2 D^{(2)})_{,r} - 2F^{(2)}}{C^{(2)} r^3}, \quad P_{66}^B = \frac{-2(F^{(2)} + 2C^{(2)})}{C^{(2)} r}.$$

Appendix B. Experimental determination of the stiffness coefficient μ_0 of the employed materials

In our qualitative experiments to detect bifurcation, we have imposed finite bending to bilayered systems made up of a natural rubber strip (3 mm thick) and a neoprene block (40 mm thick). The bilayer is obtained by gluing the neoprene block to the natural rubber strip (we have used ethyl cyanoacrylate, Pattex$^{©}$).

Four dog-bone-shaped standard ISO 5277-1/1BA 30 mm $\times$ 5 mm specimens have been sampled from the two materials to characterize them in terms of Mooney–Rivlin model. Result of the tests (performed at room temperature with a Messphysik Midi 10 testing machine equipped with Doli Edc 222 acquisition and control electronics) are shown in Figure 5, in terms of true stress versus stretch. The selected ranges of stress and stretch for the tests correspond to the values expected in the bending experiments.

In the plots, the interpolation with the Mooney–Rivlin material (which provides a nearly linear response at the stretch under consideration) selected for the calculations is also included. The least square method provides for the natural rubber $c_1^{(\text{natrub})} = 0.007$ N/mm^2, $c_2^{(\text{natrub})} = 2.68$ N/mm^2 (corresponding to $\mu_0^{(\text{natrub})} \simeq 2.687$ N/mm^2) and for the neoprene $c_1^{(\text{neopr})} = 0.09$ N/mm^2, $c_2^{(\text{neopr})} = 0.005$ N/mm^2, giving $\mu_0^{(\text{neopr})} \simeq 0.095$ N/mm^2.
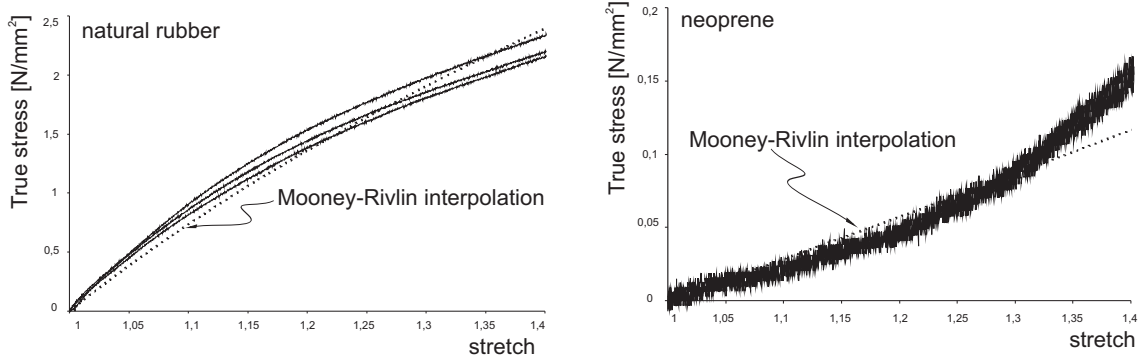
Figure 5. Uniaxial tests and material characterization of the natural rubber and the neoprene plate employed for the specimens to be subject to bending. Dotted curves represent Mooney–Rivlin interpolations employed in the analysis.

Appendix C. A justification of the fact that the region in which two neutral axes are present is bounded by a straight line in Figure 3

We provide a justification of the finding that, when two neutral axes occur in a bilayer, the stretch (at the compressed side) is independent of the global aspect ratio l_0/h_0, so that the gray zone (corresponding to the presence of two neutral axes) is bounded by a horizontal (inclined) line in the λ_{cr}–l_0/h_0 (in the $\bar{\theta}_{cr}$–l_0/h_0) representation, Figure 3.

The explanation of this effect is based on two observations:

i. During progressive bending of a bilayer with the stiff layer under compression, one neutral axis is present from the beginning of the deformation within the soft layer, while the second neutral axis always nucleates at the interface between the two layers (and then moves in the stiff layer).

ii. When the second neutral axis nucleates, the radial Cauchy stress T_r at the interface between layers takes a value independent of the initial aspect ratio l_0/h_0. We can therefore operate on a single layer by imposing, in addition to the usual bending, a pressure P^{ext} at one of its external sides (of initial length l_0) to correspond to the radial stress at the interface between layers. In particular, we can apply P^{ext} at the side where the longitudinal stretch is greater than 1.

To operate in dimensionless form, we introduce, from (2) and (4), the kinematic unknowns

$$\bar{\alpha} = \frac{2\bar{\theta}}{a}, \quad \bar{r} = \frac{r}{h_0}, \quad \bar{h} = \frac{h}{h_0}, \tag{C.1}$$

where $a = l_0/h_0$ is the aspect ratio of the undeformed configuration. The internal and external nondimensional radii, from (1), are

$$\bar{r}_i = \frac{a}{2\bar{\theta}\bar{h}} - \frac{\bar{h}}{2}, \quad \bar{r}_e = \bar{r}_i + \bar{h}. \tag{C.2}$$

As we want to write the bending problem in terms of the variable $\lambda_i = \lambda(\bar{r}_i)$, we calculate $\bar{\theta}$ as a function of a, $\bar{h}$ and λ_i, so that $(4)_2$ gives

$$\bar{\theta} = \frac{a}{\bar{h}} \left(\frac{1}{\bar{h}} - \lambda_i \right), \tag{C.3}$$

and the condition $\lambda_e = \lambda(\bar{r}_e)$ becomes

$$\lambda_e = \frac{2}{\bar{h}} - \lambda_i. \tag{C.4}$$

The boundary conditions for the layer under consideration are now

$$T_r(\bar{r}_i) = 0, \quad T_r(\bar{r}_e) = P^{\text{ext}}, \tag{C.5}$$

where

$$T_r = \frac{\mu_0}{2} \left(\lambda^2 + \frac{1}{\lambda^2} \right) + \gamma,$$

can be written from $(10)_1$. Equation $(C.5)_2$ provides the coefficient γ in the form

$$\gamma = P^{\text{ext}} - \frac{\mu_0}{2} \left(\lambda_e^2 + \frac{1}{\lambda_e^2} \right), \tag{C.6}$$

while, on the other hand, $(C.5)_1$ is equivalent to

$$\lambda_i^2 + \frac{1}{\lambda_i^2} + 2\frac{P^{\text{ext}}}{\mu_0} - \left[\left(\frac{2}{\bar{h}} - \lambda_i \right)^2 + \left(\frac{2}{\bar{h}} - \lambda_i \right)^{-2} \right] = 0, \tag{C.7}$$

from which it is clear that the unknown $\bar{h}$ is independent of a (but remains dependent on λ_i, μ_0, and P^{ext}). Therefore, since a neutral axis corresponds to

$$T_\theta(\bar{r}_e) = 0, \tag{C.8}$$

equations $(10)_2$, (C.6), and (C.4) show that the neutral axis condition is independent of a, so that the solution in terms of λ_i becomes only a function of μ_0 and P^{ext}.

References

[Anturkar et al. 1992] N. R. Anturkar, T. C. Papanastasiou, and J. O. Wilkes, "Compound matrix method for eigenvalue problems in multiple connected domains", *Comm. Appl. Numer. Methods* **8**:11 (1992), 811–818.

[Backus and Gilbert 1967] G. E. Backus and J. F. Gilbert, "Numerical applications of a formalism for geophysical inverse problems", *Geophys. J. Roy. Astron. Soc.* **13** (1967), 247–276.

[Benallal et al. 1993] A. Benallal, R. Billardon, and G. Geymonat, "Bifurcation and localization in rate-independent materials: some general considerations", pp. 1–44 in *Bifurcation and stability of dissipative systems* (Udine, 1991), edited by Q. S. Nguyen, CISM Courses and Lectures **327**, Springer, Vienna, 1993.

[Coman and Destrade 2008] C. D. Coman and M. Destrade, "Asymptotic results for bifurcations in pure bending of rubber blocks", *Quart. J. Mech. Appl. Math.* **61**:3 (2008), 395–414.

[Destrade et al. 2009] M. Destrade, A. Ní Annaidh, and C. D. Coman, "Bending instabilities of soft biological tissues", *Int. J. Solids Struct.* **46** (2009), 4322–4330.

[Destrade et al. 2010] M. Destrade, M. D. Gilchrist, J. A. Motherway, and J. G. Murphy, "Bimodular rubber buckles early in bending", *Mech. Mater.* **42** (2010), 469–476.

[Dryburgh and Ogden 1999] G. Dryburgh and R. W. Ogden, "Bifurcation of an elastic surface-coated incompressible isotropic elastic block subject to bending", *Z. Angew. Math. Phys.* **50**:5 (1999), 822–838.

[Gei and Ogden 2002] M. Gei and R. W. Ogden, "Vibration of a surface-coated elastic block subject to bending", *Math. Mech. Solids* **7**:6 (2002), 607–628.

[Gent 2005] A. N. Gent, "Elastic instabilities in rubber", *Int. J. Nonlinear Mech.* **40** (2005), 165–175.

[Gent and Cho 1999] A. N. Gent and I. S. Cho, "Surface instabilities in compressed or bent rubber blocks", *Rubber Chem. Technol.* **72** (1999), 253–262.

[Haughton 1999] D. M. Haughton, "Flexure and compression of incompressible elastic plates", *Internat. J. Engrg. Sci.* **37**:13 (1999), 1693–1708.

[Haughton and Orr 1995] D. M. Haughton and A. Orr, "On the eversion of incompressible elastic cylinders", *Int. J. Nonlinear Mech.* **30** (1995), 81–95.

[Hill and Hutchinson 1975] R. Hill and J. W. Hutchinson, "Bifurcation phenomena in the plane tension test", *J. Mech. Phys. Solids* **23**:4-5 (1975), 239–264.

[Lindsay 1992] K. A. Lindsay, "The application of compound matrices to convection problems in multi-layered continua", *Math. Models Methods Appl. Sci.* **2**:2 (1992), 121–141.

[Lindsay and Rooney 1992] K. A. Lindsay and C. E. Rooney, "A note on compound matrices", *J. Comput. Phys.* **103**:2 (1992), 472–477.

[Ng and Reid 1979a] B. S. Ng and W. H. Reid, "An initial value method for eigenvalue problems using compound matrices", *J. Comput. Phys.* **30**:1 (1979), 125–136.

[Ng and Reid 1979b] B. S. Ng and W. H. Reid, "A numerical method for linear two-point boundary value problems using compound matrices", *J. Comput. Phys.* **33**:1 (1979), 70–85.

[Ng and Reid 1985] B. S. Ng and W. H. Reid, "The compound matrix method for ordinary differential systems", *J. Comput. Phys.* **58**:2 (1985), 209–228.

[Rivlin 1949] R. S. Rivlin, "Large elastic deformations of isotropic materials, V: The problem of flexure", *Proc. Roy. Soc. London. Ser. A.* **195** (1949), 463–473.

[Roccabianca 2011] S. Roccabianca, *Multilayered structures under large bending: finite solution and bifurcation analysis*, PhD Thesis, University of Trento, 2011.

[Roccabianca et al. 2010] S. Roccabianca, M. Gei, and D. Bigoni, "Plane strain bifurcations of elastic layered structures subject to finite bending: theory versus experiments", *IMA J. Appl. Math.* **75**:4 (2010), 525–548.

[Triantafyllidis 1980] N. Triantafyllidis, "Bifurcation phenomena in pure bending", *J. Mech. Phys. Solids* **28**:3-4 (1980), 221–245.

[Yiantsios and Higgins 1988] S. G. Yiantsios and B. G. Higgins, "Numerical solution of eigenvalue problems using the compound matrix method", *J. Comput. Phys.* **74**:1 (1988), 25–40.

SARA ROCCABIANCA: sara.roccabianca@yale.edu
Department of Biomedical Engineering, Yale University, New Haven, CT 06511, United States

DAVIDE BIGONI: bigoni@ing.unitn.it
University of Trento, Department of Mechanical and Structural Engineering, Via Mesiano 77, I-38123 Trento, Italy

MASSIMILIANO GEI: mgei@ing.unitn.it
University of Trento, Department of Mechanical and Structural Engineering, Via Mesiano 77, I-38123 Trento, Italy

msp

ADVANTAGES OF FORMULATING EVOLUTION EQUATIONS FOR ELASTIC-VISCOPLASTIC MATERIALS IN TERMS OF THE VELOCITY GRADIENT INSTEAD OF THE SPIN TENSOR

M. B. Rubin and O. Papes

Dedicated to Charles and Marie-Louise Steele, who advanced the field of mechanics with their wise editorial leadership

Evolution equations for tensors that characterize elastic-viscoplastic materials are often formulated in terms of a Jaumann derivative based on the spin tensor. Typically, numerical integration algorithms for such equations split the integration operation by first calculating the response due to rate of deformation, followed by a finite rotation. Invariance under superposed rigid body motions of algorithms, incremental objectivity and strong objectivity are discussed. Specific examples of steady-state simple shear at constant rate and steady-state isochoric extension relative to a rotating coordinate system are used to analyze the robustness and accuracy of different algorithms. The results suggest that it is preferable to reformulate evolution equations in terms of the velocity gradient instead of the spin tensor, since strongly objective integration algorithms can be developed using the relative deformation gradient. Moreover, this relative deformation gradient can be calculated independently of the time dependence of the velocity gradient during a typical time step.

1. Introduction

Evolution equations with finite rotations occur naturally in continuum mechanics when history-dependent variables are expressed in terms of the present deforming configuration. Researchers in continuum mechanics typically focus attention on integrating rotations and use a representation of the rotation tensor attributed to Euler and Rodrigues; see, e.g., [Argyris 1982; Simo and Vu-Quoc 1988; Argyris and Poterasu 1993; Govindjee 1997; Becker 2006; Rubin 2007]. One of the objectives of this paper is to discuss fundamental and practical reasons for considering evolution equations based on the velocity gradient instead of on the spin tensor.

To be more specific, it is recalled that within the context of the three-dimensional theory the velocity gradient $\boldsymbol{L}$ separates into a symmetric rate of deformation tensor $\boldsymbol{D}$ and a skew-symmetric spin tensor $\boldsymbol{W}$, such that

$$\boldsymbol{L} = \boldsymbol{D} + \boldsymbol{W}, \quad \boldsymbol{D} = \tfrac{1}{2}(\boldsymbol{L} + \boldsymbol{L}^T), \quad \boldsymbol{W} = \tfrac{1}{2}(\boldsymbol{L} - \boldsymbol{L}^T). \tag{1-1}$$

Within the context of hypoelastic formulations of the elastic response for elastic-viscoplastic materials it is common to propose a constitutive equation for the time rate of change of Cauchy stress $\boldsymbol{T}$. For example, a typical constitutive structure based on the Jaumann derivative of stress suggests that

$$\overset{W}{\boldsymbol{T}} = \dot{\boldsymbol{T}} - \boldsymbol{W}\boldsymbol{T} - \boldsymbol{T}\boldsymbol{W}^T = \hat{\boldsymbol{K}}_W(\boldsymbol{T}, \boldsymbol{D}), \tag{1-2}$$

where a superposed dot denotes material time differentiation and the function $\hat{K}_W(T, D)$ characterizes elastic and inelastic rates of deformation.

Integration of (1-2) over the time period $t_1 \leq t \leq t_2$ with time increment $\triangle t = t_2 - t_1$ is usually performed in two steps. First, the initial value $T(t_1)$ of stress is augmented by an increment of stress that is determined by the quantity $\hat{K}_W$ and then the result is rotated with an orthogonal tensor associated with the spin W. This procedure can be formalized by introducing a proper orthogonal tensor Λ_W defined by the evolution equation and initial condition

$$\dot{\Lambda}_W = W \Lambda_W, \quad \Lambda_W(t_1) = I, \tag{1-3}$$

and by introducing the auxiliary tensor $\overline{T}_W$ defined by

$$T = \Lambda_W \overline{T}_W \Lambda_W^T, \tag{1-4}$$

which satisfies the evolution equation and initial condition

$$\dot{\overline{T}}_W = \Lambda_W^T \hat{K}_W(T, D) \Lambda_W, \quad \overline{T}_W(t_1) = T(t_1). \tag{1-5}$$

In particular, this procedure requires two approximations, one for the integral of the stress (1-5) and another for the integral of the rotation (1-3). Hughes and Winget [1980] developed an approximate solution of (1-2), which will be discussed in Section 2. Also, for constant spin, (1-3) can be solved exactly in terms of the exponential map; see, e.g., [Govindjee 1997].

The Jaumann derivative is one of a number of stress rates which is properly invariant under superposed rigid body motions (SRBM) and which has been used in the literature to develop evolution equations for stress. For example, Dienes [1979] used the skew-symmetric tensor associated with the material derivative of the rotation tensor R in the polar decomposition (see [Malvern 1969], for instance) of the total deformation gradient F. Also, Rashid [1993] used the skew-symmetric tensor associated with the material derivative of the rotation tensor R_r in the polar decomposition of the relative deformation gradient F_r.

The same material response as that characterized by (1-2) can be obtained using different stress rates as long as the term on the right-hand side of (1-2) is modified appropriately. For example, the Oldroyd rate (see, e.g., [Holzapfel 2000]) can be used to obtain the evolution equation

$$\overset{L}{T} = \dot{T} - LT - TL^T = \hat{K}_L(T, D), \tag{1-6}$$

where the function $\hat{K}_L$ is defined by

$$\hat{K}_L(T, D) = \hat{K}_W(T, D) - DT - TD. \tag{1-7}$$

Next, it is recalled that the relative deformation gradient F_r can be defined in terms of the total deformation gradient by the expression

$$F_r(t) = F(t) F^{-1}(t_1). \tag{1-8}$$

Moreover, it can be shown that F_r satisfies the evolution equation and initial condition

$$\dot{F}_r = LF_r, \quad F_r(t_1) = I. \tag{1-9}$$

It then follows that the solution of (1-6) can be written in the form

$$\boldsymbol{T} = \boldsymbol{F}_r \overline{\boldsymbol{T}}_L \boldsymbol{F}_r^T, \tag{1-10}$$

where the auxiliary tensor $\overline{\boldsymbol{T}}_L$ satisfies the evolution equation and initial condition

$$\dot{\overline{\boldsymbol{T}}}_L = \boldsymbol{F}_r^{-1} \hat{\boldsymbol{K}}_L(\boldsymbol{T}, \boldsymbol{D}) \boldsymbol{F}_r^{-T}, \quad \overline{\boldsymbol{T}}_L(t_1) = \boldsymbol{T}(t_1). \tag{1-11}$$

Since the material response characterized by (1-2) can be formulated using different invariant stress rates, there is no fundamental advantage of one properly invariant formulation over another. However, from a practical point of view, details of the approximate integration algorithms for $\{\boldsymbol{\Lambda}_W, \overline{\boldsymbol{T}}_W\}$ or $\{\boldsymbol{F}_r, \overline{\boldsymbol{T}}_L\}$ may present advantages due to accuracy or invariance under SRBM.

Another important consideration was discussed in [Rashid 1993]. In computer codes for integrating the equations of motion in continuum mechanics the positions of material points are known at the beginning of the time step and are determined at the end of the time step. For implicit integration algorithms the final positions are determined by iteration, whereas for explicit integration algorithms they are determined by estimations based on the positions, velocities and accelerations during the time step. Rashid argued that the time dependences of the rate of deformation tensor $\boldsymbol{D}$ and spin tensor $\boldsymbol{W}$ during the time step are never known. This means that integration of the Jaumann formulation necessarily requires approximations associated with the specification of the time dependences of $\boldsymbol{D}$ and $\boldsymbol{W}$ as well as additional approximations due to the specific integration algorithms for both the rotation (1-3) and stress (1-5).

In contrast, the value $\boldsymbol{F}_r(t_2)$ of the relative deformation gradient at the end of the time step is a unique function of the positions at the beginning and end of the time step and therefore is known. This means that the evolution equation (1-9) can be integrated exactly without assuming an approximation of the velocity gradient. Consequently, the only approximation associated with the Oldroyd formulation appears in the specified integration algorithm for the auxiliary stress value $\overline{\boldsymbol{T}}_L$ in (1-11). For this reason the Oldroyd seems to have a practical advantage over the Jaumann formulation.

Eckart [1948] seems to be the first to have proposed an evolution equation directly for elastic deformation for large deformations of elastically isotropic elastic-plastic materials. In particular, using [Flory 1961] it is possible to introduce a symmetric unimodular tensor $\boldsymbol{B}'_e$,

$$\det \boldsymbol{B}'_e = 1, \tag{1-12}$$

which is a pure measure of elastic distortional deformation. Then, for elastic-viscoplastic response $\boldsymbol{B}'_e$ can be determined by the evolution equation

$$\dot{\boldsymbol{B}}'_e = \boldsymbol{L}\boldsymbol{B}'_e + \boldsymbol{B}'_e \boldsymbol{L}^T - \tfrac{2}{3}(\boldsymbol{D} \cdot \boldsymbol{I})\boldsymbol{B}'_e - \Gamma \boldsymbol{A}_p, \quad \boldsymbol{A}_p = \boldsymbol{B}'_e - \frac{3}{\boldsymbol{B}'^{-1}_e \cdot \boldsymbol{I}}\boldsymbol{I}, \tag{1-13}$$

where Γ and $\boldsymbol{A}_p$ characterize the rate of relaxation due to plasticity. This equation must be integrated subject to the initial condition

$$\boldsymbol{B}'_e = \boldsymbol{B}'_e(t_1) \quad \text{for } t = t_1. \tag{1-14}$$

This idea was used in [Leonov 1976] for polymeric liquids, and in [Simo 1992; Rubin 1994] for elastic-plastic and elastic-viscoplastic solids. Besseling [1968] and Rubin [1994] proposed generalizations of (1-13) for elastically anisotropic response and a number of physical aspects of constitutive equations for

plasticity have been discussed in [Rubin 1994; Rubin 1996; Rubin 2001]. Also, a simple integration algorithm for (1-13) was discussed in [Rubin and Attia 1996].

Again, using [Flory 1961], it can be shown that the unimodular part F'_r of the relative deformation tensor F_r defined by

$$F'_r = (\det F_r)^{-1/3} F_r, \quad \det F'_r = 1, \tag{1-15}$$

satisfies the evolution equation and initial condition

$$\dot{F}'_r = L F'_r - \tfrac{1}{3}(D \cdot I) F'_r, \quad F'_r(t_1) = I. \tag{1-16}$$

Consequently, the solution of (1-13) can be written in terms of the auxiliary unimodular symmetric tensor $\bar{B}'_e$ defined by

$$B'_e = F'_r \bar{B}'_e F'^T_r, \quad \det \bar{B}'_e = 1, \tag{1-17}$$

with $\bar{B}'_e$ satisfying the evolution equation and initial condition

$$\dot{\bar{B}}'_e = -\Gamma F'^{-1}_r A_p F'^{-T}_r, \quad \bar{B}'_e(t_1) = B'_e(t_1). \tag{1-18}$$

In general, Γ is a nonlinear function of state variables that can include hardening. Also, it is noted that for an elastically isotropic hyperelastic material the Cauchy stress can be obtained from derivatives of a strain energy function that depends on the dilatation $J = \det F$ and the two nontrivial invariants of B'_e.

An outline of the paper is as follows. Section 2 summarizes the Hughes–Winget algorithm [Hughes and Winget 1980], Section 3 discusses invariance under SRBM and the notion of incremental objectivity. Section 4 presents integration algorithms for a simple elastic-viscoplastic material. The robustness and accuracy of these algorithms are discussed in Section 5 for the example of steady-state simple shear and in Section 6 for the example of steady-state isochoric extension relative to a rotating coordinate system. Finally, Section 7 presents conclusions.

Above and in the following, boldfaced symbols denote tensors, e_i is a right-handed orthonormal triad of fixed rectangular Cartesian base vectors, the components of all tensors are referred to e_i and $\otimes$ denotes the tensor product. Also, $A \cdot B = \text{tr}(A B^T)$ denotes the inner product between two second-order tensors $\{A, B\}$ and $a \otimes b$ denotes the tensor product between two vectors $\{a, b\}$.

2. Summary of the Hughes–Winget algorithm

Hughes and Winget [1980] developed an expression for the value $\Lambda_W(t_2)$ of a rotation tensor based on the skew-symmetric part of an approximate incremental displacement gradient. Specifically, they introduced a mapping from the position y^1 at the beginning of the time step to the position y^2 at the end of the time step and a mapping to the position y^α at an intermediate configuration by the expressions

$$y^2 = y^2(y^1), \quad y^\alpha = (1 - \alpha) y^1 + \alpha y^2. \tag{2-1}$$

Then, the incremental displacement δ, and displacement gradient G relative to this intermediate configuration are defined by

$$\delta = y^2 - y^1, \quad G = \partial \delta / \partial y^\alpha = (\partial y^2 / \partial y^1 - I)(\partial y^1 / \partial y^\alpha). \tag{2-2}$$

Moreover, with the help of (2-1) it follows that

$$\partial \boldsymbol{y}^{\alpha}/\partial \boldsymbol{y}^1 = (1-\alpha)\boldsymbol{I} + \alpha(\partial \boldsymbol{y}^2/\partial \boldsymbol{y}^1). \tag{2-3}$$

Next, using the definition of the relative deformation gradient $\boldsymbol{F}_r(t_2)$ at the end of the time step

$$\boldsymbol{F}_r(t_2) = \partial \boldsymbol{y}^2/\partial \boldsymbol{y}^1, \tag{2-4}$$

the tensor $\boldsymbol{G}$ can be expressed in the form

$$\boldsymbol{G} = \big(\boldsymbol{F}_r(t_2) - \boldsymbol{I}\big)\big(\boldsymbol{I} + \alpha\{\boldsymbol{F}_r(t_2) - \boldsymbol{I}\}\big)^{-1}. \tag{2-5}$$

Thus, for $\alpha = \frac{1}{2}$ as specified in [Hughes and Winget 1980], the tensor $\boldsymbol{G}$ is given by

$$\boldsymbol{G} = 2\big(\boldsymbol{F}_r(t_2) - \boldsymbol{I}\big)\big(\boldsymbol{F}_r(t_2) + \boldsymbol{I}\big)^{-1}. \tag{2-6}$$

Next, using the lemma quoted in [Hughes and Winget 1980],

$$\boldsymbol{A}(\boldsymbol{A}+\boldsymbol{B})^{-1}\boldsymbol{B} = \boldsymbol{B}(\boldsymbol{A}+\boldsymbol{B})^{-1}\boldsymbol{A}, \tag{2-7}$$

for all square nonsingular matrices $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{A}+\boldsymbol{B}$. Taking $\boldsymbol{A} = \boldsymbol{F}_r^T$ and $\boldsymbol{B} = \boldsymbol{I}$ it follows that

$$\boldsymbol{F}_r^T(\boldsymbol{F}_r^T+\boldsymbol{I})^{-1} = (\boldsymbol{F}_r^T+\boldsymbol{I})^{-1}\boldsymbol{F}_r^T, \quad (\boldsymbol{F}_r^T-\boldsymbol{I})(\boldsymbol{F}_r^T+\boldsymbol{I})^{-1} = (\boldsymbol{F}_r^T+\boldsymbol{I})^{-1}(\boldsymbol{F}_r^T-\boldsymbol{I}), \tag{2-8}$$

so that

$$\boldsymbol{G}^T = 2(\boldsymbol{F}_r^T+\boldsymbol{I})^{-1}(\boldsymbol{F}_r^T-\boldsymbol{I}) = 2(\boldsymbol{F}_r^T-\boldsymbol{I})(\boldsymbol{F}_r^T+\boldsymbol{I})^{-1}. \tag{2-9}$$

Thus, the symmetric part $\boldsymbol{\gamma}$ and skew-symmetric part $\boldsymbol{\omega}$ of $\boldsymbol{G}$ are given by

$$\begin{aligned} \boldsymbol{\gamma} &= \tfrac{1}{2}(\boldsymbol{G}+\boldsymbol{G}^T) = \big(\boldsymbol{F}_r(t_2)-\boldsymbol{I}\big)\big(\boldsymbol{F}_r(t_2)+\boldsymbol{I}\big)^{-1} + \big(\boldsymbol{F}_r^T(t_2)-\boldsymbol{I}\big)\big(\boldsymbol{F}_r^T(t_2)+\boldsymbol{I}\big)^{-1}, \\ \boldsymbol{\omega} &= \tfrac{1}{2}(\boldsymbol{G}-\boldsymbol{G}^T) = \big(\boldsymbol{F}_r(t_2)-\boldsymbol{I}\big)\big(\boldsymbol{F}_r(t_2)+\boldsymbol{I}\big)^{-1} - \big(\boldsymbol{F}_r^T(t_2)-\boldsymbol{I}\big)\big(\boldsymbol{F}_r^T(t_2)+\boldsymbol{I}\big)^{-1}. \end{aligned} \tag{2-10}$$

Now, the Hughes–Winget algorithm for the value $\boldsymbol{\Lambda}_W(t_2)$ of the rotation tensor $\boldsymbol{\Lambda}_W$ at the end of the time step is given by

$$\boldsymbol{\Lambda}_W(t_2) = \big(\boldsymbol{I}+\tfrac{1}{2}\boldsymbol{\omega}\big)\big(\boldsymbol{I}-\tfrac{1}{2}\boldsymbol{\omega}\big)^{-1}. \tag{2-11}$$

3. Invariance under superposed rigid body motions and incremental objectivity

Under SRBM the material point $\boldsymbol{x}$ at time t moves to the position $\boldsymbol{x}^+$ at time t^+, such that

$$\boldsymbol{x}^+ = \boldsymbol{c}(t) + \boldsymbol{Q}(t)\boldsymbol{x}, \quad t^+ = t+c, \quad \boldsymbol{Q}\boldsymbol{Q}^T = \boldsymbol{I}, \quad \det \boldsymbol{Q} = +1, \quad \dot{\boldsymbol{Q}} = \boldsymbol{\Omega}\boldsymbol{Q}, \quad \boldsymbol{\Omega}^T = -\boldsymbol{\Omega}(t), \tag{3-1}$$

where c is an arbitrary constant, $\boldsymbol{c}(t)$ is an arbitrary function of time characterizing superposed translation, $\boldsymbol{Q}(t)$ is an arbitrary proper orthogonal tensor characterizing rotation and $\boldsymbol{\Omega}(t)$ is the associated spin tensor. Moreover, under SRBM the quantities $\{\boldsymbol{F}, \boldsymbol{F}_r, \boldsymbol{F}_r', \boldsymbol{D}, \boldsymbol{W}, \boldsymbol{\Lambda}_W, \boldsymbol{T}, \boldsymbol{B}_e', \Gamma\}$ transform to $\{\boldsymbol{F}^+, \boldsymbol{F}_r^+, \boldsymbol{F}_r'^+, \boldsymbol{D}^+, \boldsymbol{W}^+, \boldsymbol{\Lambda}_W^+, \boldsymbol{T}^+, \boldsymbol{B}_e'^+, \Gamma^+\}$, respectively, by the relations

$$\begin{aligned} \boldsymbol{F}^+ &= \boldsymbol{Q}\boldsymbol{F}, & \boldsymbol{F}_r^+ &= \boldsymbol{Q}\boldsymbol{F}_r, & \boldsymbol{F}_r'^+ &= \boldsymbol{Q}\boldsymbol{F}_r', \\ \boldsymbol{D}^+ &= \boldsymbol{Q}\boldsymbol{D}\boldsymbol{Q}^T, & \boldsymbol{W}^+ &= \boldsymbol{Q}\boldsymbol{W}\boldsymbol{Q}^T+\boldsymbol{\Omega}, & \boldsymbol{\Lambda}_W^+ &= \boldsymbol{Q}\boldsymbol{\Lambda}_W, \\ \boldsymbol{T}^+ &= \boldsymbol{Q}\boldsymbol{T}\boldsymbol{Q}^T, & \boldsymbol{B}_e'^+ &= \boldsymbol{Q}\boldsymbol{B}_e'\boldsymbol{Q}^T, & \Gamma^+ &= \Gamma. \end{aligned} \tag{3-2}$$

Since the rates in (1-2) and (1-6) transform under SRBM by

$$\overset{W}{T}{}^{+} = Q\overset{W}{T}Q^{T}, \quad \overset{L}{T}{}^{+} = Q\overset{L}{T}Q^{T}, \tag{3-3}$$

it follows that the evolution equations (1-2) and (1-6) are properly invariant under SRBM provided that the functions $\hat{K}_W$ and $\hat{K}_L$ satisfy the restrictions

$$\hat{K}_W(QTQ^T, QDQ^T) = Q\hat{K}_W(T, D)Q^T, \quad \hat{K}_L(QTQ^T, QDQ^T) = Q\hat{K}_L(T, D)Q^T, \tag{3-4}$$

for all proper orthogonal Q. Moreover, the evolution equation (1-13) is properly invariant under SRBM.

Using these results together with the definitions (1-4), (1-10) and (1-17), it can also be shown that the auxiliary tensors $\{\overline{T}_W, \overline{T}_L, \overline{B}'_e\}$ transform to $\{\overline{T}_W^+, \overline{T}_L^+, \overline{B}'^+_e\}$, so they are unaffected by SRBM:

$$\overline{T}_W^+ = \overline{T}_W, \quad \overline{T}_L^+ = \overline{T}_L, \quad \overline{B}'^+_e = \overline{B}'_e. \tag{3-5}$$

Hughes and Winget [1980] introduced the notion of an algorithm being incrementally objective. Specifically, they proved that the expressions (2-10) and (2-11) are incrementally objective in the sense that they correctly produce zero strain increment ($\gamma = 0$) and the correct rotation tensor ($\Lambda_W = F_r$) when the deformation during the time step is a pure rotation with F_r being an orthogonal tensor ($F_r^T F_r = I$). It can also be shown that (2-10) and (2-11) correctly produce zero incremental spin ($\omega = 0$) and the correct rotation ($\Lambda_W = I$) when the deformation during the time step is a pure stretch with F_r being a symmetric tensor.

Rashid [1993] extended this notion of incremental objectivity by demanding that the integrator computes a stretching part that is independent of the input rotation when the incremental motion involves both stretch and rotation. Actually, the notion of strong objectivity, as discussed in [Papes and Mazza 2009], requires the estimates of all variables at the end of the time step to satisfy the same invariance properties under SRBM as their exact values. In this regard, it is noted that the tensors $\{\gamma, \omega\}$ are not properly invariant under SRBM since they retain an unphysical dependence on the arbitrary rotation tensor Q:

$$\gamma^+ = Q\big[(F_r(t_2)Q - I)(F_r(t_2)Q + I)^{-1} + (Q^T F_r^T(t_2) - I)(Q^T F_r^T(t_2) + I)^{-1}\big]Q^T,$$
$$\omega^+ = Q\big[(F_r(t_2)Q - I)(F_r(t_2)Q + I)^{-1} - (Q^T F_r^T(t_2) - I)(Q^T F_r^T(t_2) + I)^{-1}\big]Q^T. \tag{3-6}$$

Consequently, the approximation (2-11) for Λ_W is also not properly invariant:

$$\Lambda_W^+ \neq Q\Lambda_W. \tag{3-7}$$

Moreover, it is natural to consider the approximation of incremental strain

$$\Delta t\, D \approx \gamma. \tag{3-8}$$

However, this expression is also not properly invariant since

$$\Delta t\, D^+ \neq Q(\Delta t\, D)Q^T. \tag{3-9}$$

4. Integration algorithms for a simple elastic-viscoplastic material

In the remainder of this paper attention will be focused on an elastic-viscoplastic material which is characterized by the equations (1-13), (1-14), (1-17) and (1-18) with the simplification that the scalar Γ is constant. In this section three algorithms are discussed which yield approximate solutions of these evolution equations. Sections 5 and 6 will discuss examples to test the accuracy and robustness of these algorithms.

Algorithm 1. This algorithm is based on the formulation (1-17) and (1-18). This formulation has the simplicity that any approximation of $\bar{\boldsymbol{B}}'_e$ that is unaffected by SRBM will yield a solution for $\boldsymbol{B}'_e$ that transforms appropriately under SRBM. To motivate an approximate solution of (1-18) consider the value determined by the fully implicit equation

$$\bar{\boldsymbol{B}}'_e(t_2) = \bar{\boldsymbol{B}}'_e(t_1) - \Delta t \, \Gamma\Big(\bar{\boldsymbol{B}}'_e(t_2) - \frac{3}{\boldsymbol{B}'^{-1}_e(t_2)\cdot\boldsymbol{I}}\boldsymbol{C}'^{-1}_r(t_2)\Big), \quad \boldsymbol{C}'^{-1}_r(t_2) = \boldsymbol{F}'^{-1}_r(t_2)\boldsymbol{F}'^{-T}_r(t_2). \tag{4-1}$$

Next, use is made of the approximation that $\boldsymbol{B}'^{-1}_e(t_2) \cdot \boldsymbol{I} \approx 3$ to deduce that

$$\bar{\boldsymbol{B}}'_e(t_2) = \frac{1}{1+\Delta t\,\Gamma}\big(\boldsymbol{B}'_e(t_1) + \Delta t\,\Gamma\boldsymbol{C}'^{-1}_r(t_2)\big). \tag{4-2}$$

However, this expression does not ensure that $\bar{\boldsymbol{B}}'_e(t_2)$ is unimodular. Motivated by [Rubin and Attia 1996], the expression (4-2) is used to obtain an equation for the deviatoric part $\bar{\boldsymbol{B}}''_e(t_2)$ of $\bar{\boldsymbol{B}}'_e(t_2)$ of the form

$$\bar{\boldsymbol{B}}''_e(t_2) = \frac{1}{1+\Delta t\,\Gamma}\Big(\boldsymbol{B}'_e(t_1) + \Delta t\,\Gamma\boldsymbol{C}'^{-1}_r(t_2) - \tfrac{1}{3}\big[\big(\boldsymbol{B}'_e(t_1) + \Delta t\,\Gamma\boldsymbol{C}'^{-1}_r(t_2)\big)\cdot\boldsymbol{I}\big]\boldsymbol{I}\Big). \tag{4-3}$$

Then, the final value $\bar{\boldsymbol{B}}'_e(t_2)$ can be obtained using the procedure discussed in [Rubin and Attia 1996] to determine the scalar $\bar{\alpha}$, in the expression

$$\bar{\boldsymbol{B}}'_e(t_2) = \tfrac{1}{3}\bar{\alpha}\boldsymbol{I} + \bar{\boldsymbol{B}}''_e(t_2), \tag{4-4}$$

by the condition that $\bar{\boldsymbol{B}}'_e(t_2)$ is unimodular.

Algorithm 2. Typically, for elastic-viscoplastic response an equation like (1-13) is integrated in two steps. First, the elastic trial value $\bar{\boldsymbol{B}}'^*_e(t)$ is determined by the evolution equation (which is (1-13) with $\Gamma = 0$) and initial condition

$$\dot{\boldsymbol{B}}'^*_e = \boldsymbol{L}\boldsymbol{B}'^*_e + \boldsymbol{B}'^*_e\boldsymbol{L}^T - \tfrac{2}{3}(\boldsymbol{D}\cdot\boldsymbol{I})\boldsymbol{B}'^*_e, \quad \boldsymbol{B}'^*_e(t_1) = \boldsymbol{B}'_e(t_1). \tag{4-5}$$

Then, the value $\boldsymbol{B}'_e(t_2)$ at the end of the time step is determined by relaxing the elastic trial value $\boldsymbol{B}'^*_e(t_2)$ at constant total deformation with the help of an approximation of the evolution equation (1-13).

In particular, using the formulation (1-17) and (1-18) it is easy to see that the exact solution of (4-5) is given by

$$\boldsymbol{B}'^*_e(t_2) = \boldsymbol{F}'_r(t_2)\boldsymbol{B}'_e(t_1)\boldsymbol{F}'^T_r(t_2), \tag{4-6}$$

so that the solution of (1-13) can be written in the form

$$\boldsymbol{B}'_e(t_2) = \boldsymbol{B}'^*_e(t_2) - \boldsymbol{F}'_r(t_2)\Big(\int_{t_1}^{t_2}\Gamma\boldsymbol{F}'^{-1}_r\boldsymbol{A}_p\boldsymbol{F}'^{-T}_r\,dt\Big)\boldsymbol{F}'^T_r(t_2). \tag{4-7}$$

An elastic trial of the type (4-6) was used in [Simo 1992] for a nonunimodular tensor and in [Simo and Hughes 1998, p. 315] updating a unimodular tensor. In approximating the integral in (4-7) it is important to ensure that the approximation remains properly invariant under SRBM. Here, the values of $\{A_p, F_r'\}$ are approximated by their values at the end of the time step and use is made of the fully implicit form

$$B_e'(t_2) = B_e'^*(t_2) - \Delta t\, \Gamma \left(B_e'(t_2) - \frac{3}{B_e'^{-1}(t_2) \cdot I}\, I \right). \tag{4-8}$$

Again, since this result does not ensure that $B_e'(t_2)$ is unimodular, the solution is determined by the deviatoric part $B_e''(t_2)$ of $B_e'(t_2)$ given by

$$B_e''(t_2) = \frac{1}{1 + \Delta t\, \Gamma}\, B_e''^*(t_2), \tag{4-9}$$

where $B_e''^*(t_2)$ is the deviatoric part of the elastic trial $B_e'^*(t_2)$

$$B_e''^*(t_2) = B_e'^*(t_2) - \tfrac{1}{3}\big(B_e'^*(t_2) \cdot I\big)\, I. \tag{4-10}$$

Then, the final value $B_e'(t_2)$ is obtained using the procedure discussed in [Rubin and Attia 1996] to determine the scalar α in the expression

$$B_e'(t_2) = \tfrac{1}{3}\alpha I + B_e''(t_2), \tag{4-11}$$

which has the same invariance properties under SRBM as the exact value of $B_e'(t_2)$.

Algorithm 3. Algorithms 1 and 2 integrate the evolution equation (1-13) including coupling of the rates of deformation and spin $\{D, W\}$ through the expressions for the velocity gradient L and the relative deformation gradient F_r. For Algorithm 3 this evolution equation is reformulated in terms of the Jaumann derivative which focuses on spin. Specifically, (1-13) can be written in the form

$$\overset{W}{B}{}_e' = \hat{A}(B_e', D) - \Gamma A_p, \quad \hat{A}(B_e', D) = DB_e' + B_e'D - \tfrac{2}{3}(D \cdot I)B_e'. \tag{4-12}$$

Then, the solution is given by (4-9) and (4-10), where the elastic trial $B_e'^*(t_2)$ is determined by the evolution equation and initial condition

$$\overset{W}{B}{}_e'^* = \hat{A}(B_e'^*, D), \quad B_e'^*(t_1) = B_e'(t_1). \tag{4-13}$$

Next, using the Hughes–Winget algorithm the solution of (4-13) is approximated by

$$B_e'^*(t_2) = \Lambda_W\big(B_e'(t_1) + \Delta t\, \hat{A}(B_e'(t_1), D)\big)\Lambda_W^T, \tag{4-14}$$

where Λ_W is defined by (2-11). Furthermore, use is made of the approximation (3-8) to obtain the deviatoric tensor

$$B_e''^*(t_2) = \Lambda_W\big(\tilde{B}_e' - \tfrac{1}{3}(\tilde{B}_e' \cdot I)I\big)\Lambda_W^T, \quad \tilde{B}_e' = B_e'(t_1) + \gamma B_e'(t_1) + B_e'(t_1)\gamma - \tfrac{2}{3}(\gamma \cdot I)B_e'(t_1), \tag{4-15}$$

which is then used in (4-9) to obtain the final value (4-11) for this algorithm.

5. Example of steady-state simple shear at constant shear rate

The algebra required to obtain the solutions discussed in the example in this and the next sections is rather heavy so use has been made of the symbolic program Maple to derive the results.

This section presents the example of steady-state simple shear at constant shear rate to analytically analyze the robustness and accuracy of Algorithms 1, 2 and 3. For this problem the deformation gradient F and the constant velocity gradient L are specified by

$$F = I + \gamma \Gamma t (e_1 \otimes e_2), \quad L = \gamma \Gamma (e_1 \otimes e_2), \tag{5-1}$$

where the constant scalar γ should not be confused with the tensor $\boldsymbol{\gamma}$ in (2-10). It then follows that the relative deformation tensor associated with (5-1) is given by

$$F_r(t_2) = I + \kappa (e_1 \otimes e_2), \quad F_r^{-1}(t_2) = I - \kappa (e_1 \otimes e_2), \quad \kappa = \Delta t \, \gamma \Gamma. \tag{5-2}$$

Moreover, since this deformation is isochoric, F_r is a unimodular tensor with

$$F_r' = F_r. \tag{5-3}$$

Exact solution. The initial value $B_e'(t_1)$ is a steady-state solution of the evolution equation (1-13) provided that it satisfies the equation

$$L B_e'(t_1) + B_e'(t_1) L^T - \tfrac{2}{3} (D \cdot I) B_e'(t_1) - \Gamma A_p(t_1) = 0. \tag{5-4}$$

The exact solution of this algebraic equation can be written in the form

$$B_e'(t_1) = a^2 (e_1 \otimes e_1) + b^2 (e_2 \otimes e_2) + c^2 (e_3 \otimes e_3) + d(e_1 \otimes e_2 + e_2 \otimes e_1), \tag{5-5}$$

where $\{a, c\}$ take the forms

$$a = \frac{\sqrt{1 + d^2 b^2}}{b^2}, \quad c = b, \tag{5-6}$$

and the constants $\{b, d\}$ attain the values $\{b_e, d_e\}$, respectively, given by

$$b_e = \frac{1}{(1 + \gamma^2)^{1/6}}, \quad d_e = \frac{\gamma}{(1 + \gamma^2)^{1/3}}. \tag{5-7}$$

Solution of Algorithm 1. The initial value $B_e'(t_1)$ is a steady-state solution of Algorithm 1 if $\bar{B}_e'(t_2)$ and its deviatoric part $\bar{B}_e''(t_2)$ are given by

$$\bar{B}_e'(t_2) = F_r'^{-1}(t_2) B_e'(t_1) F_r'^{-T}(t_2), \quad \bar{B}_e''(t_2) = \bar{B}_e'(t_2) - \tfrac{1}{3} \big(\bar{B}_e'(t_2) \cdot I \big) I. \tag{5-8}$$

More specifically, substitution of (5-8) into (4-3) and taking $B_e'(t_1)$ in the form (5-5), with the condition (5-6), yields a system of algebraic equations for $\{b, d\}$. These equations can be solved to obtain the values $\{b_1, d_1\}$, given by

$$d_1 = (\gamma + \kappa) b_1^2 - \kappa, \tag{5-9}$$

where b_1 is the positive real root of the equation

$$(1 + \gamma^2 + \gamma \kappa) b_1^6 + \kappa^2 b_1^4 - \kappa^2 b_1^2 - 1 = 0. \tag{5-10}$$

Solution of Algorithm 2. The initial value $\boldsymbol{B}_e'(t_1)$ is a steady-state solution of Algorithm 2 if $\boldsymbol{B}_e''(t_2)$ equals the deviatoric part $\boldsymbol{B}_e'(t_1)$, given by

$$\boldsymbol{B}_e''(t_2) = \boldsymbol{B}_e''(t_1) = \boldsymbol{B}_e'(t_1) - \tfrac{1}{3}\big(\boldsymbol{B}_e'(t_1) \cdot \boldsymbol{I}\big)\boldsymbol{I}. \tag{5-11}$$

More specifically, substitution of (5-11) into (4-9) and taking $\boldsymbol{B}_e'(t_1)$ in the form (5-5), with the condition (5-6), yields a system of algebraic equations for $\{b, d\}$. These equations can be solved to obtain the values $\{b_2, d_2\}$ given by

$$b_2 = \frac{1}{(1 + \gamma^2 + \gamma\kappa)^{1/6}}, \quad d_2 = \frac{\gamma}{(1 + \gamma^2 + \gamma\kappa)^{1/3}}. \tag{5-12}$$

Solution of Algorithm 3. The initial value $\boldsymbol{B}_e'(t_1)$ will be a steady-state solution of Algorithm 3 if $\boldsymbol{B}_e''(t_2)$ in (4-9) satisfies (5-11) with the deviatoric part $\boldsymbol{B}_e''^*(t_2)$ of the elastic trial given by (4-15). The solution of this system of equations has the form (5-5), where $\{a, b, c, d\}$ obtain the values $\{a_3, b_3, c_3, d_3\}$ given by

$$a_3 = b_3\sqrt{\frac{256 + 512\gamma^2 + 512\gamma\kappa + 16(2 - \gamma^2)\kappa^2 - 16\gamma\kappa^3 + \kappa^4}{256 + 16(2 + \gamma^2)\kappa^2 + 16\gamma\kappa^3 + \kappa^4}}, \quad b_3 = \left(\frac{N_3(\gamma, \kappa)}{D_3(\gamma, \kappa)}\right)^{1/6},$$

$$d_3 = \frac{(512 - 112\kappa^2 + \kappa^4)b_3^2 - (80 - \kappa^2)\kappa^2 a_3^2}{512 + 256\gamma\kappa + 64\kappa^2 + 2\kappa^4}\gamma, \quad c_3 = \sqrt{\tfrac{1}{2}(a_3^2 + b_3^2) - \gamma d_3}. \tag{5-13}$$

In these expressions the functions N_3 and D_3 are polynomials of their arguments that can be obtained analytically by requiring $\boldsymbol{B}_e'(t_1)$ to be unimodular. Also, it is noted that the values of $\{a, c\}$ no longer satisfy the conditions (5-6) of the exact solution. Since this formulation uses the approximations (2-11) and (3-8) it is not properly invariant under SRBM.

Discussion. In these solutions the parameter γ is a normalized loading rate and the parameter κ is a normalized time increment. Figure 1 plots the exact solution as a function of the loading rate γ. From this figure it can be seen that the elastic distortional deformation is large for large values of γ. It can also be shown that in the limit that κ approaches zero all three algorithms reproduce the exact steady-state values (5-7) for all values of γ. However, for positive values of κ Algorithms 1 and 2 reproduce the exact result that $(c_1 = b_1, c_2 = b_2)$, whereas Algorithm 3 predicts that c_3 is different from b_3.
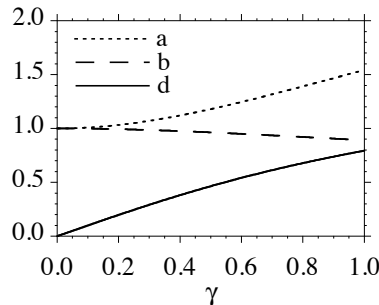
Figure 1. Simple shear: Exact steady-state solution.
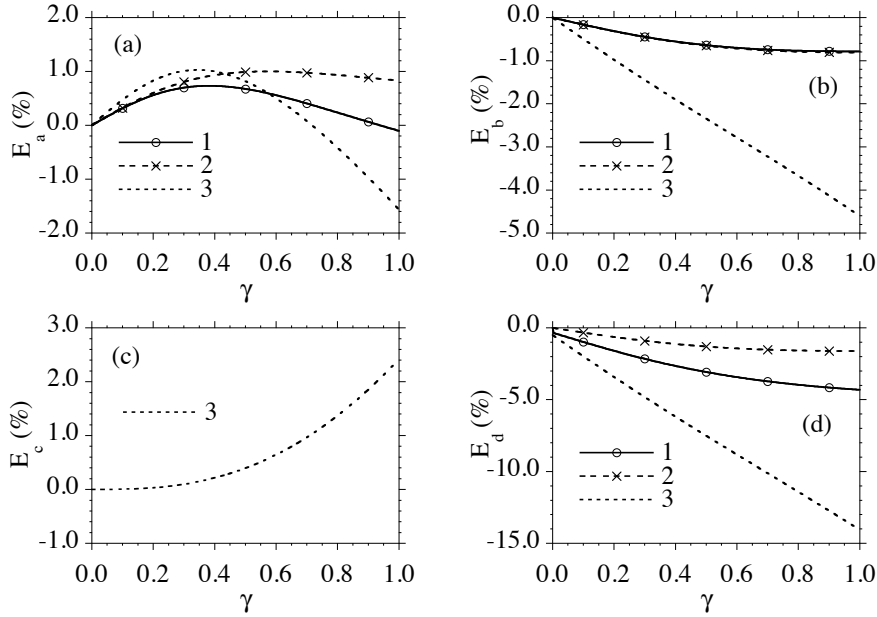
Figure 2. Simple shear: Errors in the solutions predicted by Algorithms 1, 2 and 3 for the normalized time increment $\kappa = 0.1$.

To compare the predictions of the algorithms with the exact solution it is convenient to define the errors

$$E_a = \frac{a}{a_e} - 1, \quad E_b = \frac{b}{b_e} - 1, \quad E_c = \frac{c}{c_e} - 1, \quad E_d = \frac{d}{d_e} - 1, \tag{5-14}$$

where $\{a, b, c, d\}$ are the values predicted by each of the algorithms, with the values of $\{a, c\}$ given by (5-6) for Algorithms 1 and 2. Figure 2 shows these errors as functions of γ for the relatively large normalized time increment $\kappa = 0.1$. From these results it can be seen that Algorithm 1 is slightly more accurate than Algorithm 2 for E_a. On the other hand, Algorithm 2 is more accurate than Algorithm 1 for E_d. Also, both Algorithms 1 and 2 are more accurate than Algorithm 3. Overall, it is concluded that both Algorithms 1 and 2 predict relatively robust results for the steady-state solution of simple shear at constant rate of deformation.

6. Example of steady-state isochoric extension relative to a rotating coordinate system

In order to better understand the implications of an algorithm not being strongly objective, consider the case of isochoric extension relative to a rotating coordinate system. Specifically, let e_i' be an orthonormal triad of vectors which rotates with constant angular velocity $\omega\Gamma$ about the fixed e_3 direction

$$e_1' = \cos(\omega\Gamma t)e_1 + \sin(\omega\Gamma t)e_2, \quad e_2' = -\sin(\omega\Gamma t)e_1 + \cos(\omega\Gamma t)e_2, \quad e_3' = e_3. \tag{6-1}$$

For this deformation field it is convenient to introduce the orthogonal tensor Q defined by

$$Q(t) = e_i' \otimes e_i, \quad \dot{e}_i' = \Omega e_i', \quad \Omega = \dot{Q}Q^T, \quad \Omega = \omega\Gamma(-e_1 \otimes e_2 + e_2 \otimes e_1),$$
$$Q = \cos(\omega\Gamma t)(e_1 \otimes e_1 + e_2 \otimes e_2) + (e_3 \otimes e_3) + \sin(\omega\Gamma t)(-e_1 \otimes e_2 + e_2 \otimes e_1), \tag{6-2}$$

where the usual summation convention is used for repeated indices. Moreover, the deformation gradient F and the velocity gradient L are specified by

$$F = Q\big(\lambda(e_1 \otimes e_1) + \lambda^{-1/2}(e_2 \otimes e_2 + e_3 \otimes e_3)\big), \quad \lambda = \exp(\gamma \Gamma t), \quad J = 1,$$

$$L = \gamma \Gamma Q\big((e_1 \otimes e_1) - \tfrac{1}{2}(e_2 \otimes e_2 + e_3 \otimes e_3)\big) Q^T + \Omega,$$
(6-3)

where λ is the stretch of a material line element which in the reference configuration (at $t = 0$) was oriented in the e_1 direction, and $\gamma \Gamma$ is the constant logarithmic stretch rate. It then follows that the relative deformation tensor associated with (6-3) is given by

$$F_r(t_2) = Q(t_2)\big(e^\kappa(e_1 \otimes e_1) + e^{-\kappa/2}(e_2 \otimes e_2 + e_3 \otimes e_3)\big) Q^T(t_1), \quad \kappa = \Delta t \, \gamma \Gamma.$$
(6-4)

Moreover, since this deformation is isochoric, the tensor F_r is unimodular as was the case in (5-3).

Exact solution. The tensor $B'_e(t)$ is a steady-state solution relative to the rotating basis e'_i of the evolution equation (1-13) provided that it satisfies the equation

$$\Omega B'_e + B'_e \Omega^T = L B'_e + B'_e L^T - \tfrac{2}{3}(D \cdot I) B'_e - \Gamma A_p.$$
(6-5)

The exact solution of this algebraic equation can be written in the form

$$B'_e(t) = a^2(e'_1 \otimes e'_1) + \frac{1}{a}(e'_2 \otimes e'_2 + e'_3 \otimes e'_3),$$
(6-6)

where a takes the value a_e given by

$$a_e = \left(\frac{1+\gamma}{1-2\gamma}\right)^{1/3} \quad \text{for } \gamma < \tfrac{1}{2}.$$
(6-7)

It can be seen from this solution that no steady-state solution of this form exists if the rate of extension is too large ($\gamma \geq \tfrac{1}{2}$).

Solution of Algorithm 1. The values of $\{B'_e(t_1), B'_e(t_2)\}$ correspond to a steady-state solution relative to the rotating basis e'_i of Algorithm 1 if they have the forms

$$B'_e(t_1) = Q(t_1)\hat{B}'_e Q^T(t_1), \quad B'_e(t_2) = Q(t_2)\hat{B}'_e Q^T(t_2), \quad \hat{B}'_e = a^2(e_1 \otimes e_1) + \frac{1}{a}(e_2 \otimes e_2 + e_3 \otimes e_3),$$
(6-8)

where the stretch a needs to be determined. Next, using (1-17) and (6-4) it follows that

$$\bar{B}'_e(t_2) = F'^{-1}_r B'_e(t_2) F'^{-T}_r = Q(t_1)\Big(a^2 e^{-2\kappa}(e_1 \otimes e_1) + \frac{e^\kappa}{a}(e_2 \otimes e_2 + e_3 \otimes e_3)\Big) Q^T(t_1).$$
(6-9)

These expressions will satisfy (4-3) provided that a takes the value a_1, which is the real positive root of the equation

$$\big(\gamma - (\gamma + \kappa)e^{-2\kappa}\big)a_1^3 - \kappa(e^\kappa - e^{-2\kappa})a_1 + \big((\gamma + \kappa)e^\kappa - \gamma\big) = 0$$
(6-10)

closest to unity. It can be shown that in the limit that $\kappa \to 0$ the real positive solution of (6-10) yields the exact result (6-7) for all possible values of γ. For finite values of κ, a maximum value of γ exists beyond which the real solution of (6-10) becomes negative. Thus, the maximum value of γ which produces a physical solution is a function of the size of the time step though the value of κ.
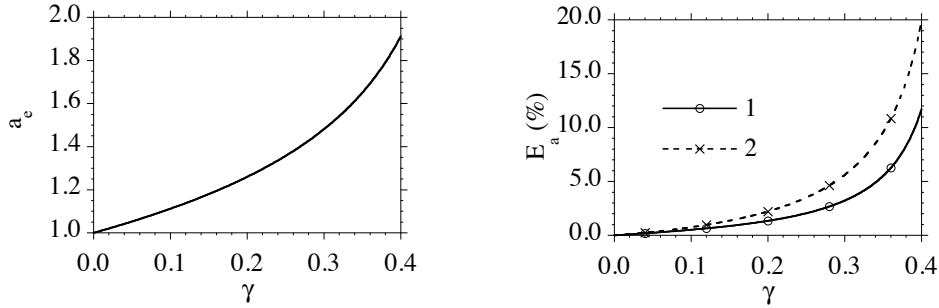
Figure 3. Isochoric extension: exact solution (left) and errors in the solutions predicted by Algorithms 1 and 2 for the normalized time increment $\kappa = 0.1$.

Solution of Algorithm 2. The values of $\{\boldsymbol{B}'_e(t_1), \boldsymbol{B}'_e(t_2)\}$ correspond to a steady-state solution relative to the rotating basis $\boldsymbol{e}'_i$ of Algorithm 2 if they have the forms (6-8). Next, using these expressions, the elastic trial value (4-6), and equating the estimate (4-9) with the deviatoric part $\boldsymbol{B}''_e(t_2)$ of $\boldsymbol{B}'_e(t_2)$ in (6-8) yields an algebraic equation for a which is solved to obtain the value a_2 given by

$$a_2 = \left(\frac{\gamma + \kappa - \gamma e^{-\kappa}}{\gamma + \kappa - \gamma e^{2\kappa}} \right)^{1/3}. \tag{6-11}$$

This solution has the same character as the solution of Algorithm 1.

Figure 3, left, plots the exact solution a_e in (6-7) as a function of the normalized stretch rate γ. The right half of the figure shows the error E_a in (5-14) predicted by Algorithms 1 and 2 for the normalized time increment $\kappa = 0.1$. From this figure is it easy to see that Algorithm 1 is more accurate than Algorithm 2. Moreover, we emphasize that both algorithms are strongly objective because the solutions (6-10) and (6-11) are uninfluenced by rate of rotation $\omega\Gamma$ of the axes $\boldsymbol{e}'_i$, which can be interpreted as a superposed rate of rigid body rotation.

Discussion of Algorithm 3. To discuss Algorithm 3 it is convenient to introduce the two tensors

$$\tilde{\boldsymbol{\Lambda}} = \boldsymbol{Q}(t_2)^T \boldsymbol{\Lambda}_W \boldsymbol{Q}(t_1), \quad \tilde{\boldsymbol{\gamma}} = \boldsymbol{Q}(t_1)^T \left(\gamma \boldsymbol{B}'_e(t_1) + \boldsymbol{B}'_e(t_1)\gamma \right) \boldsymbol{Q}(t_1). \tag{6-12}$$

It then follows from (4-9), (4-15) and (6-8) that

$$\tilde{\boldsymbol{B}}'_e = \boldsymbol{Q}(t_1)\left(\hat{\boldsymbol{B}}'_e + \tilde{\boldsymbol{\gamma}} - \tfrac{2}{3}(\boldsymbol{\gamma} \cdot \boldsymbol{I})\hat{\boldsymbol{B}}'_e \right) \boldsymbol{Q}^T(t_1),$$

$$\hat{\boldsymbol{B}}'_e - \tfrac{1}{3}(\hat{\boldsymbol{B}}'_e \cdot \boldsymbol{I})\boldsymbol{I} = \frac{1}{1+\Delta t\,\Gamma} \tilde{\boldsymbol{\Lambda}}\left(\left(\hat{\boldsymbol{B}}'_e + \tilde{\boldsymbol{\gamma}} - \tfrac{2}{3}(\boldsymbol{\gamma} \cdot \boldsymbol{I})\hat{\boldsymbol{B}}'_e \right) - \tfrac{1}{3}\left[\left(\hat{\boldsymbol{B}}'_e + \tilde{\boldsymbol{\gamma}} - \tfrac{2}{3}(\boldsymbol{\gamma} \cdot \boldsymbol{I})\hat{\boldsymbol{B}}'_e \right) \cdot \boldsymbol{I} \right]\boldsymbol{I} \right)\tilde{\boldsymbol{\Lambda}}^T. \tag{6-13}$$

Consequently, Algorithm 3 will admit solutions of the form (6-8) if the tensors $\{\tilde{\boldsymbol{\Lambda}}, \tilde{\boldsymbol{\gamma}}\}$ are independent of the value of the angular velocity $\omega\Gamma$ and if they are diagonal tensors with respect to the fixed basis $\boldsymbol{e}_i$. Specifically, it can be shown that

$$\tilde{\boldsymbol{\Lambda}} \cdot (\boldsymbol{e}_1 \otimes \boldsymbol{e}_2) = -\tilde{\boldsymbol{\Lambda}} \cdot (\boldsymbol{e}_2 \otimes \boldsymbol{e}_1) = -\frac{4\cos\dfrac{\kappa}{2}\sinh^2\dfrac{\kappa}{4}\sin\dfrac{\omega\kappa}{\gamma}}{\cosh\kappa - 8\sinh^2\dfrac{\kappa}{4}\sin^2\dfrac{\omega\kappa}{2\gamma} + \cos\dfrac{\omega\kappa}{\gamma}}$$

and

$$\tilde{\boldsymbol{\gamma}} \cdot (\boldsymbol{e}_1 \otimes \boldsymbol{e}_2) = \tilde{\boldsymbol{\gamma}} \cdot (\boldsymbol{e}_2 \otimes \boldsymbol{e}_1) = \frac{2(a^3 + 1) \sinh \dfrac{3\kappa}{4} \sin \dfrac{\omega\kappa}{\gamma}}{a \left(\cosh \dfrac{\kappa}{4} + \cosh \dfrac{3\kappa}{4} \cos \dfrac{\omega\kappa}{\gamma} \right)}. \qquad (6\text{-}14)$$

Not only are these quantities nonzero but they depend explicitly on the angular velocity $\omega\Gamma$, which is a direct consequence of the algorithm not being strongly objective.

7. Conclusions

Evolution equations based on the Jaumann derivative like (1-2) or (4-13) necessarily require three types of approximations. One for the specification of the time dependence of the spin tensor $\boldsymbol{W}$, one for the integration algorithm of the evolution equation (1-3) for the rotation tensor $\boldsymbol{\Lambda}_W$, and one for the integration algorithm of the evolution equation for an auxiliary variable like $\overline{\boldsymbol{T}}_W$ in (1-5). In contrast, when evolution equations are formulated in terms of the velocity gradient $\boldsymbol{L}$, the evolution equations (1-9) for the relative deformation gradient $\boldsymbol{F}_r$ and (1-16) for the unimodular part $\boldsymbol{F}'_r$ of the relative deformation tensor can be integrated exactly in terms of the positions of material points at the beginning and end of the time step, independently of the time dependence of $\boldsymbol{L}$ during the time step. Consequently, such integration algorithms have the advantage that approximations are limited to the algorithms for the evolution equations of the auxiliary tensors like (1-11) for $\overline{\boldsymbol{T}}_L$ and (1-18) for $\overline{\boldsymbol{B}}'_e$. Moreover, since these auxiliary tensors are unaffected by SRBM it is simple to develop algorithms which are strongly objective in the sense discussed in [Papes and Mazza 2009].

Focusing attention to the response of a simple elastic-viscoplastic material, examples of steady-state simple shear at constant shear rate and steady-state isochoric extension relative to a rotating coordinate system are considered to assess the robustness and accuracy of the three algorithms presented in Section 4. Algorithms 1 and 2 are strongly objective, but Algorithm 3, which is based on the Hughes–Winget algorithm, is only weakly incrementally objective and exhibits unphysical dependence on the arbitrary rotation tensor $\boldsymbol{Q}$ in SRBM when the rate of deformation is nonzero. Algorithms 1 and 2 give robust and relatively accurate predictions of these steady-state solutions for reasonably large time increments. From a practical point of view, Algorithm 2 is perhaps the easiest to implement. It presents an improvement over the simple formulation in [Rubin and Attia 1996] in that here the elastic trial value $\boldsymbol{B}'^*_e(t_2)$ is evaluated exactly.

Acknowledgements

References

[Argyris 1982] J. Argyris, "An excursion into large rotations", *Comput. Methods Appl. Mech. Eng.* **32** (1982), 85–155.

[Argyris and Poterasu 1993] J. Argyris and V. F. Poterasu, "Large rotations revisited application of Lie algebra", *Comput. Methods Appl. Mech. Eng.* **103** (1993), 11–42.

[Becker 2006] R. Becker, "A plasticity integration algorithm motivated by analytical integration of a generalized quadratic function", technical report UCRL-TR-21953, Livermore, CA, 2006, available at https://e-reports-ext.llnl.gov/pdf/331305.pdf.

[Besseling 1968] J. F. Besseling, "A thermodynamic approach to rheology", pp. 16–53 in *Proc. IUTAM Symp. on Irreversible Aspects of Continuum Mechanics*, edited by H. Parkus and L. I. Sedov, Springer, Vienna, 1968.

[Dienes 1979] J. K. Dienes, "On the analysis of rotation and stress rate in deforming bodies", *Acta Mech.* **32** (1979), 217–232.

[Eckart 1948] C. Eckart, "The thermodynamics of irreversible processes, IV: the theory of elasticity and anelasticity", *Phys. Rev.* **73** (1948), 373–382.

[Flory 1961] P. Flory, "Thermodynamic relations for high elastic materials", *T. Faraday Soc.* **57** (1961), 829–838.

[Govindjee 1997] S. Govindjee, "Accuracy and stability for integration of Jaumann stress rate equations in spinning bodies", *Engrg. Comput.* **14** (1997), 14–30.

[Holzapfel 2000] G. A. Holzapfel, *Nonlinear solid mechanics: a continuum approach for engineering*, Wiley, New York, 2000.

[Hughes and Winget 1980] T. J. R. Hughes and J. Winget, "Finite rotation effects in numerical integration of rate constitutive equations arising in large-deformation analysis", *Int. J. Numer. Meth. Engrg.* **15**:12 (1980), 1862–1867.

[Leonov 1976] A. I. Leonov, "Nonequilibrium thermodynamics and rheology of viscoelastic polymer media", *Rheol. Acta* **15** (1976), 85–98.

[Malvern 1969] L. E. Malvern, *Introduction to the mechanics of a continuous medium*, Prentice-Hall, Englewood Cliffs, NJ, 1969.

[Papes and Mazza 2009] O. Papes and E. Mazza, "Numerical implementation of an elasto-viscoplastic material law for FE calculations of biological tissues at finite deformations", technical report, 2009, available at http://www.zfm.ethz.ch/e/v/ncm/handouts/Implementation_Rubin_Bodner.pdf.

[Rashid 1993] M. M. Rashid, "Incremental kinematics for finite element applications", *Int. J. Numer. Meth. Engrg.* **36** (1993), 3937–3956.

[Rubin 1994] M. B. Rubin, "Plasticity theory formulated in terms of physically based microstructural variables, I: Theory", *Int. J. Solids Struct.* **31** (1994), 2615–2634.

[Rubin 1996] M. B. Rubin, "On the treatment of elastic deformation in finite elastic-viscoplastic theory", *Int. J. Plast.* **12** (1996), 951–965.

[Rubin 2001] M. B. Rubin, "Physical reasons for abandoning plastic deformation measures in finite plasticity and viscoplasticity theory", *Arch. Mech.* **53**:4-5 (2001), 519–539.

[Rubin 2007] M. B. Rubin, "A simplified implicit Newmark integration scheme for finite rotations", *Comput. Math. Appl.* **53** (2007), 219–231.

[Rubin and Attia 1996] M. B. Rubin and A. Attia, "Calculation of hyperelastic response of finitely deformed elastic-viscoplastic materials", *Int. J. Numer. Meth. Engrg.* **39** (1996), 309–320.

[Simo 1992] J. C. Simo, "Algorithms for static and dynamic multiplicative plasticity that preserve the classical return mapping schemes of the infinitesimal theory", *Comp. Meth. Appl. Mech. Engrg.* **99** (1992), 61–112.

[Simo and Hughes 1998] J. C. Simo and T. J. R. Hughes, *Computational inelasticity*, Springer, New York, 1998.

[Simo and Vu-Quoc 1988] J. C. Simo and L. Vu-Quoc, "On the dynamics in space of rods undergoing large motions: a geometrically exact approach", *Comput. Methods Appl. Mech. Eng.* **66** (1988), 125–161.

M. B. RUBIN: mbrubin@tx.technion.ac.il
Faculty of Mechanical Engineering, Technion – Israel Institute of Technology, 32000 Haifa, Israel

O. PAPES: ondrej@papes.ch
Institute of Mechanical Systems, Department of Mechanical Engineering, ETH Zentrum, CH-8092 Zurich, Switzerland

msp

msp

NUMERICAL AND EXPERIMENTAL EVALUATION OF CRYOGENIC TENSILE STRENGTH OF WOVEN FABRIC-REINFORCED GLASS/EPOXY COMPOSITES USING OPEN HOLE SPECIMENS

YASUHIDE SHINDO, SHINYA WATANABE, TOMO TAKEDA,
FUMIO NARITA, TAKUYA MATSUDA AND SATORU YAMAKI

A numerical and experimental evaluation of the tensile strength properties of woven glass fiber-reinforced polymer composite laminates at cryogenic temperatures was conducted by means of open hole specimens. Tensile tests were performed on specimens with a circular hole at room temperature, at 77 K, and at 4 K. The length of the hole edge damage zone corresponding to specimen failure was determined by microscopic examination of the fracture surfaces. A method based on finite element analysis was developed for estimating the cryogenic tensile strength of the unnotched woven laminates using the experimentally determined failure load and damage zone length. The results suggest that the tensile strength of woven composite laminates at cryogenic temperatures can be determined effectively by this approach.

1. Introduction

Fiber-reinforced polymer composites are widely used in many modern structural applications because of their unique properties. In particular, the demand for lightweight structures in the next generation of reusable launch vehicles may likely result in the use of carbon fiber-reinforced polymer (CFRP) composites in the design of the cryogenic liquid hydrogen and oxygen fuel tanks [Melcher and Johnson 2007]. Also, superconducting magnets for use in the International Thermonuclear Experimental Reactor (ITER) may use large quantities of woven glass fiber-reinforced polymer (GFRP) composite laminates as thermal and electrical insulation and structural support [Tsuji et al. 2001]. Hence, the mechanical properties of woven composite laminates must be characterized and fully understood under application-specific environmental conditions for the design process of cryogenic systems.

Accurate in-plane strength properties in tension of composite laminates are essential for material selection decision and as design data. Tensile tests have been a fundamental method of characterizing the mechanical response of composite materials for many years. The standard methods for tensile tests on composite materials — for example, American Society for Testing and Materials D 3039/D 3039M [ASTM 2000] and Japanese Industrial Standards K 7164 and K 7165 [JIS 2005; 2008] — are designed for room-temperature testing. However, there is currently no universally accepted method for tensile tests on composites at cryogenic temperatures. The determination of composite tensile properties, especially the ultimate strength, at cryogenic temperatures is not straightforward [Kumagai et al. 2003]. This is a result of the high strength, stiffness, and anisotropy of polymer matrix composite materials. Typically, cryogenic tensile tests on polymer matrix composites have resulted in noticeable inconsistencies with regard to strength values [Eisenreich and Cox 1992]. Also, cryogenic testing problems include that

the specimen slips out of the test grips during testing. Increasing the clamping force to prevent grip slippage leads to increased stress concentration within the grip that influences the measured strength data. To address the problems encountered with cryogenic tensile testing, some researchers have devoted their efforts to evaluating the cryogenic mechanical response of composite specimens under tension. For example, Kumagai et al. [2004] experimentally and numerically examined the effects of specimen geometry and gripping method on the tensile properties of woven GFRP composite laminates having high glass content at cryogenic temperatures. Shindo et al. [2006] investigated the cryogenic mechanical response of woven GFRP composite specimens under tension, and discussed the influence of specimen geometry and size on the tensile properties of the composites.

An attractive alternative to the tensile test for the determination of the composite strength at cryogenic temperatures is the open hole tensile test. The open hole tensile test was originally developed to evaluate the notched strength (failure load divided by the gross cross-sectional area of composite laminates) [ASTM 2007]. However, the notched strength is not a material property but a structural property. For an open hole specimen, the failure load is correlated with the composite strength (a material property) and the size of the damage zone near the hole edge [Kogo et al. 1998; Green et al. 2007]. In addition, since the hole causes a stress concentration, the failure of the open hole specimen occurs at loads far below the failure load of the unnotched specimen. Therefore, the open hole tensile test offers a potentially attractive approach for avoiding cryogenic testing problems such as slippage between the specimen and the grips, and evaluating the composite strength.

A combined numerical-experimental method is presented for the evaluation of the tensile strength properties of woven GFRP composite laminates at cryogenic temperatures using open hole specimens. The open hole tensile tests were carried out at room temperature, liquid nitrogen temperature (77 K), and liquid helium temperature (4 K), and the failure loads were obtained. From microscopic examination of the fracture surfaces, the length of the damage zone at the hole edge was measured. The experimentally determined failure load and damage zone length were applied to the finite element model to estimate the cryogenic tensile strength of the unnotched woven laminates. The validity of this approach is discussed by comparing the predictions with existing experimental data.

2. Experimental procedure

2.1. *Material and specimen.* In this work, National Electrical Manufacturer's Association grade G-11 woven GFRP laminates (Arisawa Mfg. Co., Ltd., Japan) were used. The basis for G-11 woven laminates is a plain weave fabric of E-glass. The plain weave is produced by interlacing warp fiber bundles and fill fiber bundles. The weave contains 43 and 32 fiber bundles per 25.4 mm in the warp and fill directions, respectively. The polymer matrix is a bisphenol A epoxy resin. The overall fiber volume fraction is 47%.

Open hole specimens were employed for experiments. The geometry and dimensions of the specimen are shown in Figure 1. The specimens were cut with the length parallel to the warp direction. The dimensions of the specimen were 120 mm in overall length (l), 27 mm in length of the grip section (g), and 25 mm in width (w). The nominal thickness of the specimen (h) was equal to 2 mm. The specimens also contain a circular hole with a diameter (d) of 4.2 mm in the center. The width to hole diameter ratio w/d was chosen based on the ASTM D 5766 standard [2007]. The tapered tabs were bonded to the sample ends by means of an adhesive system that is suitable for use at cryogenic temperatures.
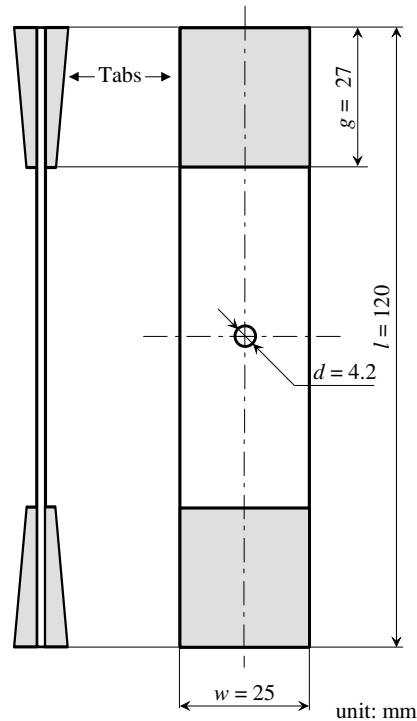
Figure 1. Geometry and dimensions of open hole specimen.

2.2. *Testing method.* Tensile tests were conducted on the open hole specimens at room temperature, 77 K, and 4 K using a 30 kN capacity servohydraulic testing machine. The strains were measured using strain gages bonded to the front and back surfaces of the specimen. The strain gages were placed near the hole as shown in Figure 2. Each specimen was tested at a displacement rate of 0.5 mm/min. Low-temperature environments were achieved by immersing the loading fixture and specimen in liquid nitrogen or liquid helium. To keep the specimen fully submerged in liquid helium during the test, an
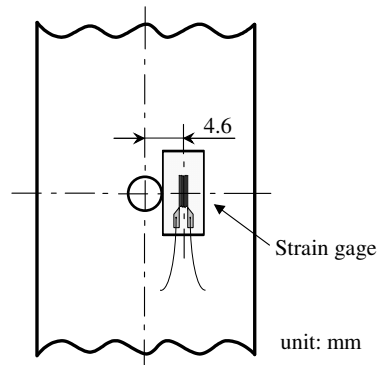
Figure 2. Position of the strain gage on open hole specimen.

automatic refill system (TRG-350D, Taiyo Nippon Sanso Co., Ltd., Japan) was incorporated. Load P and strain ε data were taken up to the catastrophic failure of the open hole specimens, and the failure loads P_c were measured.

Microscopic observations of the specimen surfaces and fracture surfaces were made with optical microscopy in order to obtain a better understanding of damage mechanisms of G-11 woven laminates at cryogenic temperatures. In addition, the fracture surfaces were examined in detail and the length of the critical damage zone (that is, the damage zone just before catastrophic failure), D_c, was determined from the fracture surface morphologies. All observations were performed after the completion of tensile testing.

3. Finite element analysis

Finite element simulations were carried out for the G-11 open hole specimens. The typical mesh of an open hole specimen is shown in Figure 3. The origin of the rectangular Cartesian coordinate system O-xyz is fixed at the center of the open hole specimen, and the x, y, and z-axes are in the fill, warp (tensile loading), and thickness directions, respectively. In the finite element calculations, plane stress conditions in the x-y plane were assumed. The specimen was meshed using 6-noded triangular elements, and a relatively fine mesh was used adjacent to the hole in order to evaluate accurately the stress state.
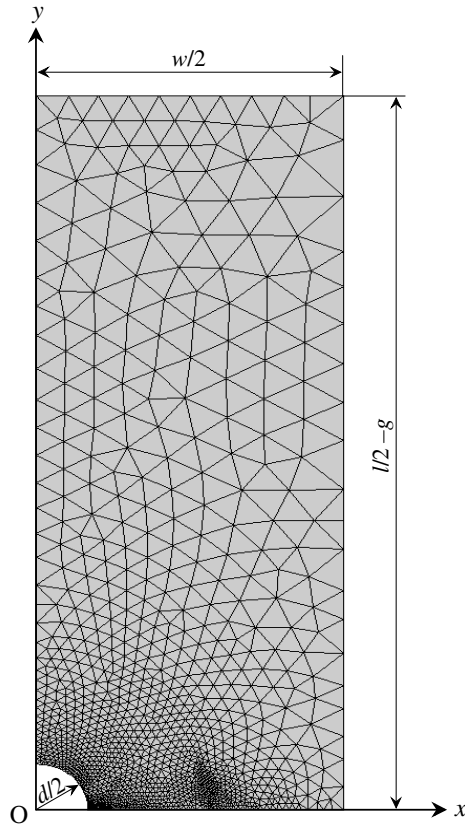
Figure 3. Finite element mesh of open hole specimen.

The finite element model treated the entire plain weave fabric composite as one homogeneous material with orthotropic material properties. Due to symmetry in geometry and loading, only one-quarter of the specimen was considered. The gripped portion was not included in the model.

The strain-stress relationships for the orthotropic composite specimen under plane stress are

$$
\begin{bmatrix} \varepsilon_{xx}(x,y) \\ \varepsilon_{yy}(x,y) \\ 2\varepsilon_{xy}(x,y) \end{bmatrix} = \begin{bmatrix} 1/E_x & -\nu_{xy}/E_x & 0 \\ -\nu_{xy}/E_x & 1/E_y & 0 \\ 0 & 0 & 1/G_{xy} \end{bmatrix} \begin{bmatrix} \sigma_{xx}(x,y) \\ \sigma_{yy}(x,y) \\ \sigma_{xy}(x,y) \end{bmatrix}, \tag{1}
$$

where $\varepsilon_{xx}(x,y)$, $\varepsilon_{yy}(x,y)$, and $\varepsilon_{xy}(x,y) = \varepsilon_{yx}(x,y)$ are the strain components; $\sigma_{xx}(x,y)$, $\sigma_{yy}(x,y)$, and $\sigma_{xy}(x,y) = \sigma_{yx}(x,y)$ are the stress components; E_x and E_y are the Young's moduli; G_{xy} is the shear modulus; and $\nu_{xy} = \nu_{yx}E_x/E_y$ is the Poisson's ratio. The Poisson's ratio ν_{xy} reflects shrinkage (expansion) in the y-direction, due to tensile (compressive) stress in the x-direction. The normal strain component ε_{zz} perpendicular to the x-y plane can be expressed in terms of the stress components $\sigma_{xx}(x,y)$ and $\sigma_{yy}(x,y)$ and the elastic properties E_x, E_y, $\nu_{xz} = \nu_{zx}E_x/E_z$, and $\nu_{yz} = \nu_{zy}E_y/E_z$ as follows:

$$
\varepsilon_{zz}(x,y) = -\frac{\nu_{xz}}{E_x}\sigma_{xx}(x,y) - \frac{\nu_{yz}}{E_y}\sigma_{yy}(x,y). \tag{2}
$$

The elastic properties E_x, E_y, E_z, G_{xy}, ν_{xy}, ν_{yz}, and ν_{zx} were determined from the micromechanics model for plain weave fabric composites [Hahn and Pandey 1994]. The model includes parameters to characterize the fabric architecture and constituent material properties. The geometrical parameters of G-11 woven laminates were measured on scanning electron microscopy photographs of the woven laminate cross-sections. The Young's modulus, shear modulus, and Poisson's ratio of the matrix and fiber were obtained from the exponential functions of temperature based on the experimental data [Dahlerup-Petersen and Perrot 1979]. The composite elastic moduli can be determined under the assumption of uniform strain inside the unit cell. Table 1 lists the predicted Young's moduli E_x, E_y, and E_z, shear modulus G_{xy}, and Poisson's ratios ν_{xy}, ν_{yz}, and ν_{zx} of G-11 woven laminates at 295 K (room temperature), 77 K, and 4 K. In the table, the values in parentheses are the experimental results [Sanada and Shindo 2005]. By comparing the predicted elastic moduli with experimental data, this model is shown to be efficient and accurate.

In [Whitney and Nuismer 1974] two stress failure criteria were proposed, the point stress criterion (PSC) and the average stress criterion (ASC), for predicting the point at which the catastrophic failure of

	295 K		77 K		4 K	
E_x (GPa)	21.07	(24.1)	28.47	(31.1)	31.95	(34.3)
E_y (GPa)	26.66	(27.9)	32.98	(32.7)	35.94	(36.9)
E_z (GPa)	10.75		20.48		25.03	
G_{xy} (GPa)	5.17	(6.2)	8.76	(9.4)	10.28	(10.1)
ν_{xy}	0.10		0.16		0.19	
ν_{yz}	0.32		0.33		0.34	
ν_{zx}	0.17		0.26		0.29	

Table 1. Predicted elastic moduli of G-11 woven laminates.

a laminated composite with a circular hole occurs. These two stress criteria use the linear elastic stress field adjacent to the hole, and are semiempirical models based on the laminate tensile strength and a characteristic dimension. The characteristic dimension seems to be related to the length of the damage zone at the hole edge [Kogo et al. 1998]. With regard to the above criteria, Belmonte et al. [2001] have pointed out that the ASC is more consistent with the experimental results on damage growth and failure at circular holes in woven GFRP composite laminates under tension at room temperature than the PSC. While the ASC has the merits of being simple to apply, its limitations are that it does not consider the evolution of damage from the hole. In practice the damage will modify, significantly, the assumed elastic stress distribution near the hole, and the ASC does not give an adequate description of the stress state. Actually, the stress distribution in the damage zone is very complicated. Although the stress level along the direction perpendicular to the applied tensile load inside the damage zone can be considered to keep nearly constant and the stress is correlated with the unnotched tensile strength [Zhen 1983], no one has evaluated exactly the stress state near the hole in the damaged composites. In the present study, the damage zone was therefore modeled as a strip with the critical damage zone length D_c ahead of the hole edge, and the normal stress σ_{yy} was assumed to be uniform within the damage zone. Let u_x and u_y be the displacement components. The boundary conditions can be written as

$$\sigma_{yy}(x, 0) = \sigma_0 \quad \text{for} \quad d/2 \leq x < d/2 + D_c, \tag{3}$$

$$u_y(x, 0) = 0 \quad \text{for} \quad d/2 + D_c \leq x \leq w/2, \tag{4}$$

$$\sigma_{yx}(x, 0) = 0 \quad \text{for} \quad d/2 \leq x \leq w/2, \tag{5}$$

$$u_y(x, l/2 - g) = u_y^* \quad \text{for} \quad 0 \leq x \leq w/2, \tag{6}$$

$$\sigma_{yx}(x, l/2 - g) = 0 \quad \text{for} \quad 0 \leq x < w/2, \tag{7}$$

$$u_x(0, y) = 0 \quad \text{for} \quad d/2 \leq y \leq l/2 - g, \tag{8}$$

$$\sigma_{xy}(0, y) = 0 \quad \text{for} \quad d/2 \leq y \leq l/2 - g, \tag{9}$$

$$\sigma_{xx}(w/2, y) = 0 \quad \text{for} \quad 0 \leq y < l/2 - g, \tag{10}$$

$$u_x(w/2, l/2 - g) = 0 \tag{11}$$

$$\sigma_{xy}(w/2, y) = 0 \quad \text{for} \quad 0 \leq y < l/2 - g, \tag{12}$$

where σ_0 is the uniform stress in the damage zone and u_y^* is the prescribed displacement in the y-direction. The prescribed displacement u_y^* was determined from the condition

$$\frac{2}{w} \int_0^{w/2} \sigma_{yy}(x, l/2 - g)\, dx = \frac{P_c}{wh}. \tag{13}$$

The displacement boundary condition given in (11) represents the constraint effect resulting from the rigid clamping of the specimen's ends. The stress distribution corresponding to the experimentally obtained failure load P_c and critical damage zone length D_c was determined from the finite element analysis, as illustrated in Figure 4. The solid and dashed lines in the figure represent the stress distributions near the hole with and without the damage zone, respectively. The solid line can be obtained using the condition of finite stress σ_0 (vanishing stress singularity) at the tip $x = d/2 + D_c$, and the σ_0 value was evaluated
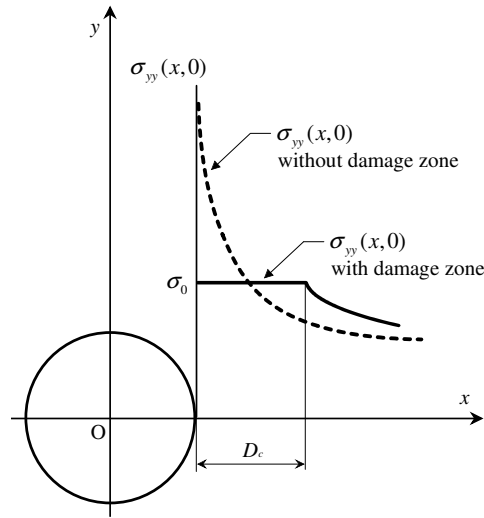
Figure 4. Stress distribution near the hole with damage zone.

for each specimen. The numerically evaluated uniform stress σ_0 was regarded as the ultimate tensile strength of the specimen material.

4. Results and discussion

Figure 5 shows the measured gage strains ε in percent as a function of the applied stress σ ($\sigma = P/wh$) at room temperature (RT), 77 K, and 4 K. For the sake of comparison, a linear elastic finite element analysis was carried out and the calculated strains are also provided in the figure (the dashed lines). The calculated strain is the normal strain in the y-direction at the center of the strain gage ($x = 4.6$, $y = 0$)
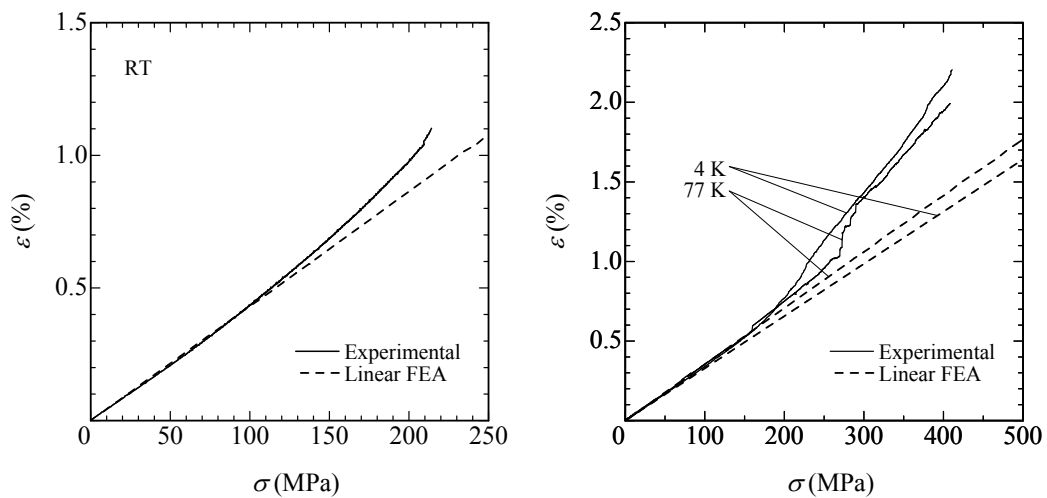
Figure 5. Measured and calculated gage strains versus applied stress at room temperature (left) and at 77 K and 4 K (right).
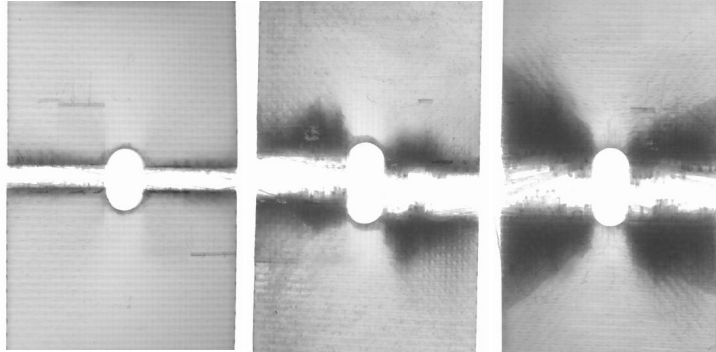
Figure 6. Photographs of failed specimens at room temperature (left), 77 K (middle), and 4 K (right).

and corresponds to the measured gage strain. The linear analysis did not consider the hole edge damage ($D_c = 0$). Each curve exhibits a linear initial region, and the calculated curves agree with the experimental data very well. With increase in the applied stress, the nonlinearity of the curves appears at both room temperature and cryogenic temperatures. Note that pronounced nonlinearity of the curves can be found at 77 K and 4 K, which can be attributed to the accumulation of damage such as microcracks in the matrix material [Takeda et al. 2005].

Photographs of the open hole specimens after the tensile tests at room temperature, 77 K, and 4 K are shown in Figure 6. Failure invariably occurs in the central section of the specimens. A straight failure plane is formed normal to the applied load at room temperature. In the specimens at 77 K and 4 K, an intense damage zone is found at the hole edge. Also, the size of the damage in the specimen at 4 K is more extensive than that at 77 K. This damage zone is generally seen on both sides of the hole. Figure 7 shows the optical microscopy images of the edge sections of failed specimens at room temperature, 77 K, and 4 K. It can be seen that the specimen at 4 K shows extensive delamination, while the specimen at 77 K displays less delamination, and the specimen at room temperature virtually none. This may be
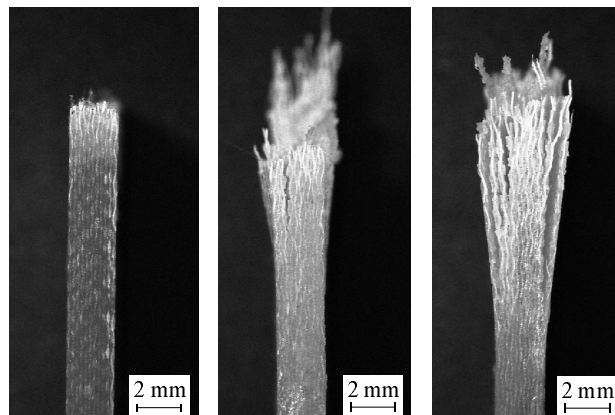
Figure 7. Edge views of failed specimens at room temperature (left), 77 K (middle), and 4 K (right).
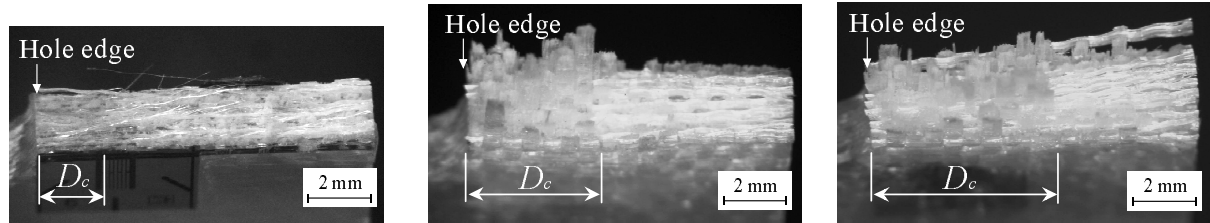
Figure 8. Fracture surface morphologies of failed specimens at room temperature (left), 77 K (middle), and 4 K (right).

due to the residual thermal stresses in the specimens generated during cooling and the brittleness of the polymer resins at cryogenic temperatures. Figure 8 shows typical fracture surface morphologies of failed specimens at room temperature, 77 K, and 4 K. The fracture surface displays the drastic difference in failure mechanism between room temperature and cryogenic temperatures. The region of the room temperature fracture surface near the hole edge is relatively smooth, and loose transverse (fill) fiber bundles are found away from the hole edge. Loose transverse fiber bundles are a characteristic of woven laminate tensile failure at room temperature [Karayaka and Kurath 1994]. Based on this finding, damage before the catastrophic failure of the open hole specimens at room temperature seems to lead to the smooth fracture surface near the hole edge and the length of the smooth region is taken as the critical damage zone length D_c at room temperature. At 77 K and 4 K, the fracture surfaces near the hole edge exhibit fiber pullout, and the fibers tend to pull out as bundles rather than individual filaments. Damage (microcracks) caused by large residual thermal stresses at cryogenic temperatures and concentrated stresses near the hole lead to the degradation in the fiber-matrix interface bonding and the fiber breakage. In the case of poor fiber-matrix bonding, the broken fibers are pulled out of the matrix [Berthelot 1999]. The fracture surfaces at 77 K and 4 K suggest that the length of the fiber pullout region corresponds to the critical damage zone length D_c. Overall the fracture surface studies indicate that the smooth region at room temperature and the fiber pullout region at 77 K and 4 K were localized near the hole edge, and the boundary of the damage zone was discerned relatively well. Table 2 shows the experimentally obtained failure loads P_c and critical damage zone lengths D_c at room temperature, 77 K, and 4 K. The critical damage zone length D_c in the table was calculated as the average of the values measured on both sides of the hole after the tensile test of each specimen. The failure loads at 77 K and 4 K are higher than that at room temperature, and the failure load at 77 K is similar to that at 4 K. The critical damage zone length increases as the temperature decreases. The difference in the critical damage zone length reflects the microdamage mechanism changes with changing temperature. Also, the critical damage zone length

RT		77 K		4 K	
P_c (kN)	D_c (mm)	P_c (kN)	D_c (mm)	P_c (kN)	D_c (mm)
10.8	1.9	20.4	5.7	20.8	6.8
10.7	2.0	21.7	4.9	22.3	6.8
10.5	2.6	21.2	5.2	20.7	6.7

Table 2. Failure loads P_c and critical damage zone lengths D_c.
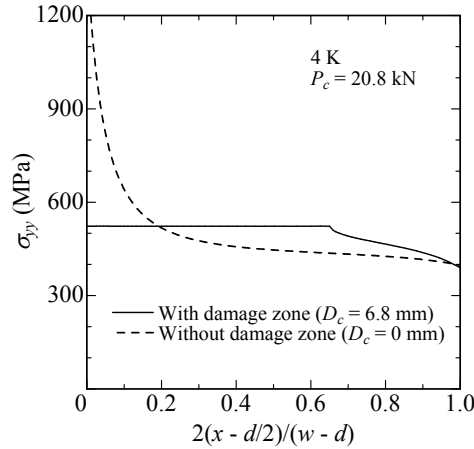
Figure 9. Stress distributions near the hole along the line $y = 0$ ($d/2 \leq x \leq w/2$).

at each temperature is reasonably constant across all specimens. This implies that the measurement of the critical damage zone length is reliable.

Figure 9 shows the distributions of the normal stresses σ_{yy} along the x-axis ($y = 0$, $d/2 \leq x \leq w/2$) with and without the damage zone at 4 K as a function of distance from the hole edge $2(x - d/2)/(w - d)$. The failure load P_c and the critical damage zone length D_c were taken as 20.8 kN and 6.8 mm, respectively. Based on the stress distribution with the damage zone, the uniform stress σ_0 in the damage zone can be determined to be 523 MPa. Table 3 presents the numerically determined uniform stresses σ_0 at room temperature, 77 K, and 4 K. The uniform stresses σ_0 in the table are averages of three values at each temperature, and the values in parentheses are the standard deviations. For comparison, the ultimate tensile strengths σ_{ult} from the tensile tests on the unnotched G-11 specimens [Shindo et al. 2006] are given in the table. The numerically determined uniform stresses σ_0 are in good agreement with the σ_{ult} values. This suggests that the uniform stress from the present combined numerical-experimental method is a reliable estimate of the composite tensile strength at cryogenic temperatures. It may be possible to carry out the cryogenic tensile tests on the unnotched specimens of the commercial G-11 woven laminates. However, cryogenic tensile testing is very difficult for high-strength advanced composite materials such as CFRP composites. Although the issue of grip slippage in unnotched tests, of course, has to be overcome, the present method is an efficient approach to measuring the cryogenic tensile strength of high-strength composite materials because high applied tensile load and high clamping force are not necessary to break the open hole specimen.

RT		77 K		4 K	
σ_0 (MPa)	σ_{ult} (MPa)	σ_0 (MPa)	σ_{ult} (MPa)	σ_0 (MPa)	σ_{ult} (MPa)
331	310	548	610	535	579
(13.9)	(28.7)	(18.8)	(73.1)	(18.4)	(14.0)

Table 3. Predicted and experimental ultimate tensile strengths of G-11 woven composite laminates. Values in parentheses are standard deviations in a sample of size three.

5. Conclusions

In this article, we studied the tensile strength properties of woven glass fiber-reinforced polymer composite laminates at cryogenic temperatures through a combined numerical-experimental approach using open hole specimens. The major conclusion from the current investigation is that it is possible to obtain reasonable values of the ultimate tensile strength of woven laminates at cryogenic temperatures from open hole specimens. Also, at cryogenic temperatures, microscopic examination of the failed specimens suggests that an intense damage zone develops at the hole edge prior to catastrophic failure. The open hole acts as sufficient stress riser to force the specimen failure from the hole, and, therefore, the present method provides a novel approach to measuring the cryogenic tensile strength of high-strength advanced composite materials.

Dedication

We dedicate this work to the memory of Marie-Louise. As Associate Editor of the *International Journal of Solids and Structures* and the *Journal of Mechanics of Materials and Structures*, she devoted herself to the success of those journals, and we owe her much for their progress. When I first met her and Charles for the first time in 1994, in Seattle, she talked to me as a friend. Since then, I enjoyed various discussions with them at conferences. I miss her. We thank Marie-Louise and Charles Steele for their efforts and dedication to two successful journals.

References

[ASTM 2000] "Standard test method for tensile properties of polymer matrix composite materials", Standard ASTM D3039/ D3039M-00$^{\epsilon 1}$, American Society for Testing and Materials, Philadelphia, 2000. Superseded by the active standard.

[ASTM 2007] "Standard test method for open hole tensile strength of polymer matrix composite laminates", Standard ASTM D5766/D5766M-02a, American Society for Testing and Materials, Philadelphia, 2007. reapproved 2007. Superseded by the active standard.

[Belmonte et al. 2001] H. M. S. Belmonte, C. I. C. Manger, S. L. Ogin, P. A. Smith, and R. Lewin, "Characterisation and modelling of the notched tensile fracture of woven quasi-isotropic GFRP laminates", *Compos. Sci. Technol.* **61**:4 (2001), 585–597.

[Berthelot 1999] J.-M. Berthelot, *Composite materials: mechanical behavior and structural analysis*, Springer, New York, 1999.

[Dahlerup-Petersen and Perrot 1979] K. Dahlerup-Petersen and A. Perrot, "Properties of organic composite materials at cryogenic temperatures", Technical report CERN-ISR-BOM-79-39, European Organization for Nuclear Research (CERN), Geneva, November 1979, available at http://cdsweb.cern.ch/record/312298.

[Eisenreich and Cox 1992] T. J. Eisenreich and D. S. Cox, "Modification of the ASTM D 3039 tensile specimen for cryogenic applications", *Adv. Cryog. Eng.* **38** (1992), 437–444.

[Green et al. 2007] B. G. Green, M. R. Wisnom, and S. R. Hallett, "An experimental investigation into the tensile strength scaling of notched composites", *Compos. A Appl. Sci. Manuf.* **38**:3 (2007), 867–878.

[Hahn and Pandey 1994] H. T. Hahn and R. Pandey, "A micromechanics model for thermoelastic properties of plain weave fabric composites", *J. Eng. Mater. Technol.* (*ASME*) **116**:4 (1994), 517–523.

[JIS 2005] "Plastics – Determination of tensile properties – Part 4: Test conditions for isotropic and orthotropic fibre-reinforced plastic composites", Standard JIS K 7164:2005, Japanese Standards Association, Tokyo, 2005.

[JIS 2008] "Plastics – Determination of tensile properties – Part 5: Test conditions for unidirectional fibre-reinforced plastic composites", Standard JIS K 7165:2008, Japanese Standards Association, Tokyo, 2008.

[Karayaka and Kurath 1994] M. Karayaka and P. Kurath, "Deformation and failure behavior of woven composite laminates", *J. Eng. Mater. Technol.* (*ASME*) **116**:2 (1994), 222–232.

[Kogo et al. 1998] Y. Kogo, H. Hatta, H. Kawada, and T. Machida, "Effect of stress concentration on tensile fracture behavior of carbon-carbon composites", *J. Compos. Mater.* **32**:13 (1998), 1273–1294.

[Kumagai et al. 2003] S. Kumagai, Y. Shindo, K. Horiguchi, and T. Takeda, "Mechanical characterization of CFRP woven laminates between room temperature and 4K", *JSME Int. J. Series A* **46**:3 (2003), 359–364.

[Kumagai et al. 2004] S. Kumagai, Y. Shindo, K. Horiguchi, and F. Narita, "Experimental and finite-element analysis of woven glass-cloth/epoxy laminate tensile specimen at room and low temperatures", *Mech. Adv. Mater. Struct.* **11**:1 (2004), 51–66.

[Melcher and Johnson 2007] R. J. Melcher and W. S. Johnson, "Mode I fracture toughness of an adhesively bonded composite-composite joint in a cryogenic environment", *Compos. Sci. Technol.* **67**:3–4 (2007), 501–506.

[Sanada and Shindo 2005] K. Sanada and Y. Shindo, "Cryogenic damage and fracture behaviors of G-11 woven glass-epoxy laminates", *JSME Int. J. Series A* **48**:2 (2005), 91–99.

[Shindo et al. 2006] Y. Shindo, F. Narita, K. Horiguchi, S. Takano, T. Takeda, and K. Sanada, "Tensile behavior and damage/acoustic emission characteristics of woven glass fiber reinforced/epoxy composite laminates at cryogenic temperatures", *Adv. Cryog. Eng.* **52** (2006), 249–256.

[Takeda et al. 2005] T. Takeda, S. Takano, Y. Shindo, and F. Narita, "Deformation and progressive failure behavior of woven-fabric-reinforced glass/epoxy composite laminates under tensile loading at cryogenic temperatures", *Compos. Sci. Technol.* **65**:11-12 (2005), 1691–1702.

[Tsuji et al. 2001] H. Tsuji, S. Egorov, J. Minervini, M. Martovetsky, K. Okuno, Y. Takahashi, and R. J. Thome, "ITER R&D: magnets: central solenoid model coil", *Fusion Eng. Des.* **55**:2–3 (2001), 153–170.

[Whitney and Nuismer 1974] J. M. Whitney and R. J. Nuismer, "Stress fracture criteria for laminated composites containing stress concentrations", *J. Compos. Mater.* **8**:3 (1974), 253–265.

[Zhen 1983] S. Zhen, "The D criterion theory in notched composite materials", *J. Reinf. Plast. Compos.* **2**:2 (1983), 98–110.

YASUHIDE SHINDO: shindo@material.tohoku.ac.jp
Tohoku University, Department of Materials Processing, Graduate School of Engineering, Aoba-ku, Sendai 980-8579, Japan

SHINYA WATANABE: *Tohoku University, Department of Materials Processing, Graduate School of Engineering, Aoba-ku, Sendai 980-8579, Japan*

TOMO TAKEDA: takeda-t@material.tohoku.ac.jp
Tohoku University, Department of Materials Processing, Graduate School of Engineering, Aoba-ku, Sendai 980-8579, Japan

FUMIO NARITA: narita@material.tohoku.ac.jp
Tohoku University, Department of Materials Processing, Graduate School of Engineering, Aoba-ku, Sendai 980-8579, Japan

TAKUYA MATSUDA: *Tohoku University, Department of Materials Processing, Graduate School of Engineering, Aoba-ku, Sendai 980-8579, Japan*

SATORU YAMAKI: *Tohoku University, Department of Materials Processing, Graduate School of Engineering, Aoba-ku, Sendai 980-8579, Japan*

▪▪
▪msp

ON TIME-DELAYED AND FEED-FORWARD TRANSMISSION LINE MODELS OF THE COCHLEA

ROBERT SZALAI, BASTIAN EPP, ALAN R. CHAMPNEYS AND MARTIN HOMER

The mammalian cochlea is a remarkable organ that is able to provide up to 60dB amplification of low amplitude sound with sharp tuning. It has been proposed that in order qualitatively to explain experimental data, models of the basilar membrane impedance must include an exponential term that represents a time-delayed feedback. There are also models that include, e.g., a spatial feed-forward mechanism, whose solution is often approximated by replacing the feed-forward term by an exponential term that yields similar qualitatively accurate results. This suggests a mathematical equivalence between time delay and the spatial feed-forward models. Using a WKB approximation to compare numerical steady-state solutions, we show that there is no such simple equivalence. An investigation of the steady-state outputs shows that both models can display sharp tuning, but that the time-delay model requires negative damping for such an effect to occur. Conversely, the feed-forward model provides the most promising results with small positive damping. These results are extended by a careful stability analysis of both models. Here it is shown that whereas a small time delay can stabilize an unstable transmission-line model (with negative damping), that the feed-forward model is stable when the damping is positive. The techniques developed in the paper are directed towards a more comprehensive analysis of nonlinear models.

1. Introduction

Hearing in mammals occurs via a complicated mechanism, in which the key organ in is the cochlea. It responds to sound pressure waves in fluid that are coupled to sound pressure waves in the outside air through the ear drum and the middle ear. Essentially, the cochlea is the stage of the auditory pathway where mechanical vibration is transformed into neural signals, to be transmitted to the brain. It is capable of sensing a wide range of frequencies and amplitudes, with a great frequency and temporal resolution; see [Hudspeth 2008], for example.

The cochlea is also active; that is, it amplifies vibration by 50–60 dB for small amplitudes. The active process displays a so-called compressive nonlinearity in its input-output characteristic; that is, the amplification decreases as the input amplitude increases. Moreover, there is sharp frequency tuning, so that individual spatial locations along the cochlea amplify different input frequencies. In this paper we shall not consider the effect of variation of input amplitude. We instead suppose that the stimuli are sufficiently small so as to, engage the cochlear amplifier in full.

Mathematical modeling of the physiology of the cochlea has a long history dating back to Helmholtz [1885], who suggested that it contains an ensemble of resonators that are sensitive to different frequencies. As the anatomy of the cochlea (illustrated schematically in Figure 1) became better understood, it was
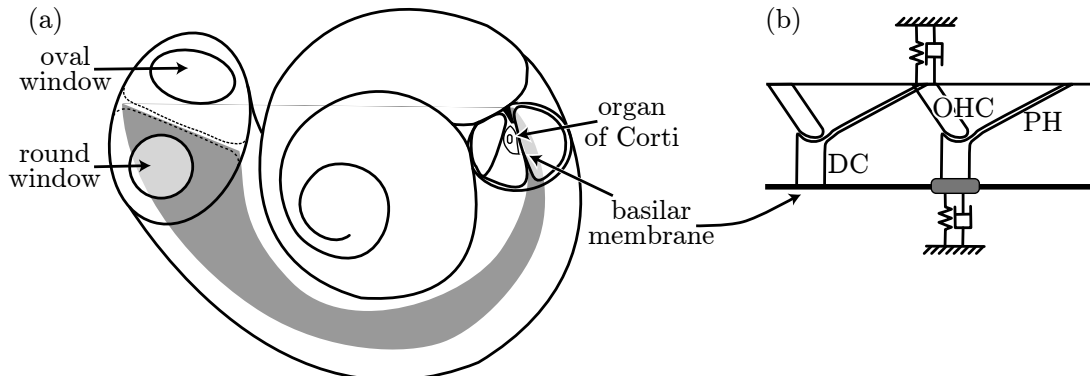
Figure 1. The basic anatomy of the mammalian cochlea. Left: the cochlea in its true representation as a coiled partitioned tube showing the basilar membrane as a shaded surface. Right: longitudinal cross section of the organ of Corti, showing the basilar membrane, the phalangeal processes (PH) of the Deiter's cells and the tilted outer hair cells (OHC).

discovered by Békésy [1960] that the cochlea supports a traveling wave that is the result of the interaction between the fluid and the resonant structures in the organ of Corti. Since then, due to the pioneering work of Charles Steele and others, over more than 35 years, passive models of the fluid-structure interaction have been proposed and embellished; see [Yoon et al. 2009] and references therein. Many of the features of the fine tuning can be explained by the way the geometry of the cochlear tube causes traveling waves to reach a maximum amplitude and stop at a frequency-dependent distance along the cochlea. More sophisticated cochlear models also exist, taking more detailed account of the fluid motion and the precise micromechanics of the cochlea based on physiological data; see [Steele and Lim 1999; Ramamoorthy et al. 2007], for example.

However, there are still significant questions to be answered, such as identifying the mechanism for amplification in the cochlea [Davis 1983; Lagarde et al. 2009]. It is widely held that outer hair cells are responsible for amplification but the mechanism is still unclear; see [Ashmore 2008] for a discussion. In particular, two broad schools of thought suggest that either hair bundle motility [Tinevez et al. 2007] or the piezoelectric elongation of the cell body [Dallos 2008] are primarily responsible. Recent modeling studies suggest that a combination of the two mechanisms may be required in order to explain observed data [Ó'Maoiléidigh and Jülicher 2010; Szalai et al. ≥ 2011]. As a conceptual model of the active process, it has been proposed that the output at a fixed longitudinal position along the basilar membrane resembles that of a nonlinear oscillator that is tuned to be at the onset of a Hopf bifurcation [Eguíluz et al. 2000].

One of the key unsolved problems associated with any active process is to explain how amplification can occur for very low inputs without activating an instability. In this paper we deal with feedback mechanisms in the cochlea that enhance tuning and provide the stability of the active process.

Another motivation for this work is the observation, due to Zweig [1991], that most models of the cochlea cannot qualitatively fit experimental data for frequencies where the decay of the phase of the basilar membrane vibration becomes steep and then flat. In order to explain this inaccuracy, Zweig

used a simple transmission line model and determined the impedance of the organ of Corti necessary to reproduce the data. This extracted impedance implied that, in order to achieve a qualitative agreement with data, a delay term must be included in the transfer function of the organ of Corti. Zweig's results are striking, as there is no obvious mechanism in the cochlea that can result in a pure time delay, independent of the dynamics.

Alternatively, there are other mechanisms that reproduce experimental data just as well, but do not explicitly include a time delay. Hubbard [1993] introduced a secondary traveling wave that could be related to either the tunnel of Corti flow [Karavitaki and Mountain 2007], or the waves of the tectorial membrane [Ghaffari et al. 2007]. Also, tilting of the outer hair cells can introduce a spatial feed-forward mechanism that can similarly explain experimental findings [Geisler 1993; Puria and Steele 2008]. However, not all mammalian species have tilting outer hair cells, but all have a phalangeal processes on their Deiter cells, resulting in longitudinal mechanical coupling (see Figure 1). Yoon and Steele argue that these processes, together with tilting outer hair cells lead to amplification via a push-pull mechanism [Yoon et al. 2009].

The purpose of this paper is neither to resolve this controversy, nor to examine the physiological origins of any active process. Rather we shall use simple mathematical models to examine the generic effects on basilar membrane dynamics of coupling in both space and time, and the relationship between them.

The outline of the rest of this paper is as follows. In Section 2 we introduce simple mathematical models of the cochlea described as a transmission-line. We describe models with time-delay and feed-forward longitudinal coupling. Further, we describe a solution technique that can determine the response of both models to oscillatory forcing, whether stable or unstable, and examine results with particular focus on the relationship between coupling in space and time. Section 3 then addresses the stability of the response of both models. Finally, Section 4 draws conclusions and suggests avenues for future work.

2. Transmission line models

Zweig's model. In the model introduced in [Zweig 1991], the cochlea is modeled as a rectangular box with two ducts of the same and constant cross-section A. The ducts are separated by the flexible basilar membrane, which is assumed in the model to have a single degree of freedom, with displacement $\xi(x, t)$, where x represents the longitudinal distance along the cochlea, measured from the stapes, and t time. The fluid in the ducts is driven by oscillations of the oval window, which generate a pressure difference between the two ducts, $p(x, t)$ and drives the basilar membrane. The equations of motion can then be derived from the one-dimensional Stokes equation assuming that the endolymph/perilymph fluid inside the cochlea is inviscid and incompressible. The fluid mechanics can then be reduced to a single equation for the pressure difference $p(x, t)$ across the membrane:

$$\frac{\partial}{\partial t^2}\xi(x, t) = \frac{\varepsilon^2}{m}\frac{\partial}{\partial x^2}p(x, t), \tag{2-1}$$

where $\varepsilon^2 = 2mA/(\rho\beta)$, m is the surface mass density of the basilar membrane, A is the cross section area of the duct, ρ is the fluid density and β is the width of the membrane. The relation between the pressure difference p and the basilar membrane displacement ξ is also determined by the mechanics of

Parameter	Description	Units	Value
t	Time	ms	$t \in \mathbb{R}$
x	Longitudinal coordinate		$0 \le x \le 1$
ξ	Membrane displacement		
p	Pressure difference	$\mathrm{kg\,ms^{-2}}$	
m	Membrane mass	kg	
A	Duct cross-section		
ρ	Fluid density	kg	
β	Membrane width		
$\theta(x)$	Interpolating function		$\mathrm{e}^{-10x} + \mathrm{e}^{-20(1-x)}$
$\zeta(x)$	Relative damping factor		$0.1\mathrm{e}^{1.1973x}\theta(x) + \zeta_0 \mathrm{e}^{-0.3914x}\,(1-\theta(x))$
ζ_0	Coefficient of $\zeta(x)$		-0.06 or 0.01
$\omega_0(x)$	Natural frequency of the membrane	$\mathrm{ms^{-1}}$	$\left(20.832\mathrm{e}^{-4.8354x} - 0.1455\right) \times 2\pi$
$\sigma(x)$	Feedback stiffness amplitude		$0.1416\mathrm{e}^{-3914x}$
ε	Coupling parameter		$\sqrt{2mA/(\rho\beta)} = 0.006$
$\tau(x)$	Feedback time delay	ms	$1.742 \times 2\pi/\omega_0(x)$

Table 1. Model parameters for (2-5).

the basilar membrane, that is,

$$\frac{p(x,t)}{m} = \frac{\partial^2}{\partial t^2}\xi(x,t) + 2\zeta(x)\omega_0(x)\frac{\partial}{\partial t}\xi(x,t) + \omega_0^2(x)\xi(x,t) + \sigma(x)\omega_0^2(x)\xi(x,t-\tau(x)). \qquad (2\text{-}2)$$

The meaning of the parameters and space-dependent coefficients in (2-1) and (2-2) are given in Table 1 along with the values that we have used in this study. Note that, apart from the damping factor ζ_0, which we allow to vary, these values are the same as those determined in [Zweig 1991] and [Epp et al. 2010], where all distances are measured in the units of the uncoiled cochlea length. We will discuss the physiological motivation of the time-delayed feedback term below.

Using the properties of linear equations, one can assume that a steady time-periodic solution to the model (2-1) and (2-2) can be expressed in the form of a series

$$p(x,t) = \sum_i \bar{p}_i(x)e^{\lambda_i t}, \qquad (2\text{-}3)$$

which is a separation of variables with exponential time dependence and spatial patterns $\bar{p}_i(x)$. Substituting this expansion into (2-1) and (2-2), we obtain

$$\lambda_i^2\bar{\xi}_i = \frac{\varepsilon^2}{m}\frac{\partial}{\partial x^2}\bar{p}_i \quad \text{and} \quad \frac{\bar{p}_i}{m} = \kappa^2(\lambda_i, x)\bar{\xi}_i, \qquad (2\text{-}4)$$

where

$$\kappa^2(\lambda_i, x) = \lambda_i^2 + 2\lambda_i\zeta(x)\omega_0(x) + \omega_0^2(x)\left(1 + \sigma(x)e^{-\lambda\tau(x)}\right).$$

Combining the two transformed equations (2-4) by eliminating $\bar{\xi}_i$ we get

$$\lambda_i^2\bar{p}_i = \varepsilon^2\kappa^2(\lambda_i, x)\frac{\partial}{\partial x^2}\bar{p}_i. \qquad (2\text{-}5)$$

This equation has to hold for each $\bar{p}_i$ and λ_i, where as yet we have not determined which values λ_i must take. (Note that our notation here is different from that of [Zweig 1991, Equation 24], and that we use a linear longitudinal coordinate. However, the essence of the model is the same.)

The natural boundary condition to take at the apex of the cochlea, where fluid is allowed to flow between the two chambers divided by the basilar membrane, is

$$p(1, 0) = 0. \tag{2-6}$$

Typically the choice of λ_i's is determined by the boundary condition at the stapes, which we assume to take the form $\frac{\partial}{\partial x} p(0, t) = -\rho v_{st}(t)$, where v_{st} is the stapes velocity.

In this study we shall assume a single frequency sound input with constant phase. It is most convenient to solve in complex coordinates, and therefore we set

$$\frac{\partial}{\partial x} p(0, t) = -\rho v_{st}(t) = e^{i\omega t}. \tag{2-7}$$

The form of the boundary conditions (2-6) and (2-7) means that we can choose a single $\lambda = \lambda_0 = i\omega$ and a single $\bar{p} = \bar{p}_0$ and hence the boundary conditions on $\bar{p}$ become

$$\frac{\partial}{\partial x} \bar{p}(0) = 1, \qquad \bar{p}(1) = 0. \tag{2-8}$$

Hence we have reduced the dynamic problem to that of finding a steady solution to the two-point boundary value problem (2-5) for $\overline{p}(x)$ subject to the boundary conditions (2-8). In order to solve such a problem we have used Chebyshev collocation [Trefethen 2000] with 2400 mesh points in x, which yields very high accuracy.

A typical numerical solution for the excitation pattern of the cochlea, defined as the maximum amplitude of the basilar membrane at each longitudinal position x, is shown in Figure 2 for the particular choice of stimulus frequency 1.6kHz. Solutions with nonzero time delay are represented by continuous green lines, and without time-delay ($\tau(x) = 0$) by dotted black lines. Similar results are obtained for a variety of different input frequencies with the sharp peak occurring at a frequency-dependent distance.

It can be seen that with negative damping $\zeta_0 = -0.06$ (left column of Figure 2), the tuning is sharper than with small positive damping ($\zeta_0 = 0.01$; right column of Figure 2). We also can see that when the delay is removed there is no sharp tuning. Moreover, the phase curves (bottom row of Figure 2) suggest that the zero-delay case causes instability in the presence of negative damping, as the phase increases rather than decreases with x. Note that a decrease in phase is not a sufficient condition for stability, and a thorough stability analysis needs to be carried out to assess the stability. Such an analysis forms the subject of Section 3.

Delay equivalency. It can be argued that time delay in the active part of the cochlea, although it cannot be directly related to a physiological effect, is in some sense equivalent to a spatial feed-forward mechanism that captures the action of the tilted hair cells and the phalangeal processes. The logic for this argument is that the cochlea supports a traveling wave whose speed is easily determined. From this speed, and starting from a time-delay model, one can calculate the distance that the wave travels during a fixed delay interval. This distance should then be equal to the spatial constant of the feed-forward mechanism.
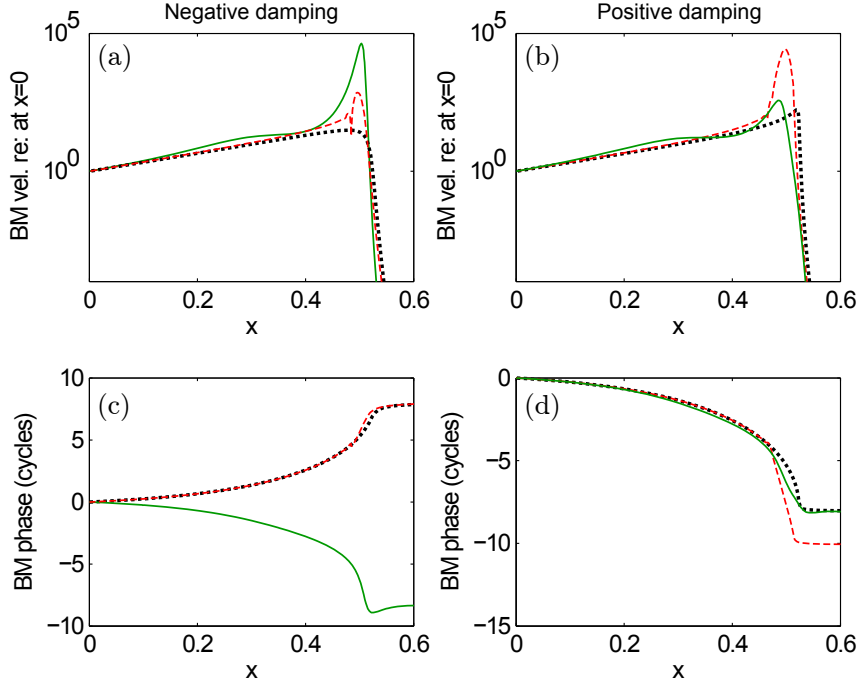
Figure 2. Steady state solutions of cochlea models for 1.6kHz harmonic stimulus: continuous green curves correspond to the time-delay model (2-5), dotted black lines correspond to the time-delay model with $\tau(x) = 0$, and dashed red lines corresponds to the feed-forward model (2-10). The damping parameter is $\zeta_0 = -0.06$ for the left column and $\zeta_0 = 0.01$ for the right column.

The reason for studying the temporal delay model is that it is much simpler to solve using the method we introduced in the previous section, and the model with delay can reproduce the cochlear response qualitatively better than without delay, as shown by the continuous green curves in Figure 2. It is useful then to consider the relation between the time-delay and feed-forward models in more detail.

We use the WKB technique to solve (2-5). The approximate solution assumes the form of $\bar{p}(x) = \sinh\phi(x)/\varepsilon$ with $\phi(x) = -\int_x^1 \frac{\lambda}{\kappa(\lambda,z)}dz$. The wavelength of the solution becomes

$$\delta(x) = \frac{2\pi\varepsilon}{\operatorname{Im}\dfrac{i\omega_0(x)}{\kappa(i\omega_0(x),x)}}.$$

With knowledge of the spatial wavelength, it is possible to calculate the distance that will have the same phase as the delayed wave, $\bar{p}(x)e^{-\lambda\tau}$, namely $\bar{p}(x + \Delta(x))$, where

$$\Delta(x) = \omega_0(x)\tau(x)\delta(x). \tag{2-9}$$

Note that since $\tau(x)\omega(x)$ is a constant, $\Delta(x)$ is a constant multiple of $\delta(x)$, and note further that $\delta(x)$ is roughly constant along the cochlea. For our parameter values, the estimated feed-forward distance $\Delta(x)$
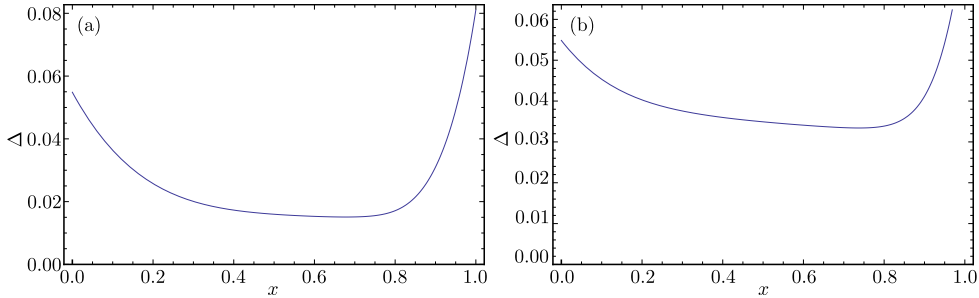
Figure 3. The equivalent feed-forward distance $\Delta(x)$ along the cochlea. The damping parameter is $\zeta = -0.06$ (a) and $\zeta = 0.01$ (b).

defined by (2-9) is plotted in Figure 3. The obtained distance values are unreasonably large, around 2% of the length of the uncoiled cochlea, which would mean 0.75 mm for the 35 mm long human cochlea.

Furthermore, this approximation of the feedback term is inaccurate. It does not take into account that the wave is decaying both to the left and to the right of its peak and hence a solution with smaller amplitude will be fed back. It also ignores the fact that the feed-forward mechanism changes the wave speed. We therefore proceed to study a model which explicitly includes a spatial feed-forward term, instead of time-delay.

Feed-forward model. Replacing the time delay in (2-5) with a spatial delay of distance $\Delta(x)$ yields the feed-forward model. In some respect this is more physiological, but it still does not properly take into account the underlying source of the feed-forward term. The governing equation after expanding the solution into the form (2-3) becomes

$$\lambda^2 \bar{p}(x) = \varepsilon^2 \kappa_0^2(\lambda, x) \frac{\partial}{\partial x^2} \bar{p}(x) + \kappa_1^2(x) \frac{\partial}{\partial x^2} \bar{p}(x + \Delta(x)), \tag{2-10}$$

where $\kappa_0^2(\lambda, x) = \lambda^2 + 2\lambda \zeta(x)\omega_0(x) + \omega_0^2(x)$ and $\kappa_1^2(x) = \omega_0^2(x)\sigma(x)$. The excitation pattern compared to the time-delay model can be seen in Figure 2. With the obtained value of $\Delta(x)$ in Figure 3 we were not able to obtain a plausible result, which is already an indication that the derived relation in (2-9) is inaccurate. Indeed, when looking for an exact value of distance $\Delta(x)$ and a constant $\hat{\sigma}$ that replaces the feedback coefficient σ, we find that $\Delta(x)$ and $\hat{\sigma}$ heavily depend on the stimulus frequency, and it is impossible to use fixed values for all frequencies. Nevertheless, we find from our simulations (indicated by dashed red lines in Figure 2) that a spatial delay of a quarter of the (inaccurately) estimated distance $\Delta(x)/4$ gives responses with the expected tall and broad peaks if ζ is positive.

It can also be seen in Figure 2 that the feed-forward effect does not stabilize the model with negative damping and has less sharp tuning. However, with small positive damping the tuning is sharp and stability appears to be preserved. Note that stability in this context means only stability with respect to single frequency response. A more thorough stability analysis is outlined in the next section.

3. Stability

Stability of the delayed model. Stability calculations for a delay equation can be demanding, depending on the type of the equation [Stépán 1989]. For delayed ordinary differential equations (DDEs) the

problem reduces to finding roots of an exponential polynomial, followed by a check that all of these (typically infinitely many) roots have negative real part to ensure stability. In a numerical scheme one can count the number of unstable roots using the so-called argument principle [Marsden and Hoffman 1999]. The argument principle states that along a contour in the complex plane the argument variation of the function will be a constant times the difference between the number of poles and roots of the function inside the contour. This is very helpful, but only in the case when the function does not have poles.

Here we construct a function (the characteristic function) for the time-delayed cochlea model, whose roots determine the stability. However, since our underlying model is a delayed partial differential equation, it is not possible to guarantee that this function is without poles.

The construction of the characteristic function goes as follows. We say that λ is a characteristic root of (2-5) if it has a nontrivial solution that satisfies the boundary conditions $(\partial \bar{p}/\partial x)(0) = 0$ and $\bar{p}(1) = 0$. Since (2-5) is an ODE, one can solve it uniquely by specifying two boundary conditions $\bar{p}(0) = a_1$ and $(\partial \bar{p}/\partial x)(1) = a_2$. If for nonzero a_1 and a_2 the conditions $(\partial \bar{p}/\partial x)(0) = 0$ and $\bar{p}(1) = 0$ hold, there is a nontrivial solution and hence λ is a characteristic root.

One can therefore generate a characteristic matrix M such that

$$\begin{pmatrix} \dfrac{\partial}{\partial x} \bar{p}(0) \\[2mm] \bar{p}(1) \end{pmatrix} = M(\lambda) \begin{pmatrix} \dfrac{\partial}{\partial x} \bar{p}(1) \\[2mm] \bar{p}(0) \end{pmatrix}.$$

The left side of this equation can only be zero for nonzero right-hand side vector if $\det M(\lambda) = 0$. Hence, the roots of $\det M(\lambda) = 0$ are the characteristic roots of (2-5). Since this is a delayed partial differential equation there may be a continuous spectrum, so finding the roots is a challenge. One can however approximate the roots by finding local maxima of the function $1/\det M(\lambda)$ that are above a certain threshold.

We solve the boundary value problem with a central finite difference scheme with 1200 equidistant points that is a rather fast and moderately accurate method. A plot of the roots can be seen as the black dots in Figure 4. In the nondelay case, we can also approximate the roots by a (slower but more accurate) direct method using Chebyshev collocation to check the accuracy of our computation; the results are the color dots in Figure 4.

We see that in the case of negative damping ($\zeta_0 = -0.06$, shown in the left column of Figure 4) adding time delay stabilizes the system. On the other hand, with positive damping ($\zeta_0 = 0.01$, shown in the right column of Figure 4) the solution is always stable, with or without time-delay.

We note that both numerical methods provide spectra that are essentially indistinguishable in the case of no time-delay, which gives confidence in the accuracy of our computations. Furthermore, every root of our characteristic function is accompanied by a pole, therefore counting the difference between the number of roots and poles gives identically zero result, hence the argument principle cannot be used.

Stability of the feed-forward model. The stability of the feed-forward model can be calculated more efficiently than that of the time-delay model. In order to formulate the spectral problem we introduce the differential operators

$$D_0^2 : (D_0^2 \varphi)(x) = \frac{\partial}{\partial x^2} \varphi(x), \quad \varphi \in \left\{ \phi \in C^2[0, 1] : \frac{\partial}{\partial x} \phi(0) = 0, \ \phi(1) = 0 \right\}$$
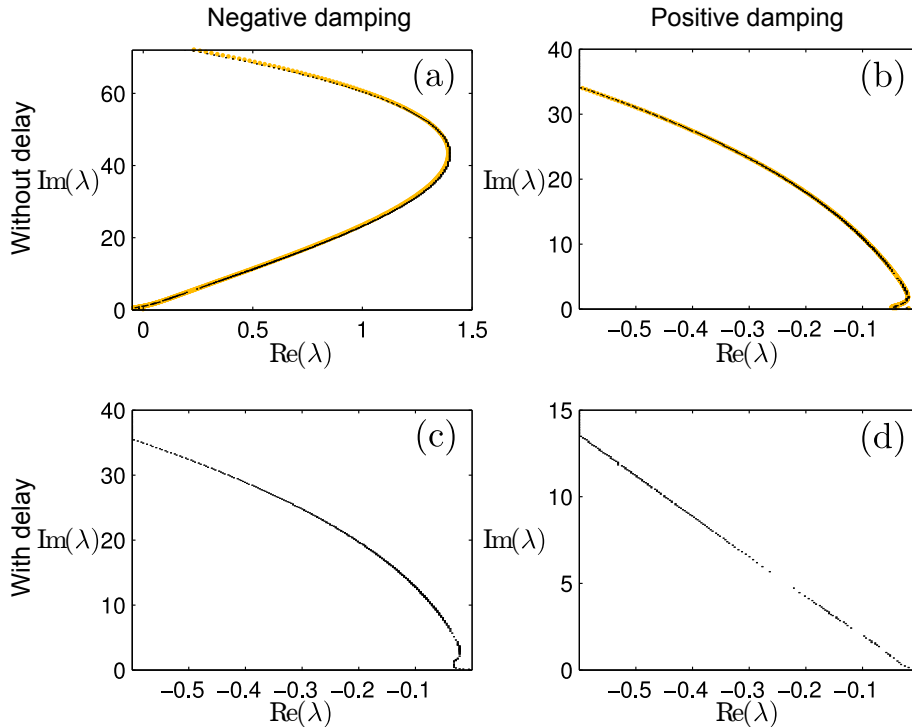
Figure 4. Stability of the time-delay model. Black points show the characteristic roots of (2-5), computed from the characteristic function $\det M = 0$, without delay (top) and with delay (bottom). In the nondelay case, the roots are also computed with a direct method using Chebyshev collocation, as a check; these values are indicated with color dots. The damping parameter is $\zeta_0 = -0.06$ for the left column and $\zeta_0 = 0.01$ for the right column.

and

$$D_\Delta^2 \; : \; (D_\Delta^2 \varphi)(x) = \frac{\partial}{\partial x^2}\varphi(x + \Delta(x)), \quad \varphi \in \left\{\phi \in C^2[0,1] \; : \; \frac{\partial}{\partial x}\phi(0) = 0, \; \phi(1) = 0\right\},$$

noting that the boundary conditions for (2-10) are $(\partial \bar{p}/\partial x)(0) = 0$ and $\bar{p}(1) = 0$. The spectral problem is not an ODE any more, but is instead a neutral delay-differential equation. In contrast to the time-delay case the equation is polynomial in λ, and therefore can be transformed into a regular eigenvalue/eigenfunction problem:

$$\left[\begin{pmatrix} 0 & I \\ \varepsilon^2\omega_0^2\left(D_0^2 + \sigma D_\Delta^2\right) & 2\varepsilon^2\zeta\omega_0 D_0^2 \end{pmatrix} - \lambda\begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}\right]\begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{3-1}$$

We discretize this boundary problem with Chebyshev collocation, using 2400 points, exactly as for the steady state solution, and put into an eigenvalue solver to obtain the spectrum.

The spectrum of (3-1) is shown in Figure 5. We see that the feed-forward mechanism does not stabilize the system with negative damping, but that stability is preserved with the combination of positive damping and spatial feed forward.
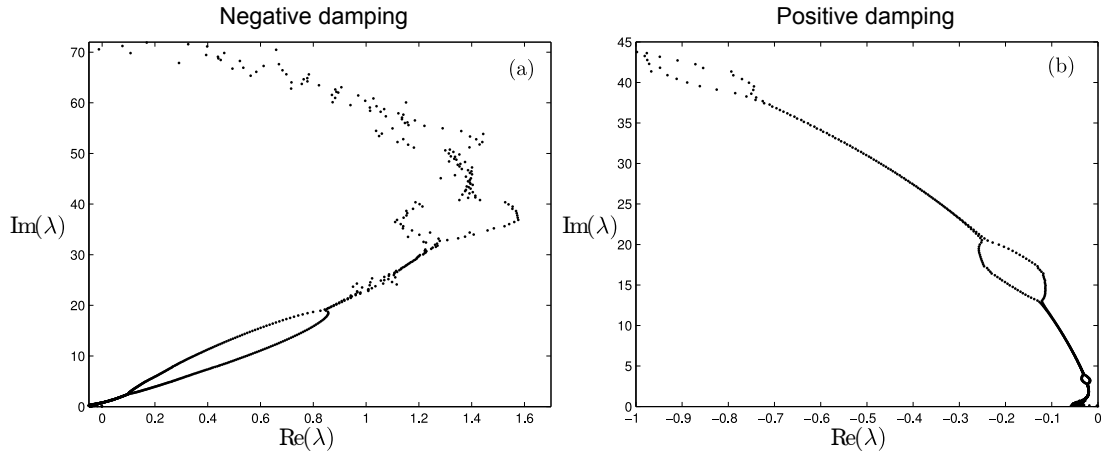
Figure 5. Characteristic roots of the feed-forward model (2-10). The damping parameter is $\zeta = -0.06$ (a) and $\zeta = 0.01$ (b).

4. Conclusion

In this paper we have compared two cochlea models, one with time delay, and another with a spatial feed-forward mechanism. We calculated an approximate relation between the temporal and spatial delays and found no simple relation between the two such models can exist. Moreover, we showed that the two models do not yield similar excitation profiles for equivalent physiological parameters. From a mathematical point of view, it is easier to control stability with a time-delay than with the feed-forward mechanism. In particular, we were not able to find a value for the feed-forward distance that made the model stable. However, the restriction on the damping does not restrict the sharpness of tuning with a stable model. In fact, with small positive damping one can achieve sharp tuning and qualitatively good agreement with data, as also shown in [Yoon et al. 2009].

Furthermore, our stability analysis has given information on the alignment of the characteristic roots along the complex imaginary axis, which might also give insight into spontaneous otoacoustic emissions and even explain synchronization of the emitted frequencies. The distance of the roots from the imaginary axis might also indicate the sensitivity to different frequencies.

A natural next step that will be investigated in future work would be to introduce a nonlinear model for the outer hair cells themselves. Introduction of such a nonlinear term into the transmission line equations would enable us to investigate the effect of the loss of stability on the dynamics. More importantly, we would like to understand what kinds of local hair cell dynamics lead to compressive active nonlinearity when the spatial feed-forward or temporal delay are taken into account. In particular, through matching with experimental data, this would enable us to gain a better global understanding of the relative importance of the longitudinal effects considered in this paper and the local hair cell dynamics.

Acknowledgements

References

[Ashmore 2008] J. Ashmore, "Cochlear outer hair cell motility", *Physiol. Rev.* **88**:1 (2008), 173–210.

[Békésy 1960] G. von Békésy, *Experiments in hearing*, McGraw-Hill, New York, 1960.

[Dallos 2008] P. Dallos, "Cochlear amplification, outer hair cells and prestin", *Curr. Opin. Neurobiol.* **18**:4 (2008), 370–376.

[Davis 1983] H. Davis, "An active process in cochlear mechanics", *Hearing Res.* **9**:1 (1983), 79–90.

[Eguíluz et al. 2000] V. M. Eguíluz, M. Ospeck, Y. Choe, A. J. Hudspeth, and M. O. Magnasco, "Essential nonlinearities in hearing", *Phys. Rev. Lett.* **84**:22 (2000), 5232–5235.

[Epp et al. 2010] B. Epp, J. L. Verhey, and M. Mauermann, "Modeling cochlear dynamics: interrelation between cochlea mechanics and psychoacoustics", *J. Acoust. Soc. Am.* **128**:4 (2010), 1870–1883.

[Geisler 1993] C. D. Geisler, "A realizable cochlear model using feedback from motile outer hair cells", *Hearing Res.* **68**:2 (1993), 253–262.

[Ghaffari et al. 2007] R. Ghaffari, A. J. Aranyosi, and D. M. Freeman, "Longitudinally propagating traveling waves of the mammalian tectorial membrane", *Proc. Nat. Acad. Sci. USA* **104**:42 (2007), 16510–16515.

[Helmholtz 1885] H. Helmholtz, *Die Lehre von den Tonempfindungen*, 4th ed., Longmans, London, 1885. Translated by A. J. Ellis in *On the sensations of tone as a physiological basis for the theory of music*, 2nd English ed., Dover, New York, 1954.

[Hubbard 1993] A. Hubbard, "A traveling-wave amplifier model of the cochlea", *Science* **259**:5091 (1993), 68–71.

[Hudspeth 2008] A. J. Hudspeth, "Making an effort to listen: mechanical amplification in the ear", *Neuron* **59**:4 (2008), 530–545.

[Karavitaki and Mountain 2007] K. D. Karavitaki and D. C. Mountain, "Evidence for outer hair cell driven oscillatory fluid flow in the tunnel of Corti", *Biophys. J.* **92**:9 (2007), 3284–3293.

[Lagarde et al. 2009] M. M. M. Lagarde, M. Drexl, V. A. Lukashkina, A. N. Lukashkin, and I. J. Russell, "Determining the identity of the cochlear amplifier: electrical stimulation of the *Tecta* mouse cochlea", pp. 106–112 in *Concepts and challenges in the biophysics of hearing* (Staffordshire, 2008), edited by N. P. Cooper and D. T. Kemp, World Scientific, Singapore, 2009.

[Marsden and Hoffman 1999] J. E. Marsden and M. J. Hoffman, *Basic complex analysis*, 3rd ed., W. H. Freeman, New York, 1999.

[Ó'Maoiléidigh and Jülicher 2010] D. Ó'Maoiléidigh and F. Jülicher, "The interplay between active hair bundle motility and electromotility in the cochlea", *J. Acoust. Soc. Am.* **128**:3 (2010), 1175–1190.

[Puria and Steele 2008] S. Puria and C. R. Steele, "Mechano-acoustical transformations", pp. 165–202 in *The senses: a comprehensive reference, 3: Audition*, edited by P. Dallos and D. Oertel, Elsevier, Amsterdam, 2008.

[Ramamoorthy et al. 2007] S. Ramamoorthy, N. V. Deo, and K. Grosh, "A mechano-electro-acoustical model for the cochlea: response to acoustic stimuli", *J. Acoust. Soc. Am.* **121**:5 (2007), 2758–2773.

[Steele and Lim 1999] C. R. Steele and K.-M. Lim, "Cochlear model with three-dimensional fluid, inner sulcus and feed-forward mechanism", *Audiol. Neurotol.* **4**:3-4 (1999), 197–203.

[Stépán 1989] G. Stépán, *Retarded dynamical systems: stability and characteristic functions*, Pitman Research Notes in Mathematics Series **210**, Longman, Harlow, 1989.

[Szalai et al. ≥ 2011] R. Szalai, D. Ó'Maoiléidigh, H. Kennedy, N. P. Cooper, A. R. Champneys, and M. Homer, "On the origins of the compressive cochlear nonlinearity", in preparation.

[Tinevez et al. 2007] J.-Y. Tinevez, F. Jülicher, and P. Martin, "Unifying the various incarnations of active hair-bundle motility by the vertebrate hair cell", *Biophys. J.* **93**:11 (2007), 4053–4067.

[Trefethen 2000] L. N. Trefethen, *Spectral methods in MATLAB*, Software, Environments, and Tools **10**, SIAM, Philadelphia, 2000.

[Yoon et al. 2009] Y. Yoon, S. Puria, and C. R. Steele, "A cochlear model using the time-averaged Lagrangian and the push-pull mechanism in the organ of Corti", *J. Mech. Mater. Struct.* **4**:5 (2009), 977–986.

[Zweig 1991] G. Zweig, "Finding the impedance of the organ of Corti", *J. Acoust. Soc. Am.* **89**:3 (1991), 1229–1254.

ROBERT SZALAI: r.szalai@bristol.ac.uk
Department of Engineering Mathematics, University of Bristol, Queen's Building, University Walk, Bristol, BS8 1TR,
United Kingdom

BASTIAN EPP: bastian.epp@uni-oldenburg.de
Institute of Physics, Neuroacoustics, Carl von Ossietzky Universität Oldenburg, Carl-von-Ossietzky-Str. 9-11,
26111 Oldenburg, Germany

ALAN R. CHAMPNEYS: a.r.champneys@bristol.ac.uk
Department of Engineering Mathematics, University of Bristol, Queen's Building, University Walk, Bristol, BS8 1TR,
United Kingdom

MARTIN HOMER: Martin.Homer@bristol.ac.uk
Department of Engineering Mathematics, University of Bristol, Queen's Building, University Walk, Bristol, BS8 1TR,
United Kingdom

■
■msp

A POROELASTIC MODEL FOR CELL CRAWLING INCLUDING MECHANICAL COUPLING BETWEEN CYTOSKELETAL CONTRACTION AND ACTIN POLYMERIZATION

LARRY A. TABER, YUNFEI SHI, LE YANG AND PHILIP V. BAYLY

Much is known about the biophysical mechanisms involved in cell crawling, but how these processes are coordinated to produce directed motion is not well understood. Here, we propose a new hypothesis whereby local cytoskeletal contraction generates fluid flow through the lamellipodium, with the pressure at the front of the cell facilitating actin polymerization which pushes the leading edge forward. The contraction, in turn, is regulated by stress in the cytoskeleton. To test this hypothesis, finite element models for a crawling cell are presented. These models are based on nonlinear poroelasticity theory, modified to include the effects of active contraction and growth, which are regulated by mechanical feedback laws. Results from the models agree reasonably well with published experimental data for cell speed, actin flow, and cytoskeletal deformation in migrating fish epidermal keratocytes. The models also suggest that oscillations can occur for certain ranges of parameter values.

1. Introduction

Cell motility plays an important role in many biological processes, including the response to disease and injury, morphogenesis, and the progression of cancer (see [Stossel 1994], for example). Much is known about the molecular and physical mechanisms involved in cell crawling, but how these mechanisms combine to produce directed motion remains incompletely understood. According to the traditional view, cell crawling involves three sequential steps [Flaherty et al. 2007]: (1) actin polymerization pushes a lamellipodium forward from the leading edge; (2) the lamellipodium adheres to the substrate; and (3) cytoskeletal contraction generates tension that breaks weaker adhesions at the rear of the cell and pulls the cell body forward.

Recent work has questioned the order of these events, however, suggesting that contraction precedes polymerization and begins the crawling cycle [Yam et al. 2007; Loitto et al. 2009]. The reason that contraction may be needed to initiate normal cell crawling is unknown. One possibility is that it generates hydrostatic pressure, which helps drive the protrusion at the front of the cell [Charras et al. 2005; Yam et al. 2007]. This idea is a key element of the present study.

The notion that fluid pressure is an important factor in cell crawling is not new. Diffusion alone may not provide a forward flow of actin monomers sufficiently fast to give observed protrusion rates [Oster and Perelson 1987; Zicha et al. 2003], and researchers have speculated that pressure may foster this flow, as well as drive lamellipodial extension [Zhu and Skalak 1988; Stossel 1994; Keren et al. 2009].

This idea has fallen out of favor for most cell types, however, as mounting evidence indicates that actin polymerization is the dominant mechanism in lamellipodial extension [Bray 2001]. On the other hand, recent studies have shown that substantial pressure gradients and fluid flow can occur in cells [Charras et al. 2005; Iwasaki and Wang 2008; Mitchison et al. 2008; Keren et al. 2009], suggesting that fluid dynamics may play a larger role in cell motility than commonly thought.

The purpose of this paper is to investigate a new hypothesis for cell crawling that integrates the effects of actin polymerization, cytoskeletal contractility, and intracellular fluid flow. To illustrate the plausibility of our hypothesis, these mechanisms are included in 1-D and 2-D continuum models of a migrating cell. Here, the main focus is on fish epidermal keratocytes, which move relatively quickly (about 0.05–0.2 μm/s) while maintaining a nearly constant shape, in contrast to most other types of cells, which exhibit oscillatory behavior [Bray 2001; Keren et al. 2008]. For realistic parameter values, results from the models agree reasonably well with published experimental data for cell speed, actin flow, and cytoskeletal deformation. In addition, the models predict oscillations (similar to other cell types) for certain parameter values, thereby illuminating some of the inherent feedback control problems that apparently have been solved by crawling keratocytes in order to reach a steady state.

2. Preliminaries

General characteristics of crawling cells. The shape of an epidermal keratocyte (EK) depends on whether it is stationary or crawling, as well as on whether it is located on a 2-D surface or in a 3-D matrix. Stationary cells on a surface resemble the contents of an egg poured onto a flat surface (Figure 1b). The hemispherical yolk in the center corresponds to the cell body, which contains the nucleus and other organelles. The egg white, spread out into a thin disk around the yolk, represents the lamellipodium, which is filled with actin and myosin.

Some cells crawl in a cyclic manner like a worm, e.g., amebas and fibroblasts. They first extend their front end, grab the substrate, and then pull their rear ends forward. In contrast, EKs move more like a centipede, which crawl without changing shape (except to turn). Unlike a centipede, however, EKs
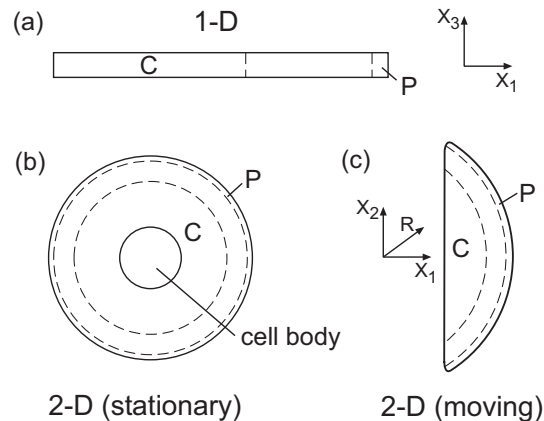
Figure 1. Models for crawling cells (P = actin polymerization zone of lamellipodium; C = CSK contraction zone). (a) 1-D model. (b) Stationary 2-D model. (c) Crawling 2-D model.

apparently are not propelled by dozens of tiny legs. How these cells can crawl steadily at speeds up to $10\,\mu$m/min or more is an intriguing and unanswered question.

While crawling on a surface at steady state, the EK lamellipodium becomes shaped like a fan or wing (approximately $40\,\mu$m wide, $10\,\mu$m long, and $0.2\,\mu$m thick; see Figure 1c). Experiments have shown that lamellipodia isolated from the cell body can crawl on their own with a similar speed and shape as intact cells (see Figure 6a) [Verkhovsky et al. 1999]. Hence, this structure is the engine that drives the cell, with the cell body pulled along behind it. The cell body, therefore, can be neglected in models for crawling EKs.

For technical reasons, most experimental studies and models of cell crawling have been two-dimensional. Crawling in three dimensions is quite different, however. For example, cells moving through a 3-D matrix have fewer stress fibers and adhesions, and they must create space in which to move [Friedl and Bröcker 2000]. Moreover, whereas cells on surfaces spread out into a relatively flat configuration, they become spindle shaped as they crawl though tissue. Thus, a 1-D model may be a good approximation for a cell crawling through 3-D tissue.

The cytoskeleton of crawling cells. Cytoskeletal (CSK) dynamics are key to the migration process. Within a relatively narrow region (about 1–$2\,\mu$m wide) near the leading edge, the lamellipodium is filled with a relatively dense meshwork of branched actin filaments (F-actin). Actin at the front of this region polymerizes to push the cell membrane forward, while actin at the rear depolymerizes into monomers (G-actin), which then move forward to provide a pool for further polymerization. This sets up a treadmilling mechanism [Bray 2001].

Behind this network, F-actin is organized into bundles, and interaction with myosin generates contraction in the crawling direction. Researchers have speculated that this contraction enhances the rate of actin depolymerization and may drive fluid, along with G-actin, toward the leading edge [Svitkina et al. 1997; Zicha et al. 2003; Yam et al. 2007; Keren et al. 2009; Okeyo et al. 2009]. It is notable that actin density decreases and myosin density increases from front to back, and a relatively thick stress fiber forms along the back edge of the lamellipodium.

Proposed crawling mechanism. Combining previous ideas, we herein propose a relatively simple crawling mechanism that integrates the effects of actin polymerization and cytoskeletal contraction. First, CSK contraction elevates the fluid pressure at the rear of the cell, driving fluid (and actin monomers) forward. Next, pressure at the front of the cell pushes the cell membrane forward, creating extra space for monomers and thereby increasing the rate of actin polymerization at the leading edge. Finally, in conjunction with cell-substrate adhesion, the tension generated in the contracting CSK pulls the cell forward.

To achieve steady-state motion, these processes must be continuous and balanced. Hence, coupling between these processes is crucial to the global dynamic behavior of the cell. We suggest that this coupling is regulated by mechanical feedback that depends on the interaction between fluid pressure, CSK stress, and adhesive forces.

3. Governing equations

To model fluid-solid interaction, we take the cytoplasm as a porous elastic material (the CSK) saturated with a viscous fluid, i.e., a poroelastic material. Laying the foundation for the model, this section summarizes the governing equations of nonlinear poroelasticity theory. Then, modifications are discussed for

including growth, which is used to simulate actin polymerization and CSK contraction. The presentation is relatively brief; further details on poroelasticity can be found in [Biot 1972; Kenyon 1976; Yang et al. 1994], while detailed discussions of the growth theory are given, for example, in [Rodriguez et al. 1994; Taber 2001].

Nonlinear poroelasticity theory. In poroelasticity theory, as in the more general mixture theory [Bowen 1976], all points in the bulk material are assumed to be occupied simultaneously by both solid and fluid particles. Here, both the fluid and solid phases are assumed to be incompressible. The total Cauchy stress tensor is given by

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}_s + \boldsymbol{\sigma}_f,$$

where the respective partial stresses of the solid and fluid (per unit area of the bulk material in the current configuration) are

$$\boldsymbol{\sigma}_s = \phi_s \hat{\boldsymbol{\sigma}}_s, \quad \boldsymbol{\sigma}_f = -\phi_f p \, \boldsymbol{I}.$$

In these relations, ϕ_s and ϕ_f are volume fractions, $\hat{\boldsymbol{\sigma}}_s$ is the true solid stress, p is the fluid (pore) pressure, and $\boldsymbol{I}$ is the identity tensor. If there is no mass exchange with the surroundings, $\phi_s + \phi_f = 1$.

Computations are based on a Lagrangian formulation, with all quantities defined relative to the initial (reference) configuration for a cell at rest (taken as stress free). As described below, growth is simulated by introducing an intermediate zero-stress configuration, but we disregard this for the moment. With position vectors to a material point of the solid in the initial and current configurations given by $\boldsymbol{X}$ and $\boldsymbol{x} = \boldsymbol{x}(\boldsymbol{X}, t)$, respectively, the total deformation gradient tensor for the solid skeleton is

$$\boldsymbol{F} = (\nabla \boldsymbol{x})^T, \tag{3-1}$$

where $\nabla \equiv \partial / \partial \boldsymbol{X}$ and superscript T denotes the transpose. The corresponding Lagrange strain tensor is $\boldsymbol{E} = \frac{1}{2}(\boldsymbol{F}^T \cdot \boldsymbol{F} - \boldsymbol{I})$.

The total first Piola–Kirchhoff stress tensor $\boldsymbol{P}$, defined relative to the bulk area in the initial state, is given by [Taber 2004]

$$\boldsymbol{P} = J \boldsymbol{F}^{-1} \cdot \boldsymbol{\sigma}, \tag{3-2}$$

where $J = \det \boldsymbol{F}$ is the dilatation ratio. The equilibrium and constitutive equations can be written in the forms

$$\nabla \cdot \boldsymbol{P} + \boldsymbol{f}_0 = \boldsymbol{0}, \tag{3-3}$$

$$\boldsymbol{P} = \frac{\partial W}{\partial \boldsymbol{F}^T} - J p \, \boldsymbol{F}^{-1}, \tag{3-4}$$

where $\boldsymbol{f}_0$ and W are the respective body force vector and strain-energy density function per unit bulk reference volume.

Fluid flow is governed by a generalized Darcy's law of the form [Yang et al. 1994]

$$\dot{\boldsymbol{U}} = -\boldsymbol{K} \cdot \nabla p, \tag{3-5}$$

where $\dot{\boldsymbol{U}}$ is the volume flow rate of fluid relative to the solid, per unit bulk material area in the reference state, and dot denotes time differentiation. In terms of $\dot{\boldsymbol{U}}$, the relative fluid velocity is $\boldsymbol{v}_f = J^{-1} \dot{\boldsymbol{U}} \cdot \boldsymbol{F}^T$. In addition,

$$\boldsymbol{K} = \boldsymbol{F}^{-1} \cdot \boldsymbol{k} \cdot \boldsymbol{F}^{-T} J / \mu_f$$

is the effective permeability tensor, with k being the intrinsic permeability tensor and μ_f the fluid viscosity.

Finally, for incompressible fluid and solid constituents, equating the net volume flow rate of fluid entering a material element to the rate of change of element volume gives the incompressibility (continuity) condition

$$-\nabla \cdot \dot{U} = \dot{J}.$$

This equation can be integrated to yield

$$\nabla \cdot U + J - 1 = 0. \tag{3-6}$$

Given appropriate boundary and initial conditions, equations (3-1), (3-3), (3-4), (3-5) and (3-6) can be solved for x, F, P, U, and p. Note that if $K = 0$, then $U = 0$ and these equations reduce to the governing equations for an incompressible solid, with p becoming a Lagrange multiplier [Taber 2004].

Growth. Polymerization and contraction are simulated using a theory for volumetric growth of an elastic solid [Rodriguez et al. 1994; Taber 2001]. Positive growth along an actin fiber simulates polymerization, while negative growth simulates contraction.

In this theory, the growth tensor G transforms the zero-stress reference state for each material element into an intermediate (virtual) zero-stress state of different size and shape. Geometric compatibility between growing elements is ensured by the elastic deformation gradient tensor F^*, and the total deformation gradient tensor is written is the form

$$F = F^* \cdot G. \tag{3-7}$$

With this formulation, growth occurs relative to the reference configuration. In addition, stress depends only on F^*, and the Cauchy stress tensor is given by [Taber 2001]

$$\sigma = J^{*-1} F^* \cdot \frac{\partial W}{\partial E^*} \cdot F^{*T} - p\, I, \tag{3-8}$$

from which (3-2) provides P. In the resulting expression, which replaces (3-4), $J^* = \det F^*$, $W = W(E^*)$, and $E^* = \frac{1}{2}(F^{*T} \cdot F^* - I)$.

In our models, the growth tensor is determined by mechanical feedback laws. These laws have the general rate form $\dot{G} = \dot{G}(X, t; F; \hat{\sigma}_s, p)$, which is specialized below.

4. Proposed models for cell crawling

In this paper, we present 1-D and 2-D poroelastic models for a cell crawling on a flat surface S. To define the geometry, let X_i represent a Cartesian coordinate system fixed to S, with X_1 directed along the crawling direction and X_3 normal to S.

Geometry and material properties. The 1-D model represents a narrow strip along the centerline of the lamellipodium in the X_1-X_3 plane. In the initial resting state, the cell is taken as a rectangular bar (or plate) of length $10\,\mu$m and height $1\,\mu$m (Figure 1a). This model also can be considered a first approximation for a spindle-shaped cell crawling through a 3-D matrix. (The cell height, which is

irrelevant in the present study, is taken unrealistically large to improve numerical convergence in the 1-D model.)

The 2-D model represents an EK lamellipodium in the X_1-X_2 plane. The stationary lamellipodium is taken as a circular annulus of outer radius 10 μm surrounding a cell body of radius 3 μm (Figure 1b). For a crawling cell, the lamellipodium initially has the shape of a section of a circle with rounded corners and approximately the same planform area as the lamellipodium in the stationary model (Figure 1c); the cell body is not included. For convenience, the center of the circle defining the leading edge at $t = 0$ is placed at the origin of the X_i system.

Actually, both models are analyzed as two-dimensional objects, just in different planes, with the 1-D and 2-D models assumed to be under conditions of plane strain and plane stress, respectively. The cell membrane is not included explicitly, but the boundaries are assumed to be impermeable to flow. Hence, the total cell volume remains constant, although fluid can flow from one region to another.

The strain-energy density function for the solid skeleton must have a compressible form. For a first approximation, we assume that the CSK is transversely isotropic and take

$$W = C\left[I^* - 3 + \frac{1 - 2\nu}{\nu}\left(J^{*-2\nu/(1-2\nu)} - 1\right)\right] + C_3 E_{33}^{*2}, \qquad (4\text{-}1)$$

where the term in brackets corresponds to a Blatz–Ko material [1962] and the last term provides a stiffening in the X_3-direction that prevents the lamellipodium from thickening unrealistically during crawling. We speculate that relatively short and stiff crosslinks between actin filaments, as well as connections between actin and the cell membrane, provide this additional transverse stiffness. The parameters C and ν represent the small-strain modulus and Poisson's ratio, respectively, of the drained solid skeleton, C_3 is the transverse modulus, and $I^* = \text{tr}\,(I + 2E^*)$ is a strain invariant.

Actin polymerization. Polymerization of actin filaments is simulated by growth of the solid skeleton in a narrow zone near the leading edge of the model, along the local fiber coordinate (X_f) normal to the edge in the reference configuration. In the 1-D model, X_f is the longitudinal coordinate X_1, whereas in the 2-D models, it is the radial direction R (Figure 1b,c). Since growth in our models is one-dimensional, the growth tensor G has only one non-zero component, denoted here simply by G. With the initial value of G set to unity everywhere, $G > 1$ gives polymerization and $G < 1$ gives contraction (see below).

It is important to note that simulating actin polymerization by volumetric growth of filaments is a relatively crude approximation for the actual process. At the molecular level, the addition of monomers to the end of a filament actually is a type of surface growth. To more accurately model polymerization in a continuum formulation, a multiscale analysis is needed. This would be a natural extension of the present model.

Feedback control of polymerization in the models is based on the work of Peskin et al. [1993], who postulate that, at the leading edge of a crawling cell, actin filaments push the cell membrane forward via a Brownian ratchet mechanism. According to this idea, thermal fluctuations in the membrane create a small gap between the membrane and actin filaments, allowing actin monomers to enter and add to the length of the filaments. Here, we propose that positive fluid pressure at the leading edge facilitates this process by pushing the cell membrane forward, enlarging the gap and increasing the rate of polymerization. This

mechanism is implemented through the feedback law

$$\ddot{G}_p = A_p \, \dot{p}_L \, g_p(x_f), \tag{4-2}$$

where G_p represents the growth due to polymerization, p_L is the fluid pressure at the leading edge, and A_p is a positive rate constant. With this law taken as second order in time, the polymerization rate ($\dot{G}_p$) becomes constant when pressure reaches a steady state ($\dot{p}_L = 0$).

In the equation above, the function g_p provides the spatial distribution of growth in the current configuration. With x_L denoting the current location of the leading edge, we take

$$g_p(x_f) = H_s(x_f - (x_L - \Delta x_p, \Delta x_{jp}))(1 - \beta X_2^2), \tag{4-3}$$

in which H_s is a smoothed Heaviside step function defined in the region $x_f \geq x_L - \Delta x_p$. In this expression, x_f is the local fiber coordinate (the growth direction), Δx_p defines the (specified) width of the growth zone, and Δx_{jp} is the length of the smoothed "jump" from zero to one. This form for g_p keeps the width of the active polymerization zone fixed in the crawling direction, rather than growing longer with G. In addition, the term with the specified constant β allows for a gradient in polymerization force along the leading edge of the 2-D crawling model ($g_c = 0$ for $|X_2| > 1/\sqrt{\beta}$ and $\beta = 0$ for the 1-D and 2-D stationary models). Some investigators have speculated that the observed regional variation in actin density, being highest near the center and lowest near the sides of the cell, plays an important role in maintaining cell shape [Keren et al. 2008; Rubinstein et al. 2009].

Cytoskeletal contraction. Actomyosin contraction occurs in a wider zone toward the back of the cell [Yam et al. 2007]. To achieve steady-state conditions, we speculate that the contraction is continuous in time, i.e., the contracting region never relaxes. For this to occur, contraction is assumed to be accompanied by CSK remodeling, similar to the contractile mechanism of smooth muscle [An and Fredberg 2007; Bossé et al. 2008]. As a smooth muscle cell contracts, its CSK remodels to reset the zero-stress configuration so that shortening can continue to very large magnitudes. This process is included in the present models by progressively enlarging the contractile region so it maintains a constant width as it contracts. In the 1-D and stationary 2-D models, contraction is assumed to occur in the same direction as actin polymerization, i.e., normal to the leading edge. In the 2-D crawling model, contraction occurs in the direction of motion.

To obtain steady-state motion, feedback is included in the contractile machinery. A growing amount of evidence suggests that cells and tissues respond to mechanical perturbations in ways that tend to restore a homeostatic state of stress [Humphrey 2008]. Correspondingly, we assume that a homeostatic (target) stress $\hat{\sigma}_0$ exists for contractile fibers and propose a contraction feedback law of the form

$$\ddot{G}_c = -A_c^2 (1 - \hat{\sigma}_s / \hat{\sigma}_0) g_c(x_f), \tag{4-4}$$

where G_c is the (negative) growth due to contraction, $\hat{\sigma}_s$ is the true solid stress in the fiber direction, and A_c is a rate constant. With this law, the contractile rate ($\dot{G}_c$) becomes constant when $\hat{\sigma}_s = \hat{\sigma}_0$.

The spatial distribution for contraction is taken in a form similar to that given above for g_p, i.e.,

$$g_c = [1 - H_s(x_f - (x_0 + \Delta x_c, \Delta x_{jc}))](1 - \beta X_2^2), \tag{4-5}$$

where x_0 is the current location of the trailing edge of the lamellipodium, and the lengths Δx_c and Δx_{jc} provide an active contractile zone of constant geometry in the current configuration. In the axisymmetric model for a stationary cell, we take $\beta = 0$, as well as $g_c = 0$ in the cell body (Figure 1b).

Once G_p and G_c are computed, the total growth is given by $G = G_p G_c$.

Adhesion dynamics. Adhesion between the CSK and the substrate is modeled by a frictional body force [Alt and Dembo 1999; Gracheva and Othmer 2004; Kruse et al. 2006; Kuusela and Alt 2009]. The proportionality between velocity and force represents the cumulative effects of a number of adhesions. The force on each adhesion depends on the displacement of the cytoskeleton relative to the undeformed substrate, over the time since the adhesion was formed. If adhesion lifetime is approximately constant, net adhesion force will increase roughly in proportion to velocity, and the coefficient in this relationship will increase with density of adhesions, adhesion stiffness, and substrate stiffness. Other dissipative forces, such as those caused by viscous resistance between the CSK and cell membrane, as well as between the cell membrane and the surrounding environment, are included through an additional friction force. The total friction force is given by

$$\boldsymbol{f}_0 = -\eta \boldsymbol{v}_s - \eta_c \boldsymbol{v}_c, \tag{4-6}$$

where η and η_c are friction coefficients for the adhesions and cell membrane interaction, respectively. In addition, $\boldsymbol{v}_s$ is the local velocity vector for the solid skeleton (CSK), and $\boldsymbol{v}_c$ is the forward velocity of the cell as a whole, taken as the velocity at the center of the leading edge.

5. Computational methods

To express the governing equations in dimensionless form, the following nondimensional variables are defined:

$$\bar{X} = \frac{X}{L}, \quad \bar{x} = \frac{x}{L}, \quad \bar{\nabla} = L\nabla, \quad \bar{t} = \frac{t}{T}, \quad \bar{U} = \frac{U}{L}, \quad \bar{v} = \frac{vT}{L},$$

$$\bar{\sigma} = \frac{\sigma}{C}, \quad \bar{P} = \frac{P}{C}, \quad \bar{f}_0 = \frac{f_0 L}{C}, \quad \bar{p} = \frac{p}{C}, \quad \bar{W} = \frac{W}{C}, \quad \bar{C}_3 = \frac{C_3}{C},$$

$$\bar{K} = \frac{CT}{L^2} K, \quad \bar{\eta} = \frac{L^2}{CT} \eta, \quad \bar{A}_p = CTA_p, \quad \bar{A}_c = TA_c.$$

The characteristic length L is taken as the initial cell length ($10 \, \mu$m), and the characteristic time is defined as $T = L/v_c = 50$ s for a cell velocity $v_c = 0.2 \, \mu$m/s. Thus, it typically takes an EK about 50 s to move a distance equal to its own length. In terms of these quantities, the governing equations (3-1), (3-3), (3-5), (3-6), (3-8), (4-2), (4-4), and (4-6) maintain their dimensional forms if we let dot denote $d/d\bar{t}$ and remove all bars. Unless stated otherwise, all results are given in terms of nondimensional variables with the bars implied.

Finite element modeling. Finite element models were developed using the commercial software Comsol Multiphysics (v. 3.5; Comsol AB, Burlington, MA). The Structural Mechanics Module provided the basis for the nonlinear elastic part of the model. The governing equations were modified to implement volumetric growth as described in [Taber 2008], with Darcy's law and the feedback laws included as auxiliary differential equations. In addition, the following changes were made to enforce incompressibility:

- The extra term $-p\,\delta$ was added to the strain-energy density function, where δ is the expression on the left side of (3-6) and p is a Lagrange multiplier with the physical meaning of fluid pressure.

- The expression $(p/M + \delta)\,p_t$ was included as a weak constraint condition, where p_t is a test function and M is a (large) penalty parameter (set to 10^5).

Triangular second-order and first-order elements were used for displacement (solid elasticity) and pressure (fluid flow), respectively. The mesh was refined until the model yielded sufficient solution accuracy. The function *flc2hs* provided the smoothed Heaviside function used in the feedback laws.

Finally, we note that the model is unstable when it is required to meet the target stress at all points in the CSK. To improve stability, we used the average value of the true solid stress over the rear $4\,\mu$m (1-D and 2-D crawling models) or $3\,\mu$m (2-D stationary model) along the cell centerline for $\hat{\sigma}_s$ in (4-4). Testing showed that the length of the averaging region has a relatively small impact on the results. Similarly, the value of p_L in (4-2) was taken at the middle of the leading edge.

Simulation procedure. Initially, the model is at rest with no stresses and $\boldsymbol{G} = \boldsymbol{I}$. All boundaries are stress free and impermeable to flow ($\boldsymbol{P} \cdot \boldsymbol{N} = \dot{\boldsymbol{U}} \cdot \boldsymbol{N} = 0$, where $\boldsymbol{N}$ is the local normal to the boundary in the reference configuration). There are no specified loads or displacements. At the outset of the simulation, the feedback law (4-4) initiates a contraction at the rear of the cell, generating fluid pressure, and the other feedback law (4-2) causes growth near the front that pushes the leading edge forward. Except for time plots and results from the 2-D crawling model, all results are plotted after the cell reaches steady-state conditions.

Parameter values. The geometries of the polymerization and contraction zones were estimated from published distributions of F-actin, myosin II, and contractile strains in crawling EKs [Svitkina et al. 1997; Adachi et al. 2009]. The values used in the models are $\Delta x_p = \Delta x_{jp} = 1\,\mu$m, $\Delta x_c = 4\,\mu$m (1-D & 2-D crawling) or $2\,\mu$m (2-D stationary), $\Delta x_{jc} = 4\,\mu$m (1-D & 2-D crawling) or $3\,\mu$m (2-D stationary), and $\beta = 0.25$.

Otherwise, with $\boldsymbol{K} = K\boldsymbol{I}$ (isotropic permeability), the models contain ten physical parameters: C, C_3, ν, K, ϕ_s ($\phi_f = 1 - \phi_s$), η, η_c, A_p, A_c, and $\hat{\sigma}_0$. Unfortunately the range of reported parameter values is quite large in some cases. For example, estimates for the hydraulic permeability (K) range over at least five orders of magnitude, from about 10^{-10} to 10^{-15} m^4/N s [Guilak et al. 2006]. Hence, after first obtaining first-approximation parameter values from published data as described below, we iteratively correlated model and experimental results to refine these values in order to define a baseline model (Table 1). We also used this method to guide our choices for the functions g_p and g_c in (4-2) and (4-4).

In defining the mechanical properties of the cytoskeleton, we note that recent studies have shown that the CSK is capable of very rapid remodeling [An and Fredberg 2007; Bossé et al. 2008]. Hence, we assume that CSK remodeling keeps up with the dynamics of the crawling cell and take the porosity and permeability as constant. The representative values $\phi_s = \phi_f = 0.5$ are used for the volume fractions, and the value taken for K is near the low end of the range given above (Table 1). The values chosen for C and ν are typical for a compressible cell [Guilak et al. 2006], while C_3 is taken large enough to prevent significant changes in lamellipodium thickness.

Parameter	Physical meaning	Value	Nondimensional Value
C	CSK modulus	1 kPa	1
C_3	Transverse fiber modulus	20 kPa	20
ν	Poisson's ratio	0.3	0.3
K	Effective hydraulic permeability	4×10^{-15} m^4/N s	2
η	Adhesion friction coefficient	200 Pa s/μm^2	0.4
η_c	Cell membrane friction coefficient	50 Pa s/μm^2	0.1
A_p	Actin polymerization rate constant	1 (kPa s)$^{-1}$	50
A_c	CSK contraction rate constant	0.2 s^{-1}	10
$\hat{\sigma}_0$	Target stress	0.5 kPa	0.5

Table 1. Parameter values for baseline model.

The value used for the friction coefficient (η) was measured by [Fournier et al. 2010]. The membrane friction coefficient (η_c) is assumed to be similar to η, but smaller to keep the cell moving forward.

The rate constants (A_p and A_c) and the target stress ($\hat{\sigma}_0$) were estimated as follows. The value of $\hat{\sigma}_0$ is based on measurements of actomyosin contractile stress (~ 0.4 kPa) in EKs [Galbraith and Sheetz 1999]. The constant A_c in (4-4) characterizes the rate of contraction under zero load conditions. With the unloaded speed of myosin relative to actin being approximately $v_m = 6\,\mu$m/s [Howard 2001], the fractional contraction rate is roughly $A_c = v_m/\Delta x_c \cong 1.5\,\text{s}^{-1}$, but we ended up using a somewhat smaller value (Table 1).

Finally, to estimate the actin polymerization rate constant A_p, we use two pieces of information: (1) the unloaded polymerization velocity for actin is approximately $v_0 = 0.8\,\mu$m/s [Peskin et al. 1993]; and (2) the stall stress for actin polymerization is about $\sigma_s = 2$ kPa in EKs [Prass et al. 2006]. Then, taking the force-velocity relation as linear (a very rough approximation) yields $A_p = (v_0/\Delta x_p)/\sigma_s = 0.4$ (kPa s)$^{-1}$.

Again, all of these values were modified somewhat to obtain more accurate model behavior, and, unless stated otherwise, the results for both the 1-D and 2-D models are based on the parameters given in Table 1. Because of incomplete and inconsistent data, it was important to conduct a parametric sensitivity analysis (see Results section).

6. Results

One-dimensional model. Steady-state distributions along the cell are shown for actin velocity ($v_s = \dot{x}_1$), CSK stretch rate ($\dot{F}_{11}$), actin stress ($\hat{\sigma}_s$), fluid pressure (p), and relative fluid velocity (v_f) (Figure 2). For clarity, except in the plots of the deformed cell (Figure 2a), all quantities are plotted as functions of the undeformed coordinate X_1. To illustrate the effects of fluid flow, curves are shown for various values of the effective hydraulic permeability K.

Actin velocity is forward at the front and back of the cell, but backward elsewhere (Figure 2b). The retrograde flow of actin is caused by growing actin filaments at the front pushing backwards as they meet frictional resistance to forward motion (due to adhesions). The magnitude of the actin velocity increases with K in each region.
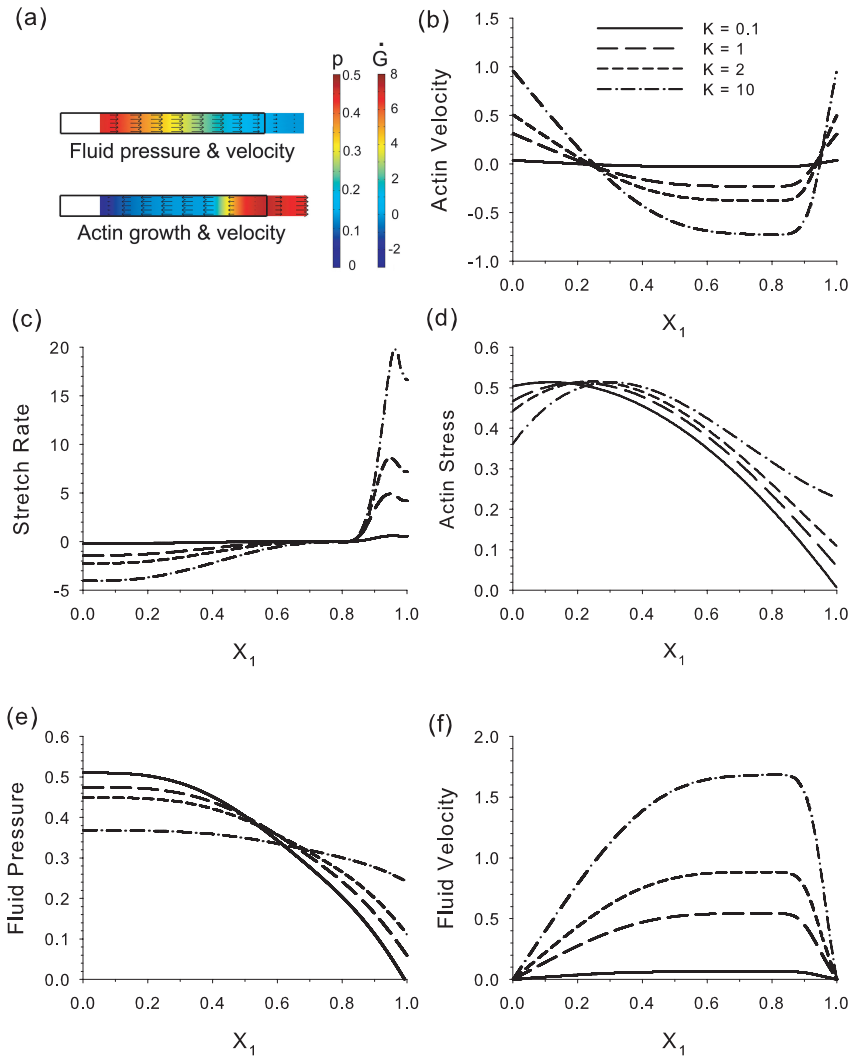
Figure 2. Steady-state distributions in 1-D model. Except in (a), quantities are plotted as functions of the undeformed longitudinal coordinate. Results are shown for various values of the effective hydraulic permeability K, with $K = 2$ giving the baseline case and the legend in (b) applying to all plots. (a) Deformed cell geometry (outline is initial geometry). Top: pressure (colors) and relative fluid velocity (arrows); bottom: growth rate (colors) and actin velocity (arrows). (b) Actin velocity. (c) Stretch rate. (d) Actin (solid) stress. (e) Fluid pressure. (f) Relative fluid velocity.

Actin polymerization causes a positive stretch rate near the leading edge, while CSK contraction generates increasingly negative stretch rates toward the rear (Figure 2c). The stretch-rate magnitudes increase with K. In addition, the actin (solid) stress distribution peaks near the rear of the cell and drops off toward the front (Figure 2d).
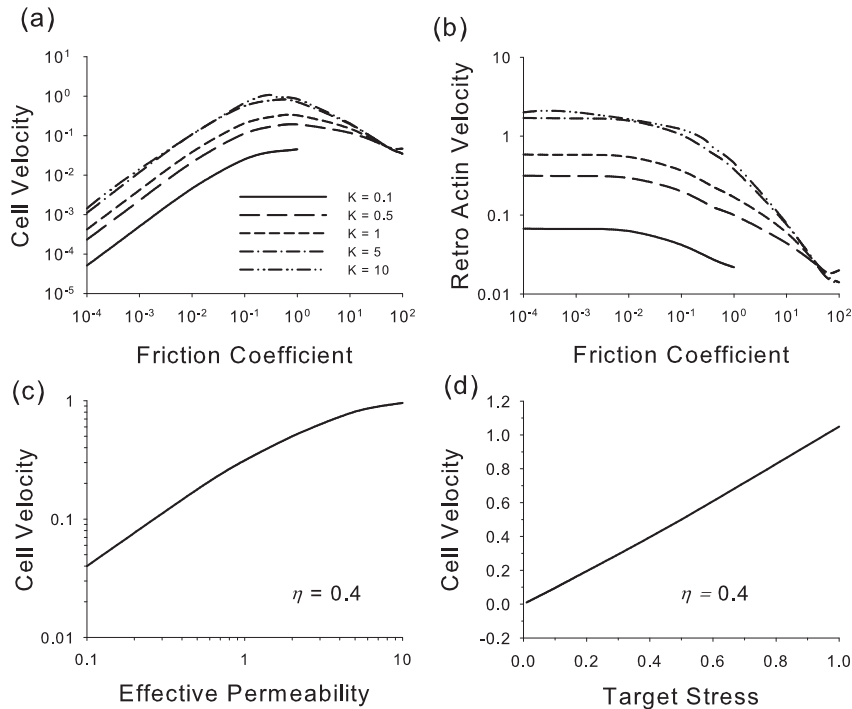
Figure 3. Effects of adhesiveness (friction) and target stress on cell velocity and retrograde actin flow in 1-D model. (a) Cell velocity vs friction coefficient η for various values of effective hydraulic permeability K. (b) Retrograde actin velocity vs η for various K (same legend as in (a)). (c) Cell velocity vs K for $\eta = 0.4$. (d) Cell velocity vs target stress $\hat{\sigma}_0$ for $\eta = 0.4$.

The fluid pressure is positive throughout the cell (Figure 2e). It is greatest at the trailing edge and decreases to a minimum at the leading edge. The negative pressure gradient, which is produced by the combined effects of contraction at the rear and growth at the front, drives the fluid forward. The fluid velocity relative to the solid skeleton peaks near the leading edge and meets the boundary conditions of zero relative flow at the front and back (Figure 2f). Although the pressure gradient decreases with increasing values of K (Figure 2e), the fluid velocity increases because the higher permeability offers less resistance to flow.

Cell speed (actin speed at the leading/trailing edge) depends in a biphasic manner on adhesiveness, as represented by the friction coefficient η (Figure 3a). It is small at low and high values of η and reaches a peak between $\eta = 0.1$ and 1. In contrast, the retrograde actin velocity is nearly constant for $\eta < 0.01$ but drops off for larger values of η (Figure 3b). For the baseline value $\eta = 0.4$, cell speed increases with K and approaches a constant value for $K > 10$ (Figure 3c). However, the speed increases in proportion to the target stress $\hat{\sigma}_0$ over the range of values studied (Figure 3d).

Finally, results are presented to illustrate the temporal behavior of the model. At the beginning of the simulation, the start-up conditions produce oscillations which eventually damp out in most cases, leaving
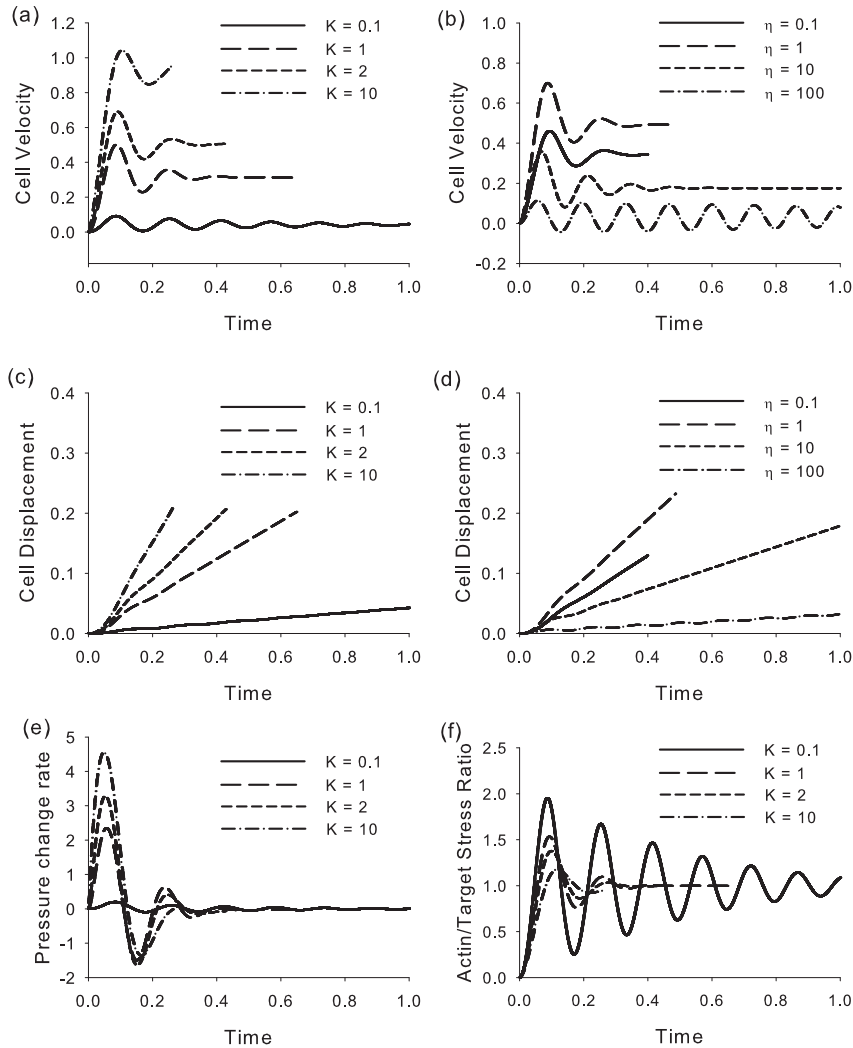
Figure 4. Time-dependent behavior of 1-D model. (a) Cell velocity for various values of the effective hydraulic permeability K. (b) Cell velocity for various values of the friction coefficient η. (c) Cell displacement for various K. (d) Cell displacement for various η. (e) Rate of change in fluid pressure for various K. (f) Ratio of actin (solid) stress to target stress ($\hat{\sigma}_s/\hat{\sigma}_0$) for various K.

a steady-state condition (Figure 4a,b). At steady state, the fluid pressure becomes constant and $\hat{\sigma}_s \to \hat{\sigma}_0$ (Figure 4e,f).

For certain parameter values, however, the oscillations persist and may continue indefinitely. We found that this behavior is especially sensitive to the values of K and η, as persistent oscillations appear for low values of K and high values of η (Figure 4a,b). Even when cell speed oscillates, however, the net motion is forward in all cases (Figure 4c,d).
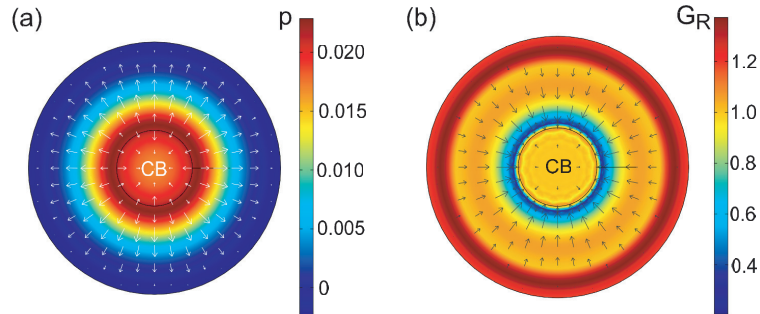
Figure 5. Results for stationary 2-D model ($t = 0.45$; CB = cell body). (a) Pressure (colors) and relative fluid velocity (arrows). (b) Radial growth (colors) and actin velocity (arrows).

Two-dimensional model. Unfortunately, the present 2-D models do not reach steady state. (A parameter study failed to find a suitable combination of parameter values.) Hence, we present only illustrative results at times when nearly steady-state conditions occur.

Although the fluid pressure in the stationary model is relatively small, the pressure gradient is significant enough to drive fluid radially outward from the cell body everywhere except near the leading edge and within the cell body (Figure 5a). At the same time, the actin CSK flows inward throughout the lamellipodium except for a small outward motion due to polymerization near the outer edge (Figure 5b).

The crawling model acquires the characteristic fan-like shape of an EK (Figure 6). Near the symmetry axis of the cell, the results are similar to those from the 1-D model. In particular, a pressure gradient drives fluid toward the leading edge (Figure 6b), while actin flows backward in the middle of the cell and forward near the front and back edges (Figure 6c). Near the tips of the "wing," however, relatively little fluid flow occurs, but the actin moves radially inward at considerable speed. The continued lateral deformation is one reason for the failure to achieve steady state. Also shown is a map of the frictional body force exerted by adhesions on the cell (Figure 6d). Peaks in force occur near the front and back edges, with strong concentrations at the wing tips.

7. Discussion

Theoretical modeling has played a central role in studies of cell migration. Models promote understanding of the physical processes underlying observed behavior. They also can be used to test whether a given hypothetical mechanism is consistent with physical law.

Previous models for cell crawling. Here, we provide a brief summary of some of the models most pertinent to the present work, focusing primarily on 1-D and 2-D models with a significant mechanical component. More comprehensive recent reviews of this topic include those of [Flaherty et al. 2007; Carlsson and Sept 2008].

DiMilla et al. [1991] proposed one of the first models for cell crawling that includes adhesion dynamics. The CSK is modeled as a 1-D viscoelastic system of springs and dashpots, with contractile elements providing the driving force, which mediates binding affinities. Notably, this model predicts the experimentally observed biphasic behavior of cell speed with increasing substrate adhesiveness. In particular,
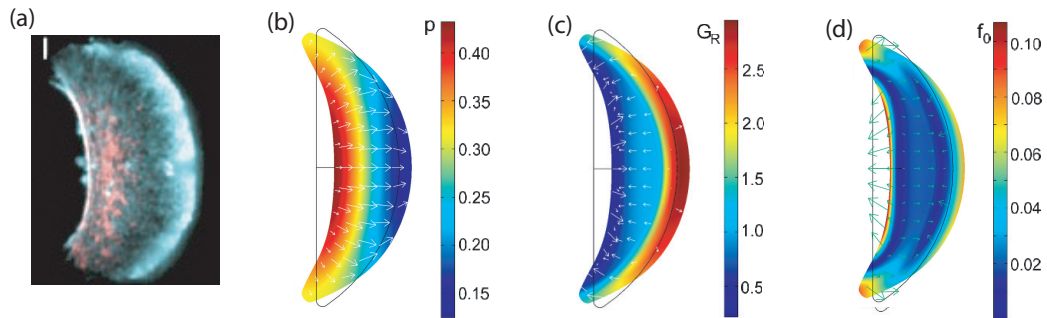
Figure 6. Results for 2-D model crawling toward the right ($t = 0.3$; outlines indicate initial configuration). (a) Image of isolated EK lamellipodium crawling on a surface (red: myosin II; cyan: actin) (reprinted with permission from [Verkhovsky et al. 1999]). (b) Pressure (colors) and relative fluid velocity (arrows). (c) Radial growth (colors) and actin velocity (arrows). (d) Frictional body force (colors and arrows) exerted by substrate on the cell.

cell speed is low for either very high or very low adhesiveness, and it peaks at some intermediate value [Palecek et al. 1997]. A common interpretation for this behavior is that cells cannot gain enough traction to move quickly on very slippery surfaces, while they have difficulty letting go of very sticky surfaces [DiMilla et al. 1991]. The present model captures this fundamental behavior (Figure 3a).

In [Alt and Dembo 1999; Kuusela and Alt 2009] the cytoplasm is treated as a biphasic fluid-like mixture. One phase contains actin and myosin filaments, and the other phase is a solvent containing actin monomers. The phases exchange monomers through chemical reactions, and the forces of actin polymerization and CSK contraction are simulated by positive and negative pressures, respectively. Other investigators also have considered the cytoplasm as a fluid-like mixture (e.g., [Kruse et al. 2006; Rubinstein et al. 2009].

Modeling the CSK as a fluid is based on the argument that the time scales for CSK remodeling and function are similar [Alt and Dembo 1999]. At low Reynolds numbers like those in a cell, however, fluid stress is dominated by pressure, which is intrinsically isotropic, whereas solid-like behavior is needed to produce directed (anisotropic) stress like that generated by actin polymerization and CSK contraction. Our model includes rapid remodeling of contractile fibers, but they are considered as elastic solids.

Hydrostatic pressure provides the driving force in some published models for pseudopod extension and blebbing [Young and Mitran 2010; Oster and Perelson 1987]. In our model, pressure plays a central but more indirect role, as it modulates the rate of actin polymerization at the leading edge, but actin filaments do the actual "pushing." The pseudopod extension model of [Zhu and Skalak 1988] is similar in some respects to our model. In their model, actin polymerization pushes the cell membrane forward, creating a pressure drop that draws fluid and actin monomers toward the leading edge. As in the present model, fluid flow is governed by Darcy's law.

Stolarska et al. [2009] presented a 3-D finite element model for an EK crawling on a surface. The cell is treated as a viscoelastic solid, and focal adhesions are included at the boundary with the substrate. As in our model, actin polymerization and CSK contraction are simulated using volumetric growth. However, growth is specified in their model, rather than being governed by feedback.

The 1-D model of [Larripa and Mogilner 2006] treats the cell as a viscoelastic actin-myosin gel. Contractile stress is assumed to be proportional to the concentration of myosin bound to actin, as computed from chemical kinetics. This model achieves steady-state motion like an EK.

Perhaps the most biochemically comprehensive model for a crawling cell is the 2-D model of [Rubinstein et al. 2005] for an EK lamellipodium. This model includes turnover of actin and actomyosin fibers, forces generated by actin polymerization and actomyosin contraction, and actin transport. External loads and displacements are applied at the boundaries in an ad hoc manner, whereas all loads in our models arise automatically from mechanical feedback.

Comparison of numerical and experimental results. A theoretical model is only as good as the experimental data used to support and test it. Over the years, numerous experimental studies of cell crawling have been published. Many of these studies have been qualitative, but recent reports have emphasized quantitative measurements. We used some of these data to estimate parameter values for our models, but, as noted above, the available data are not always consistent. In addition to natural biological variability, the use of different cell and substrate types contribute to this inconsistency. Nevertheless, it is important to compare numerical and experimental results when possible.

In general, results given by previous models for cell crawling have been compared with only limited portions of the available experimental data. Here, we evaluate our models using the following data for EKs under steady-state conditions:

(1) The speed and shape of a crawling EK are relatively constant. At steady state, our 1-D model is consistent with this observation. The steady-state cell speed in the 1-D baseline model is approximately 0.1 μm/s, which is within the range of reported values (about 0.05–0.2 μm/s; see, e.g., [Adachi et al. 2009; Okeyo et al. 2009; Wilson et al. 2010]. The speed is similar in the 2-D model, but this model does not achieve steady state.

(2) As discussed above, cell speed depends on adhesion strength in a biphasic manner. (Actually, this behavior has been found in at least two cell types [Palecek et al. 1997; Peyton and Putnam 2005]; to our knowledge this effect has not been explored in EKs.) Our 1-D model captures this behavior, including the shapes of the cell velocity vs adhesiveness (η) curves and the result that cell velocity increases by approximately an order of magnitude for a similar increase in adhesiveness (Figure 3a).

(3) In crawling cells, the actin CSK moves forward relative to the substrate at the front and back of the lamellipodium and backwards in between [Schaub et al. 2007; Yam et al. 2007; Fournier et al. 2010]. In stationary (circular) EKs, on the other hand, actin flows radially inward toward the cell body throughout the lamellipodium [Yam et al. 2007]. Both the 1-D and 2-D models are consistent with these results (Figures 2a,b and 5b). In the 1-D baseline model, the retrograde actin speed is about 0.07 μm/s, which is similar to measured values for EKs [Schaub et al. 2007; Adachi et al. 2009].

(4) In the direction of motion, strain rate along the midline of the lamellipodium is positive near the leading edge and negative toward the rear [Adachi et al. 2009]. Results from our models are consistent with this pattern (Figure 2c). In addition, the magnitude of the peak negative strain rate in the 1-D baseline model is about $-0.04\ \mathrm{s}^{-1}$, which is nearly the same as that measured in EKs by Adachi et al.

(5) Relative to the CSK, intracellular fluid flows forward through the lamellipodium [Keren et al. 2009]. Both the 1-D and 2-D models give this result (Figures 2f and 6b). However, the fluid and solid velocities in our models are similar, whereas Keren at al. found that relative fluid velocity is about 40% that of the cell.

(6) The largest traction stresses between a crawling EK and the substrate occur along the leading and trailing edges of the cell, as well as near the tips of the "wings" [Jurado et al. 2005; Fournier et al. 2010]. Our 2-D model is consistent with these data, including the result that the maximum traction occurs at the wing tips (Figure 6d), but the measured traction direction at the tips is opposite to that predicted by the model.

(7) According to our 1-D model, cell velocity increases with the value of the target stress ($\hat{\sigma}_0$) in the contractile fibers (Figure 3d). This prediction is consistent with the experiments of Okeyo et al. [2009], who found that EKs move faster when exposed to the contraction enhancer calyculin and slower when exposed to the contraction inhibitor blebbistatin.

In addition to these key results, Barnhart et al. [2010] recently have found that crawling EKs exhibit oscillations under certain conditions. They speculate that these oscillations are caused by dynamic interactions between cell elasticity and cell-substrate adhesions. Moreover, most other types of cells, e.g., fibroblasts, crawl in a cyclic manner. Interestingly, our 1-D model oscillates for relatively small values of the hydraulic permeability or large values of the friction coefficient (Figure 4a,b). These results suggest that the oscillations are caused by delays in the feedback loops in our models, i.e., phase differences that often characterize poroelastic and frictional systems. This type of behavior is commonly encountered in feedback control systems.

Limitations. This study focuses solely on mechanical aspects of cell crawling. Biochemistry and detailed molecular mechanisms, e.g., adhesion dynamics, receptor-ligand binding, and stress fiber assembly, are not considered, although they clearly play major roles in this problem [DiMilla et al. 1991; Boulbitch et al. 2001; Schaub et al. 2007; Yam et al. 2007; Vicente-Manzanares et al. 2009]. We also do not consider some of the more detailed mechanical events, such as lamellipodial ruffling [Zhang 2009], the effects of substrate stiffness on cell speed and direction [Pelham and Wang 1997; Lo et al. 2000; Ghosh et al. 2007], and how crawling cells probe their environment [Mattila and Lappalainen 2008; Petrie et al. 2009]. These effects are beyond the scope of the present study.

Concluding remarks. Our models reproduce quite well a relatively wide range of experimental data obtained for crawling EKs. There are some discrepancies, however. A crucial unresolved issue is how the cell achieves steady state motion in two dimensions. A number of investigators have pondered this problem, with some suggesting that a gradient in actin polymerization along the leading edge is a necessary ingredient [Lee et al. 1993; Grimm et al. 2003; Rubinstein et al. 2009]. Although our 2-D crawling model includes such a gradient, it clearly is still missing something essential. One possibility is that the stress fiber along the trailing edge and the strong adhesions at the ends of the fiber are important factors. This problem remains fertile ground for future work.

A personal note

This paper is dedicated to Marie-Louise and Charles Steele. The first author (Taber) has had the privilege of knowing them since his days as a graduate student. MarieLu treated JoMMS as her "baby" and was instrumental in helping Charles establish and maintain its high standards. She was loved by all and is greatly missed. Taber learned the art of mathematical modeling from Professor Steele, who was his doctoral research advisor. Although the models presented in this paper cannot rival the extraordinarily creative and mathematically rigorous models of Professor Steele, we have striven to make them consistent with his teachings. That is, our models are based on fundamental physical principles, include (we hope) the most essential elements, and have been tested in multiple ways using experimental data. We wish him a pleasant retirement.

Acknowledgements

References

[Adachi et al. 2009] T. Adachi, K. O. Okeyo, Y. Shitagawa, and M. Hojo, "Strain field in actin filament network in lamellipodia of migrating cells: implication for network reorganization", *J. Biomech.* **42**:3 (2009), 297–302.

[Alt and Dembo 1999] W. Alt and M. Dembo, "Cytoplasm dynamics and cell motion: two-phase flow models", *Math. Biosci.* **156**:1–2 (1999), 207–228.

[An and Fredberg 2007] S. S. An and J. J. Fredberg, "Biophysical basis for airway hyperresponsiveness", *Can. J. Physiol. Pharmacol.* **85**:7 (2007), 700–714.

[Barnhart et al. 2010] E. L. Barnhart, G. M. Allen, F. Jülicher, and J. A. Theriot, "Bipedal locomotion in crawling cells", *Biophys. J.* **98**:6 (2010), 933–942.

[Biot 1972] M. Biot, "Theory of finite deformations of porous solids", *Indiana Univ. Math. J.* **21**:7 (1972), 597–620.

[Blatz and Ko 1962] P. J. Blatz and W. L. Ko, "Application of finite elastic theory to the deformation of rubbery materials", *Trans. Soc. Rheol.* **6**:1 (1962), 223–251.

[Bossé et al. 2008] Y. Bossé, A. Sobieszek, P. D. Paré, and C. Y. Seow, "Length adaptation of airway smooth muscle", *Proc. Am. Thorac. Soc.* **5**:1 (2008), 62–67.

[Boulbitch et al. 2001] A. Boulbitch, Z. Guttenberg, and E. Sackmann, "Kinetics of membrane adhesion mediated by ligand-receptor interaction studied with a biomimetic system", *Biophys. J.* **81**:5 (2001), 2743–2751.

[Bowen 1976] R. M. Bowen, "Theory of mixtures", pp. 1–127 in *Continuum physics, III: Mixtures and EM field theories*, edited by A. C. Eringen, Academic Press, New York, 1976.

[Bray 2001] D. Bray, *Cell movements: from molecules to motility*, Garland, New York, 2001.

[Carlsson and Sept 2008] A. E. Carlsson and D. Sept, "Mathematical modeling of cell migration", *Methods Cell Biol.* **84** (2008), 911–937.

[Charras et al. 2005] G. T. Charras, J. C. Yarrow, M. A. Horton, L. Mahadevan, and T. J. Mitchison, "Non-equilibration of hydrostatic pressure in blebbing cells", *Nature* **435**:7040 (2005), 365–369.

[DiMilla et al. 1991] P. A. DiMilla, K. Barbee, and D. A. Lauffenburger, "Mathematical model for the effects of adhesion and mechanics on cell migration speed", *Biophys. J.* **60**:1 (1991), 15–37.

[Flaherty et al. 2007] B. Flaherty, J. P. McGarry, and P. E. McHugh, "Mathematical models of cell motility", *Cell Biochem. Biophys.* **49**:1 (2007), 14–28.

[Fournier et al. 2010] M. F. Fournier, R. Sauser, D. Ambrosi, J.-J. Meister, and A. B. Verkhovsky, "Force transmission in migrating cells", *J. Cell Biol.* **188**:2 (2010), 287–297.

[Friedl and Bröcker 2000] P. Friedl and E.-B. Bröcker, "The biology of cell locomotion within three-dimensional extracellular matrix", *Cell Mol. Life Sci.* **57**:1 (2000), 41–64.

[Galbraith and Sheetz 1999] C. G. Galbraith and M. P. Sheetz, "Keratocytes pull with similar forces on their dorsal and ventral surfaces", *J. Cell Biol.* **147**:6 (1999), 1313–1324.

[Ghosh et al. 2007] K. Ghosh, Z. Pan, E. Guan, S. R. Ge, Y. J. Liu, T. Nakamura, X.-D. Ren, M. Rafailovich, and R. A. F. Clark, "Cell adaptation to a physiologically relevant ECM mimic with different viscoelastic properties", *Biomater.* **28**:4 (2007), 671–679.

[Gracheva and Othmer 2004] M. E. Gracheva and H. G. Othmer, "A continuum model of motility in ameboid cells", *Bull. Math. Biol.* **66**:1 (2004), 167–193.

[Grimm et al. 2003] H. P. Grimm, A. B. Verkhovsky, A. Mogilner, and J.-J. Meister, "Analysis of actin dynamics at the leading edge of crawling cells: implications for the shape of keratocyte lamellipodia", *Eur. Biophys. J.* **32**:6 (2003), 563–577.

[Guilak et al. 2006] F. Guilak, M. A. Haider, L. A. Setton, T. A. Laursen, and F. P. T. Baaijens, "Multiphasic models of cell mechanics", pp. 84–102 in *Cytoskeletal mechanics: models and measurements*, edited by M. R. K. Mofrad and R. D. Kamm, Cambridge University Press, New York, 2006.

[Howard 2001] J. Howard, *Mechanics of motor proteins and the cytoskeleton*, Sinauer, Sunderland, MA, 2001.

[Humphrey 2008] J. D. Humphrey, "Vascular adaptation and mechanical homeostasis at tissue, cellular, and sub-cellular levels", *Cell Biochem. Biophys.* **50**:2 (2008), 53–78.

[Iwasaki and Wang 2008] T. Iwasaki and Y.-L. Wang, "Cytoplasmic force gradient in migrating adhesive cells", *Biophys. J.* **94**:5 (2008), L35–L37.

[Jurado et al. 2005] C. Jurado, J. R. Haserick, and J. Lee, "Slipping or gripping? Fluorescent speckle microscopy in fish keratocytes reveals two different mechanisms for generating a retrograde flow of actin", *Mol. Biol. Cell* **16**:2 (2005), 507–518.

[Kenyon 1976] D. E. Kenyon, "The theory of an incompressible solid-fluid mixture", *Arch. Ration. Mech. Anal.* **62**:2 (1976), 131–147.

[Keren et al. 2008] K. Keren, Z. Pincus, G. M. Allen, E. L. Barnhart, G. Marriott, A. Mogilner, and J. A. Theriot, "Mechanism of shape determination in motile cells", *Nature* **453**:7194 (2008), 475–480.

[Keren et al. 2009] K. Keren, P. T. Yam, A. Kinkhabwala, A. Mogilner, and J. A. Theriot, "Intracellular fluid flow in rapidly moving cells", *Nat. Cell Biol.* **11**:10 (2009), 1219–1224.

[Kruse et al. 2006] K. Kruse, J. F. Joanny, F. Jülicher, and J. Prost, "Contractility and retrograde flow in lamellipodium motion", *Phys. Biol.* **3**:2 (2006), 130–137.

[Kuusela and Alt 2009] E. Kuusela and W. Alt, "Continuum model of cell adhesion and migration", *J. Math. Biol.* **58**:1-2 (2009), 135–161.

[Larripa and Mogilner 2006] K. Larripa and A. Mogilner, "Transport of a 1D viscoelastic actin-myosin strip of gel as a model of a crawling cell", *Physica A* **372**:1 (2006), 113–123.

[Lee et al. 1993] J. Lee, A. Ishihara, J. A. Theriot, and K. Jacobson, "Principles of locomotion for simple-shaped cells", *Nature* **362**:6416 (1993), 167–171.

[Lo et al. 2000] C.-M. Lo, H.-B. Wang, M. Dembo, and Y.-L. Wang, "Cell movement is guided by the rigidity of the substrate", *Biophys. J.* **79**:1 (2000), 144–152.

[Loitto et al. 2009] V. M. Loitto, T. Karlsson, and K.-E. Magnusson, "Water flux in cell motility: expanding the mechanisms of membrane protrusion", *Cell Motil. Cytoskelet.* **66**:5 (2009), 237–247.

[Mattila and Lappalainen 2008] P. K. Mattila and P. Lappalainen, "Filopodia: molecular architecture and cellular functions", *Nat. Rev. Mol. Cell Biol.* **9**:6 (2008), 446–454.

[Mitchison et al. 2008] T. J. Mitchison, G. T. Charras, and L. Mahadevan, "Implications of a poroelastic cytoplasm for the dynamics of animal cell shape", *Sem. Cell Dev. Biol.* **19**:3 (2008), 215–223.

[Okeyo et al. 2009] K. O. Okeyo, T. Adachi, J. Sunaga, and M. Hojo, "Actomyosin contractility spatiotemporally regulates actin network dynamics in migrating cells", *J. Biomech.* **42**:15 (2009), 2540–2548.

[Oster and Perelson 1987] G. F. Oster and A. S. Perelson, "The physics of cell motility", *J. Cell Sci. Suppl.* **8** (1987), 35–54.

[Palecek et al. 1997] S. P. Palecek, J. C. Loftus, M. H. Ginsberg, D. A. Lauffenburger, and A. F. Horwitz, "Integrin-ligand binding properties govern cell migration speed through cell-substratum adhesiveness", *Nature* **385**:6616 (1997), 537–540.

[Pelham and Wang 1997] R. J. Pelham, Jr. and Y.-L. Wang, "Cell locomotion and focal adhesions are regulated by substrate flexibility", *Proc. Nat. Acad. Sci. USA* **94**:25 (1997), 13661–13665.

[Peskin et al. 1993] C. S. Peskin, G. M. Odell, and G. F. Oster, "Cellular motions and thermal fluctuations: the Brownian ratchet", *Biophys. J.* **65**:1 (1993), 316–324.

[Petrie et al. 2009] R. J. Petrie, A. D. Doyle, and K. M. Yamada, "Random versus directionally persistent cell migration", *Nat. Rev. Mol. Cell Biol.* **10**:8 (2009), 538–549.

[Peyton and Putnam 2005] S. R. Peyton and A. J. Putnam, "Extracellular matrix rigidity governs smooth muscle cell motility in a biphasic fashion", *J. Cell. Physiol.* **204**:1 (2005), 198–209.

[Prass et al. 2006] M. Prass, K. Jacobson, A. Mogilner, and M. Radmacher, "Direct measurement of the lamellipodial protrusive force in a migrating cell", *J. Cell Biol.* **174**:6 (2006), 767–772.

[Rodriguez et al. 1994] E. K. Rodriguez, A. Hoger, and A. D. McCulloch, "Stress-dependent finite growth in soft elastic tissues", *J. Biomech.* **27**:4 (1994), 455–467.

[Rubinstein et al. 2005] B. Rubinstein, K. Jacobson, and A. Mogilner, "Multiscale two-dimensional modeling of a motile simple-shaped cell", *Multiscale Model. Simul.* **3**:2 (2005), 413–439.

[Rubinstein et al. 2009] B. Rubinstein, M. F. Fournier, K. Jacobson, A. B. Verkhovsky, and A. Mogilner, "Actin-myosin viscoelastic flow in the keratocyte lamellipod", *Biophys. J.* **97**:7 (2009), 1853–1863.

[Schaub et al. 2007] S. Schaub, S. Bohnet, V. M. Laurent, J.-J. Meister, and A. B. Verkhovsky, "Comparative maps of motion and assembly of filamentous actin and myosin II in migrating cells", *Mol. Biol. Cell* **18**:10 (2007), 3723–3732.

[Stolarska et al. 2009] M. A. Stolarska, Y. Kim, and H. G. Othmer, "Multi-scale models of cell and tissue dynamics", *Phil. Trans. R. Soc. A* **367**:1902 (2009), 3525–3553.

[Stossel 1994] T. P. Stossel, "The machinery of cell crawling", *Sci. Am.* **271**:3 (1994), 54–63.

[Svitkina et al. 1997] T. M. Svitkina, A. B. Verkhovsky, K. M. McQuade, and G. G. Borisy, "Analysis of the actin-myosin II system in fish epidermal keratocytes: mechanism of cell body translocation", *J. Cell Biol.* **139**:2 (1997), 397–415.

[Taber 2001] L. A. Taber, "Biomechanics of cardiovascular development", *Annu. Rev. Biomed. Eng.* **3** (2001), 1–25.

[Taber 2004] L. A. Taber, *Nonlinear theory of elasticity: applications in biomechanics*, World Scientific, River Edge, NJ, 2004.

[Taber 2008] L. A. Taber, "Theoretical study of Beloussov's hyper-restoration hypothesis for mechanical regulation of morphogenesis", *Biomech. Model. Mechanobiol.* **7** (2008), 427–441.

[Verkhovsky et al. 1999] A. B. Verkhovsky, T. M. Svitkina, and G. G. Borisy, "Self-polarization and directional motility of cytoplasm", *Curr. Biol.* **9**:1 (1999), 11–20.

[Vicente-Manzanares et al. 2009] M. Vicente-Manzanares, X. F. Ma, R. S. Adelstein, and A. R. Horwitz, "Non-muscle myosin II takes centre stage in cell adhesion and migration", *Nat. Rev. Mol. Cell Biol.* **10**:11 (2009), 778–790.

[Wilson et al. 2010] C. A. Wilson, M. A. Tsuchida, G. M. Allen, E. L. Barnhart, K. T. Applegate, P. T. Yam, L. Ji, K. Keren, G. Danuser, and J. A. Theriot, "Myosin II contributes to cell-scale actin network treadmilling through network disassembly", *Nature* **465**:7296 (2010), 373–377.

[Yam et al. 2007] P. T. Yam, C. A. Wilson, L. Ji, B. Hebert, E. L. Barnhart, N. A. Dye, P. W. Wiseman, G. Danuser, and J. A. Theriot, "Actin-myosin network reorganization breaks symmetry at the cell rear to spontaneously initiate polarized cell motility", *J. Cell Biol.* **178**:7 (2007), 1207–1221.

[Yang et al. 1994] M. Yang, L. A. Taber, and E. B. Clark, "A nonlinear poroelastic model for the trabecular embryonic heart", *J. Biomech. Eng.* (*ASME*) **116**:2 (1994), 213–223.

[Young and Mitran 2010] J. Young and S. A. Mitran, "A numerical model of cellular blebbing: a volume-conserving, fluid-structure interaction model of the entire cell", *J. Biomech.* **43**:2 (2010), 210–220.

[Zhang 2009] Y.-W. Zhang, "Mechanics of membrane instability in biological cells", *Appl. Phys. Lett.* **94**:16 (2009), 163903.

[Zhu and Skalak 1988] C. Zhu and R. Skalak, "A continuum model of protrusion of pseudopod in leukocytes", *Biophys. J.* **54**:6 (1988), 1115–1137.

[Zicha et al. 2003] D. Zicha, I. M. Dobbie, M. R. Holt, J. Monypenny, D. Y. H. Soong, C. Gray, and G. A. Dunn, "Rapid actin transport during cell protrusion", *Science* **300**:5616 (2003), 142–145.

LARRY A. TABER: lat@wustl.edu
Department of Biomedical Engineering, Washington University, 1 Brookings Drive, Box 1097, St. Louis, MO 63130, United States

YUNFEI SHI: ys12@cec.wustl.edu
Department of Biomedical Engineering, Washington University, 1 Brookings Drive, Box 1097, St. Louis, MO 63130, United States

LE YANG: le@biomed.wustl.edu
Department of Biomedical Engineering, Washington University, 1 Brookings Drive, Box 1097, St. Louis, MO 63130, United States

PHILIP V. BAYLY: baylyp@seas.wustl.edu
Department of Mechanical Engineering and Materials Science, Washington University, 1 Brookings Drive, Box 1097, St. Louis, MO 63130, United States

msp

A REMARKABLE STRUCTURE OF LEONARDO AND A HIGHER-ORDER INFINITESIMAL MECHANISM

TIBOR TARNAI AND ANDRÁS LENGYEL

Dedicated to the memory of Marie-Louise Steele

This paper is concerned with the static and kinematic behavior of two chain-like bar-and-joint assemblies which have the same topology. One is a structure which is both statically and kinematically indeterminate, and constitutes a higher-order infinitesimal mechanism. The other is a structure which is both statically and kinematically determinate, introduced by Leonardo da Vinci in the *Codex Madrid*. Proceeding along the internal joints from the bottom to the top of the assembly, the lateral components of the displacements of the internal joints of the infinitesimal mechanism show an exponential decay, and the forces in the internal bars of Leonardo's structure show an exponential growth. It is pointed out that, in the elastic model of Leonardo's structure, the propagation of displacements of internal joints and the propagation of forces in internal bars also show an exponential character in a modified form. This work also provides some hints for overcoming difficulties arising in higher-order infinitesimal mechanisms, and corrects minor mistakes made by Leonardo.

1. Introduction

Cable nets, cable domes, and tensegrity structures, from the mechanical point of view, are usually bar-and-joint assemblies that constitute infinitesimal mechanisms. The answer to the important question of whether self-stress may impart first-order stiffness to such a structure depends on whether or not it is an infinitesimal mechanism of first or higher order [Pellegrino and Calladine 1986]. (An infinitesimal mechanism is of the n-th order ($n = 1, 2, \ldots$) if it involves no elongation of any bar up to and including the n-th order but exhibits an elongation of order $n + 1$ in at least one bar.)

In [Tarnai 1989], we discussed problems of definition and determination of the order of an infinitesimal mechanism, and presented, among other things, an example of a higher-order infinitesimal mechanism where the propagation of the displacements of the joints shows exponential decay. On the other hand, Nielsen [1998] called attention to a drawing of a structure by Leonardo da Vinci (see page 597) taken from the *Codex Madrid* [Leonardo 1493], which contains notes and drawings of his mechanical studies, and which was believed to be lost until the stunning announcement, in 1967, of its accidental rediscovery in the National Library in Madrid. In this structure, the propagation of forces in the bars shows an exponential growth.

In this paper, the accidental similarity between the two structures, as well as the duality between the exponential decay of the displacements and the exponential growth of the bar forces, are investigated.

2. Higher-order infinitesimal mechanisms

A structure is called an infinitesimal mechanism if it has only infinitesimal free motions. Koiter defined

> "an infinitesimal mechanism of the first order by its property that any infinitesimal displacement of the mechanism is accompanied by second order elongations of at least some of the bars. An infinitesimal mechanism is called of second (or higher) order if there exists an infinitesimal motion such that no bar undergoes an elongation of lower than the third (or higher) order" [Koiter 1984].

Koiter's definition can be mathematically formulated as follows. Consider a bar-and-joint assembly which contains b bars and consider a system of infinitesimal displacements of the joints. Let us denote an infinitesimal displacement component of a characteristic joint by u and the elongation of the bar k due to u by e_k. Write the power-series expansion of e_k in u:

$$e_k = a_{1k}u + a_{2k}u^2 + a_{3k}u^3 + \cdots \quad (k = 1, 2, \ldots, b). \tag{1}$$

For an infinitesimal mechanism, there always exists a system of infinitesimal displacements of joints such that, in (1), $a_{1k} = 0$ for $k = 1, 2, \ldots, b$. Such a system of infinitesimal displacements is obtained as the solution of a set of homogeneous compatibility equations. The displacement vector coordinates depend on the chosen coordinate system but the positions of the displaced joints do not. However, it is advantageous if one axis of the coordinate system is chosen in the direction of the displacement of the characteristic joint.

Definition 1 [Tarnai 1984]. An infinitesimal mechanism is *of order at least n* ($n \geq 1$) if there exists a system of infinitesimal displacements of joints such that, in (1), $a_{1k} = a_{2k} = \cdots = a_{nk} = 0$ for $k = 1, 2, \ldots, b$. An infinitesimal mechanism is *of order n* if it is of order at least n, but not of order at least $n + 1$.

Consider, for example, the mechanism in Figure 1, left. If we move it as in the middle figure, the lengthening of the horizontal bars is quadratic in the displacement u (for small u; see next section for details). The length of the vertical bars does not change. Therefore, for this displacement, all the coefficients a_{1k} in (1) vanish, so the system has order at least 1.

However, we can also consider the displacement in Figure 1, right. Here, if the central joint moves by u keeping the length of the bars attached to it constant and the far ends of these bars along the center line, the far ends move up and down by an amount proportional to u^2, and so the length of the horizontal bars increases by an amount proportional to u^4. For this displacement, then, a_{1k}, a_{2k} and a_{3k} vanish for all k. So in fact this system has order at least 3.
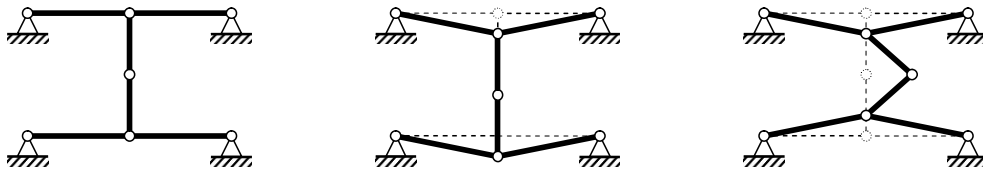
Figure 1. A two-DOF infinitesimal mechanism in the plane. Left: initial state. Middle: a displacement where the horizontal bars undergo second-order elongation. Right: a displacement where the same bars undergo fourth-order elongation.

Remark. We see from this example that, for multi-degree of freedom (DOF) infinitesimal mechanisms, displacements corresponding to different DOFs can result in different numbers n; but according to Definition 1, the maximum n will be the order of the infinitesimal mechanism. Thus, the definition can be reformulated as follows. For a system of infinitesimal displacements, consider the exponent of the first nonvanishing term in the series (1) for each bar, and take the minimum of these exponents. Then determine this minimum for each possible system of infinitesimal displacements, and take the maximum of these minima. The number obtained is one more than the order n of the infinitesimal mechanism.

Thus, in the assembly of Figure 1, we can say that the order is at least 1, based on the DOF whose displacement is illustrated in the middle diagram. But another DOF, illustrated in the diagram on the right, ensures that the order is at least 3. Since we have to take the maximum, we can declare that the assembly in Figure 1, left is a third-order infinitesimal mechanism. (It can be shown that no displacement leads to a higher order.)

In mathematics, the term "n-th order infinitesimal mechanism" is not used. Mathematicians usually deal with n-th order rigidity and n-th order flexes instead [Connelly 1980].

3. A series of infinitesimal mechanisms

Consider the bar-and-joint assembly in Figure 2a, all bars being of unit length. As shown in the figure, we denote by x the horizontal component of the displacement of joint B and by $2y$ the vertical component of the displacement of joint A. We constrain the displacements so that bars 1 and 2 undergo the same elongation, denoted by e, and bars 3 and 4 are inextensional (elongation zero). We can express the displacement component y in terms of x, and the elongation e in terms of y, as

$$y = 1 - \sqrt{1 - x^2} \quad \text{and} \quad e = \sqrt{1 + (2y)^2} - 1.$$

Taking the binomial series expansion of these functions, we get

$$y = \tfrac{1}{2} x^2 + \tfrac{1}{8} x^4 + \tfrac{1}{16} x^6 + \cdots, \tag{2}$$

$$e = \tfrac{1}{2} (2y)^2 - \tfrac{1}{8} (2y)^4 + \cdots. \tag{3}$$
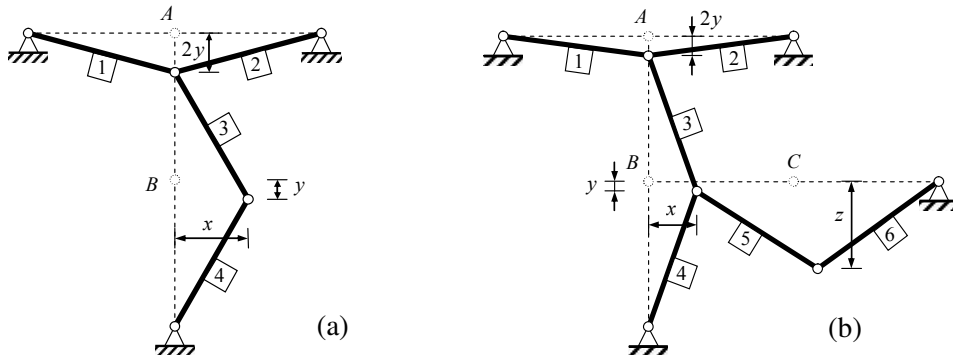
Figure 2. A planar assembly composed of (a) two and (b) three three-pinned frame units of collinear joints, and their infinitesimal free displacements.

Introduction of (2) into (3) yields

$$e = \tfrac{1}{2}x^4 + \tfrac{1}{4}x^6 + \tfrac{1}{32}x^8 + \cdots ,$$

that is, fourth-order elongations arise in bars 1 and 2. Bars 3 and 4 have no elongation, by assumption. Thus, $n + 1 = 4$, and so $n = 3$: the assembly in Figure 2a is an infinitesimal mechanism of order at least 3. It can be shown that no displacement leads to a higher order, so the assembly in Figure 2a is a third-order infinitesimal mechanism. (A similar calculation applied to the rightmost configuration of the structure in Figure 1 justifies our earlier assertion that that structure has order at least 3.)

Now let us supplement the bar-and-joint assembly in Figure 2a with a horizontal three-pinned frame of collinear joints, connected to joint B, as in Figure 2b. All bars have unit length, and bars 3, 4, 5, and 6 are taken as inextensional. Let us apply a vertical displacement z at joint C. In this case, the displacement components x and y shown in Figure 2b and the elongation e of bars 1 and 2 are related by

$$x = \sqrt{1 - (1 - y)^2}, \qquad (z - y)^2 + \big(2 - x - \sqrt{1 - z^2}\,\big)^2 = 1, \qquad e = \sqrt{1 + (2y)^2} - 1.$$

Using Maple to eliminate x and y and to expand the resulting formula for e in powers of z, we obtain for the elongation of bars 1 and 2 the series

$$e = \tfrac{1}{2}z^8 + \tfrac{1}{2}z^{10} - z^{11} + \tfrac{11}{16}z^{12} + \cdots ;$$

that is, eighth-order elongations arise in bars 1 and 2. Thus, the assembly in Figure 2b is an infinitesimal mechanism of order at least 7. It can be shown that no displacement leads to a higher order, so the assembly in Figure 2b is a seventh-order infinitesimal mechanism.

More generally, it can be shown [Tarnai 1989] that an the assembly composed of M three-pinned frame units with collinear joints, as in Figure 3, is an infinitesimal mechanism of order $2^M - 1$.
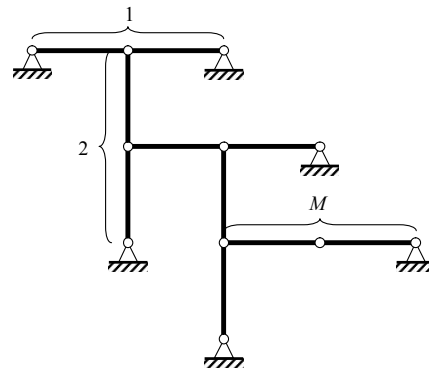
Figure 3. A planar assembly composed of M three-pinned frame units of collinear joints.

4. A chain-like infinitesimal mechanism

Let the value of M be equal to 9; thus we have the assembly in Figure 4, which is both statically and kinematically indeterminate, constituting an infinitesimal mechanism of order $n = 2^9 - 1 = 511$. We selected $M = 9$ because later we want to compare this infinitesimal mechanism to the similar structure by Leonardo, mentioned in the Section 1. Let each bar be 6 m long. We suppose that all bars are inextensional except the uppermost two, where extensions are allowed to develop freely. Let us investigate
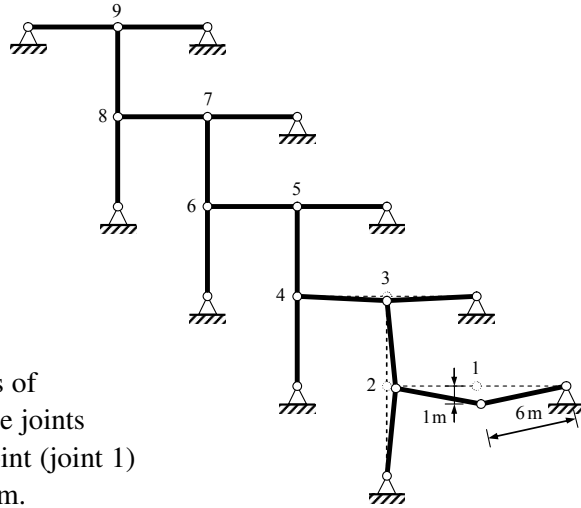
Figure 4. A chain-like planar assembly composed of nine three-pinned frame units of collinear joints, and displacements of middle joints (numbered 1 through 9) when the lowest joint (joint 1) has a prescribed vertical displacement of 1 m.

the motion of the assembly if the lowermost joint (joint 1) has a vertical displacement of 1 m. In this investigation, we apply the approximation $\sqrt{1+a} \approx 1 + \frac{1}{2}a$, and consider only the lateral displacement of the middle joint of a three-pinned frame unit, that is, the vertical displacement for a horizontal three-pinned unit and the horizontal displacement for a vertical three-pinned unit. The axial displacements of the middle joints are neglected. If the lateral displacement of joint i is δ_i then, from this, a displacement δ_{i+1} arises at joint $i + 1$, and the directions of the two displacements are perpendicular to each other. According to the approximations above, we have

$$\delta_{i+1} = 2\left(6 - \sqrt{6^2 - \delta_i^2}\right),$$

which yields $\delta_{i+1} = \frac{1}{6}\delta_i^2$, that is,

$$\delta_i = \left(\tfrac{1}{6}\right)^{2^{i-1}-1} \quad i = 1, 2, \ldots, 9. \tag{4}$$

It is easy to see that, proceeding from the bottom to the top of the assembly, the displacements are *exponentially decreasing*.

It is also possible to calculate the exact position and displacement of the joints, taking into account the displacement in the axial direction in addition to that in the lateral direction. In fact, the numbered joints move along circular arcs centered at the supported joints (not numbered) to which they are connected by 6-meter bars. Once the position of joint 1 is defined, the next one can be calculated geometrically as the intersection of two circles centered at joint 1 and the next supported joint, respectively. The procedure goes on this way for all the numbered joints. The results are only slightly different from those of the approximate calculations.

The displacements δ_i obtained both by approximate calculations, according to (4), and by exact ones are given for the first few joints in Table 1. Thus the 1 m displacement of the first joint decreases to approximately 5 mm at the third joint, and 4 μm at the fourth joint. The displacement of the fifth joint is only 2 pm, less than the size of an atom. (Atomic radii lie in the range 50–200 pm.) From a practical point of view, this means that from the fourth joint on there is basically no displacement. This surprising result illustrates the most important property of higher-order infinitesimal mechanisms: mathematically

i	δ_i^{approx} (m)	δ_i^{exact} (m)
1	1	1
2	$6^{-1}\ \ = 1.667 \times 10^{-1}$	1.674×10^{-1}
3	$6^{-3}\ \ = 4.630 \times 10^{-3}$	4.674×10^{-3}
4	$6^{-7}\ \ = 3.572 \times 10^{-6}$	3.641×10^{-6}
5	$6^{-15} = 2.127 \times 10^{-12}$	2.210×10^{-12}

Table 1. Lateral displacements of the middle joints of three-pinned units of the assembly in Figure 4. (At joint i, the approximate value δ_i^{approx} and the exact value δ_i^{exact} of the lateral displacement are given).

they have only infinitesimal free motions, but physically they behave locally like finite mechanisms. The motion of mechanisms of this kind was also investigated in [Hortobágyi 2000].

5. Leonardo's structure

5.1. *Structure in the Codex Madrid.* On folio 75R of the Codex Madrid [Leonardo 1493], which we reproduce in Figure 5, there is a drawing of a chain-like structure composed of 12-unit-long cables. Each cable is V-shaped, and the distance between the midpoint of the cable and the straight line connecting the end points of the cable is 1 unit. The straight lines connecting the end points of the cables are alternately horizontal and vertical. (The bar-and-joint model of Leonardo's structure shown in Figure 6 is both statically and kinematically determinate, unlike the infinitesimal mechanism studied in Section 4.)

Leonardo investigated the propagation of forces in the structure if the bottom cable is loaded at its midpoint with a weight of one pound. He made an approximate analysis. He considered the cables to be inextensional and neglected the fact that a cable segment, joining another cable at its midpoint, has an inclination: he calculated the forces in the cables as if the direction of the joining cable segment were perpendicular to the straight line connecting the end points of the cable. In this way, Leonardo obtained equal forces in both segments of a cable. Let φ denote the angle of inclination of cable segments. From the geometrical data it follows that

$$\sin \varphi = \tfrac{1}{6}.$$

By resolving the forces horizontally and vertically at the node loaded by the one-pound weight, force S_1 in the lowermost cable is obtained: $S_1 = 3$ pounds. For the second cable, this member force of 3 pounds acts as a load and, because of the above-mentioned approximation, the force in the second cable is $S_2 = 3S_1$. The force in the k-th cable from below is $S_k = 3S_{k-1}$, that is,

$$S_k = 3^k \text{ pounds.}$$

Therefore, the propagation of forces in cables is *exponentially increasing*: the forces form a geometric progression with ratio 3. From the 1 pound load applied at the midpoint of the lowermost cable, a force S_9 of 19,683 pounds ($S_9 = 3^9$ pounds) arises in the uppermost cable. (Leonardo wrote 19,530 pounds for that force; indeed, we see in Figure 5 that in two steps of the progression his multiplications by 3 are slightly off: $27 \rightarrow 80$, instead of 81, and $2160 \rightarrow 6510$, instead of 6480. The reason for this is not known.)
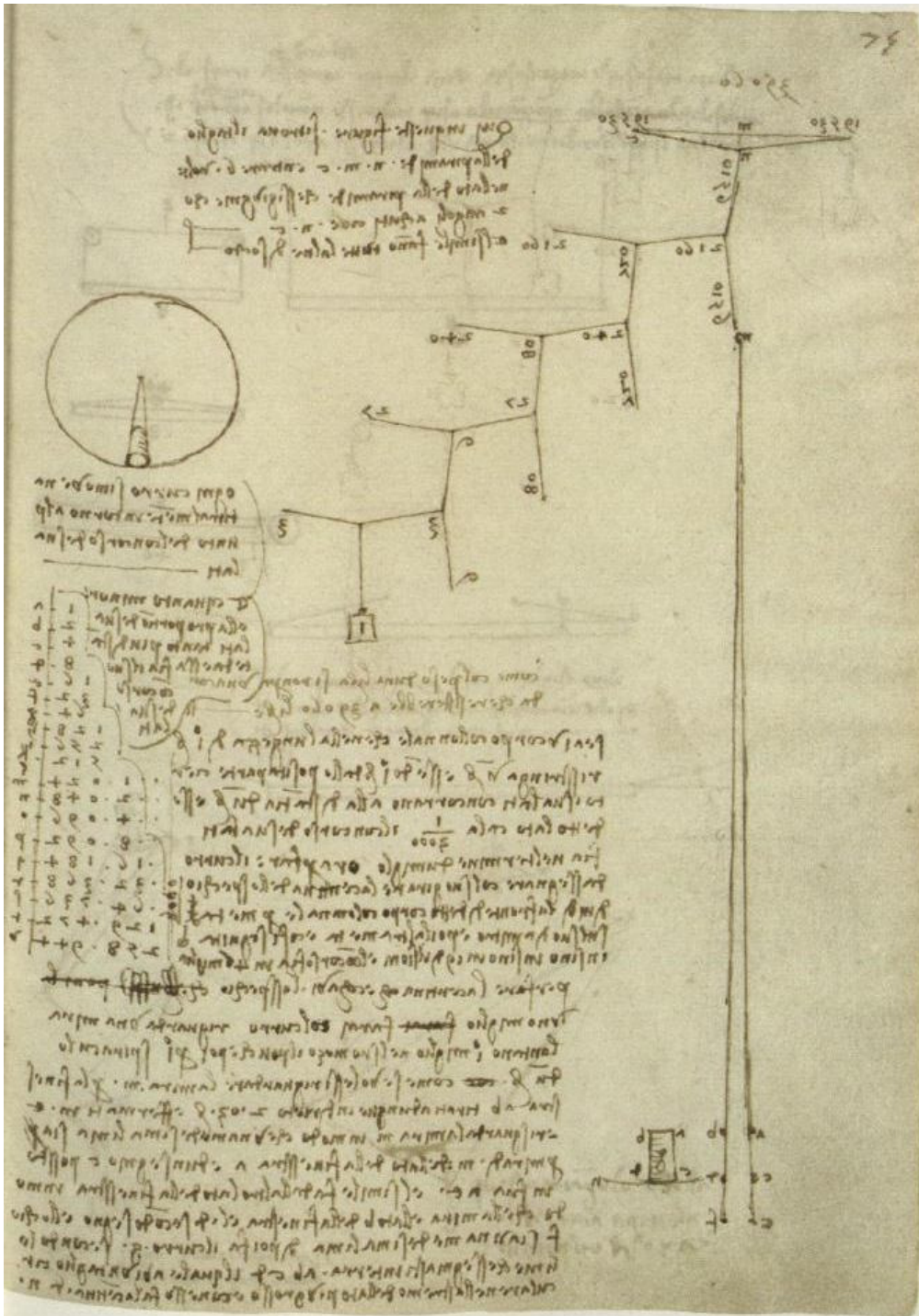
Figure 5. Leonardo da Vinci's drawing of a planar assembly [Leonardo 1493, folio 75R]. The forces in the cables, arising from a one-pound load at the lowest cable, are written on the respective cables. Leonardo, as usual, made this note in mirror image.
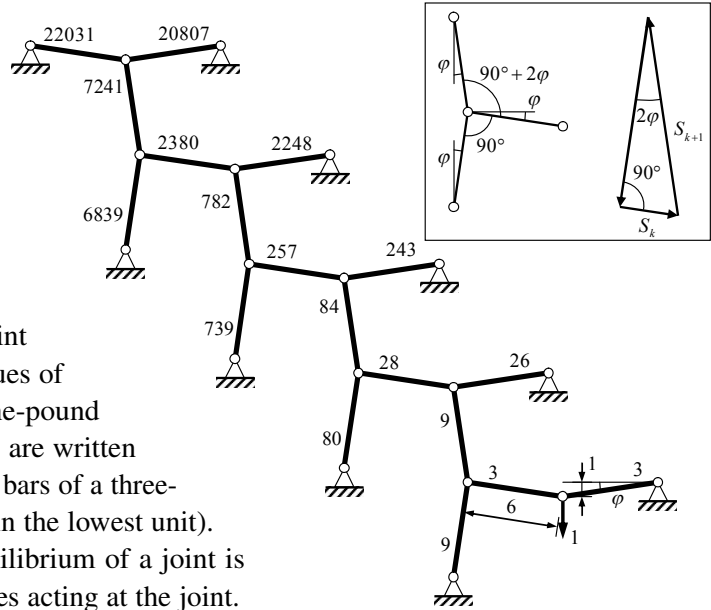
Figure 6. The inextensional bar-and-joint model of Leonardo's structure. The values of the forces in the bars, arising from a one-pound vertical load at the lowest internal joint, are written on the respective bars. Forces in the two bars of a three-pinned frame unit are different (except in the lowest unit). In the upper right-hand corner, the equilibrium of a joint is shown by the vector triangle of bar forces acting at the joint.

5.2. Structure with rigid members subject to unit load. These minor inaccuracies compel us to analyse the equilibrium of Leonardo's structure in detail. We consider all nodes in their exact positions and all members in their exact directions. We call the middle points of the three-pinned frame units internal joints and the bars connecting two internal joints internal bars. Regarding all members as infinitely rigid, one can obtain the bar forces at all nodes from a vector triangle. A typical equilibrium system is shown in the upper right-hand corner of Figure 6. The bar forces are displayed at their corresponding locations rounded to the nearest integer. Note that the exponential growth of the forces in the internal bars follows the quotient

$$\frac{S_{k+1}}{S_k} = \frac{1}{\sin 2\varphi} = \frac{1}{2\sin\varphi\cos\varphi} = \frac{1}{2\frac{1}{6}\frac{\sqrt{35}}{6}} = \frac{18}{\sqrt{35}} = 3.042555\ldots,$$

which is slightly larger than 3.

5.3. Structure with elastic members subject to a load. Leonardo's discovery that a single load at the lowermost node produces an approximately twenty thousand times magnified cable force at the top indicates that the elastic deformations of the members ought to be considered.

Let all members of length L be linearly elastic with Young's modulus E and cross-sectional area A. Let us apply a single load at the lowermost node similarly to Leonardo's drawing. We expect an exponential growth in the bar forces upwards. The elastic elongations $e_k = S_k L/(EA)$ of the members due to bar forces S_k result in increasing nodal displacements downwards in a twofold way. On the one hand, longer bars of three-pinned frame units result in larger heights, and, on the other hand, they produce a shorter base length (and consequently a larger height) in the next three-pinned frame unit attached. Thus, this double exponential change in the bar forces and the displacements implies that a surprisingly small force can produce large displacements even if the structure is quite stiff.

EA (kN)	D (mm)	$S_{9,10}$ (kN)	$\sigma_{9,10}$ (MPa)	δ_1 (m)
10^0	7.979×10^{-2}	4.410×10^{-3}	8.819×10^2	4.596×10^0
10^1	2.523×10^{-1}	1.721×10^{-2}	3.441×10^2	4.293×10^0
10^2	7.979×10^{-1}	6.306×10^{-2}	1.261×10^2	3.847×10^0
10^3	2.523×10^0	2.201×10^{-1}	4.402×10^1	3.253×10^0
10^4	7.979×10^0	7.365×10^{-1}	1.473×10^1	2.533×10^0
10^5	2.523×10^1	2.326×10^0	4.652×10^0	1.728×10^0
10^6	7.979×10^1	6.538×10^0	1.308×10^0	9.295×10^{-1}
10^7	2.523×10^2	1.427×10^1	2.853×10^{-1}	3.183×10^{-1}
10^8	7.979×10^2	2.038×10^1	4.076×10^{-2}	5.493×10^{-2}
10^9	2.523×10^3	2.184×10^1	4.368×10^{-3}	6.096×10^{-3}

Table 2. Extremal values of bar forces and nodal lateral displacements under the vertical load $F = 1\,\text{N}$ applied at joint 1. The cross-section diameter D, the maximal force $S_{9,10}$, the normal stress $\sigma_{9,10}$ in bar $\{9, 10\}$ (according to the numbering of Figure 7), and the maximal vertical displacement δ_1 of joint 1 are given for different values of the normal stiffness EA of the bars.

In order to bring this structure closer to reality and have a better picture of its behavior, let us choose steel as our material, with Young's modulus $E = 200\,\text{GPa}$, apply a small unit load ($F = 1\,\text{N}$) at the lowermost node, and consider a series of different values of tensile stiffness: $EA = 10^0, 10^1, \ldots, 10^9\,\text{kN}$. Let the 6 m-long bars be of circular cross-section. Table 2 shows the cross-section diameter, the largest bar force (at the top of the structure), the stress, and the largest vertical displacement (at the lowermost node) as functions of the tensile stiffness. As the tensile stiffness increases, the bar forces also increase but at a lesser rate. Consequently, the stresses and deformations decrease, resulting in smaller displacements. The larger the tensile stiffness, the more closely the force-displacement system approximates that of the infinitely rigid structure of Section 5.2. The maximal force in the table also gives a hint of the rate of the force propagation in the chain, which is less than that in the case of the rigid structure. This is because when the displacements are large the vector triangles change significantly.

However, a relatively good match (for example, in the range $EA = 10^8$–$10^9\,\text{kN}$) can only be obtained for unrealistically large cross-sections, which means that this cable structure with ordinary cross-sections (such as $EA = 10^3$–$10^4\,\text{kN}$) indeed undergoes large displacements even for a small load. Figure 7 shows the deformed structure with tensile stiffness $EA = 10^4\,\text{kN}$. The member forces $S_{i,i+1}^F$ and the lateral components δ_i^F of the nodal displacements are listed in Table 3. The results were obtained by a large displacement iterative analysis.

According to Table 3, the forces $S_{i,i+1}^F$ in internal bars form a monotonically increasing sequence where the ratio $S_{i+1,i+2}^F / S_{i,i+1}^F$ is monotonically increasing with an increase in i ($i = 1, 2, \ldots, 8$). Thus, the bar forces show an exponential-like increase, but the change is faster than in a geometric progression. A similar observation can be made for the displacements, but in an inverse manner. The lateral components δ_i^F of the displacements of the internal joints form a monotonically decreasing sequence where the ratio $\delta_{i+1}^F / \delta_i^F$ is monotonically decreasing with an increase in i. Thus, the lateral components of the displacements show an exponential-like decrease, but the change is faster than in a geometric progression.
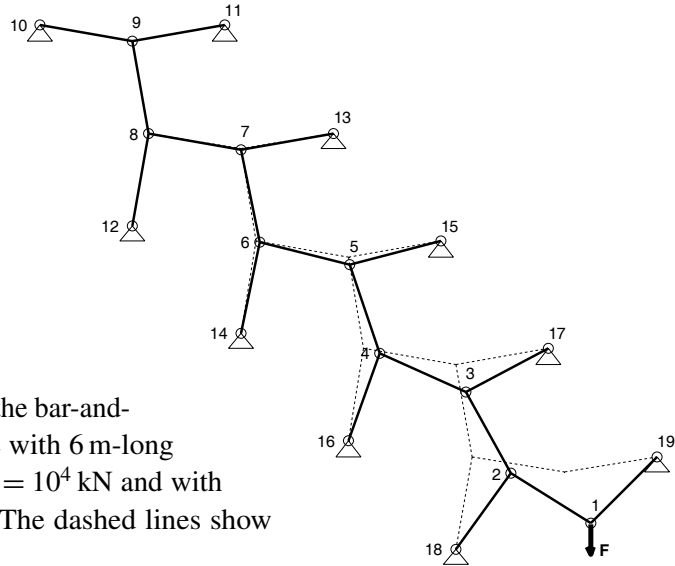
Figure 7. The equilibrium shape of the bar-and-joint model of Leonardo's structure with 6 m-long linearly elastic bars of stiffness $EA = 10^4$ kN and with a vertical load $F = 1$ N at joint 1. The dashed lines show the structure at the rest position.

$\{i, i+1\}$	$S_{i,i+1}^F$ (kN)	$S_{i,i+1}^{\Delta T}$ (kN)	i	δ_i^F (m)	$\delta_i^{\Delta T}$ (m)	$\delta_i^{\Delta T\,\text{lim}}$ (m)
$\{1, 2\}$	8.884×10^{-4}	7.365×10^{-4}	1	2.533×10^0	-1.011×10^0	-1.000×10^0
$\{2, 3\}$	1.130×10^{-3}	2.630×10^{-3}	2	1.756×10^0	-1.621×10^{-1}	-1.636×10^{-1}
$\{3, 4\}$	1.804×10^{-3}	8.357×10^{-3}	3	1.040×10^0	-4.381×10^{-2}	-4.931×10^{-2}
$\{4, 5\}$	3.741×10^{-3}	2.568×10^{-2}	4	5.088×10^{-1}	-8.882×10^{-3}	-1.579×10^{-2}
$\{5, 6\}$	9.526×10^{-3}	7.806×10^{-2}	5	2.078×10^{-1}	2.275×10^{-3}	-5.139×10^{-3}
$\{6, 7\}$	2.707×10^{-2}	2.365×10^{-1}	6	7.509×10^{-2}	5.880×10^{-3}	-1.675×10^{-3}
$\{7, 8\}$	8.041×10^{-2}	7.154×10^{-1}	7	2.555×10^{-2}	6.931×10^{-3}	-5.409×10^{-4}
$\{8, 9\}$	2.427×10^{-1}	2.162×10^0	8	8.432×10^{-3}	6.864×10^{-3}	-1.684×10^{-4}
$\{9, 10\}$	7.365×10^{-1}	6.542×10^0	9	2.574×10^{-3}	5.595×10^{-3}	-4.255×10^{-5}

Table 3. Forces in internal bars and lateral displacements of internal joints of the structure with bar stiffness $EA = 10^4$ kN. Force $S_{i,i+1}^F$ in bar $\{i, i+1\}$, the lateral component δ_i^F of displacement of internal joint i for a vertical load $F = 1$ N applied at joint 1, force $S_{i,i+1}^{\Delta T}$ in bar $\{i, i+1\}$, and the lateral component $\delta_i^{\Delta T}$ of the displacement of internal joint i for temperature drop $\Delta T = -40°$ C are given. $\delta_i^{\Delta T\,\text{lim}}$ provides the lateral component of displacement of internal joint i in the limit state due to a decrease in temperature. Downward and rightward displacement components are considered positive.

5.4. *Thermal effect.* We have seen that Leonardo's structure is extremely sensitive to small deformations of the members. Let us now examine the effects of the shortening of the elements contrary to the elongations due to tensile forces.

Consider a simultaneous shortening of the lengths L as a kinematic load, for example, due to uniform cooling of all bars. Denoting the thermal expansion coefficients by α [1/°C], all bars undergo elongation $e = L \cdot \alpha \cdot \Delta T$ due to a temperature rise ΔT [°C] (or shortening due to a temperature drop, as in our case).
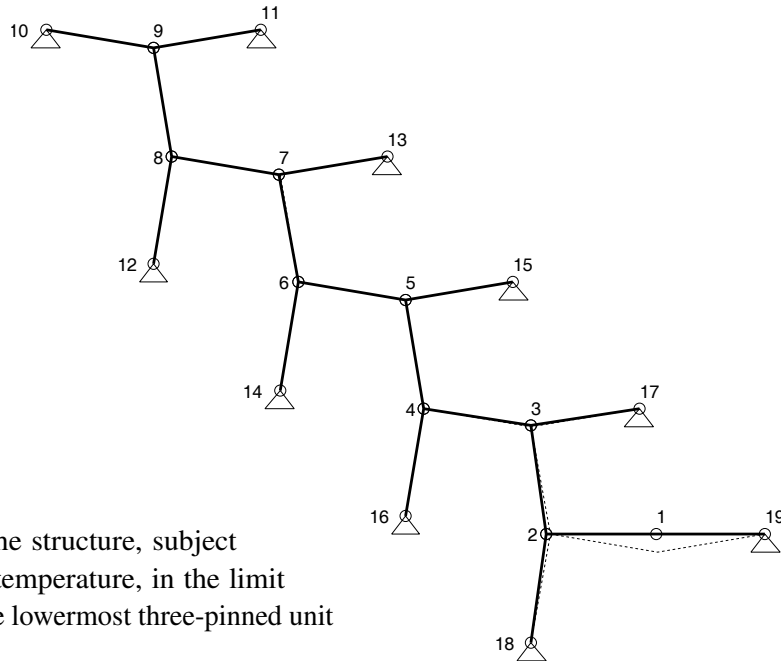
Figure 8. The shape of the structure, subject to a uniform decrease in temperature, in the limit state where the joints of the lowermost three-pinned unit become collinear.

The new position of the internal nodes can be calculated, starting with the uppermost, proceeding to the one below it, and then to the one to the right, and so on. The shortening of the bars makes the height of a three-pinned frame unit smaller, and at the same time the base length of the next unit longer. Thus in each step the nodal displacements are magnified. This phenomenon is similar to the one described in Section 5.3 but in the opposite direction. This means that the internal nodes now move towards the base of the frame units, that is, upwards and leftwards, alternately.

Note that none of the base lengths can exceed $2L$, which sets a limit on the measure of the cooling. This limit occurs at a small $e = -0.00000709091\ldots$ m shortening (which is equivalent to a $\Delta T = -0.0984849\ldots °$C change in temperature in the case of steel with $\alpha = 1.2 \cdot 10^{-5} 1/°$C). Now the joints of the last three-pinned frame unit are collinear (as shown in Figure 8). This is a limit state where the originally both statically and kinematically determinate structure becomes a statically and kinematically indeterminate structure. Until this decrease in temperature, the assembly is free of stresses.

At this limit state, the lateral displacements $\delta_i^{\Delta T \, \lim}$ of all the internal nodes are summarized in the last column of Table 3. The direction is emphasized by the use of the minus signs before the displacement values. Proceeding from the bottom to the top of the structure, the lateral displacements are exponentially decreasing, which can be easily illustrated in a semilogarithmic scale plot (Figure 9). The line in the figure is approximately linear except at the first and last nodes. The ratio of decrease is around $3.1^{-1} = 0.32\ldots$.

If the cooling continues, stresses appear in the structure. In further calculations, the elastic properties of the structure, considered in Section 5.3, should be taken into account. The same algorithm has been applied to calculate the equilibrium shape as in Section 5.3. Our calculations show that the structure basically tightens up, at this point having only tensile forces, and only minor displacements occur beyond that. The shape of the structure is similar to the one shown in Figure 8. The bar forces and the stresses naturally increase with the change in temperature but the elongations are countered with the thermal
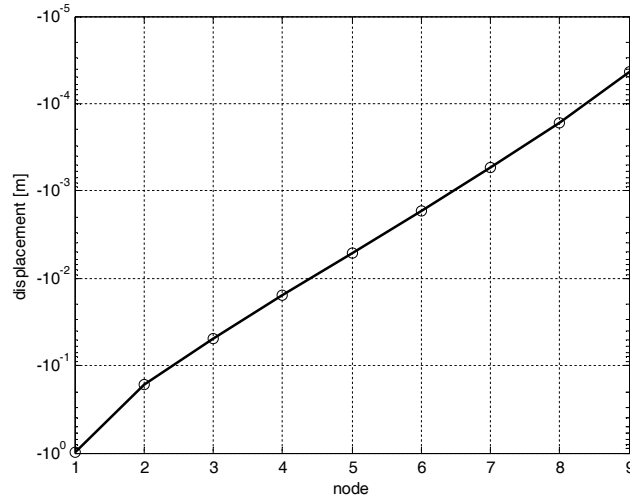
Figure 9. Lateral displacement of the internal nodes in the limit state.

shortenings. The member forces $S_{i,i+1}^{\Delta T}$ and the lateral components $\delta_i^{\Delta T}$ of the nodal displacements are listed in Table 3 for $\Delta T = -40°\,\text{C}$. In accordance with our earlier notation, the positive displacement values refer to rightward or downward motion, while the negative displacement values refer to leftward or upward motion. In the progression of the forces in the internal bars, the ratio is decreasing from $3.572\ldots$ to $3.0265\ldots$; in a longer range, it is around 3.03, that is, close to the value of $3.0425\ldots$ found for the infinitely rigid structure loaded with a concentrated force at joint 1.

6. Concluding remarks

It is ascertained that, proceeding from bottom to top, the lateral components of the displacements of internal joints of the inextensional mechanism in Figure 4 show an exponential decay, where the displacements decrease much faster than a geometric progression, and the forces in the internal bars of Leonardo's infinitely rigid structure in Figure 6 show an exponential growth according to a geometric progression with common ratio $3.0425\ldots$. If the members in Leonardo's structure are linearly elastic equal bars, then elastic deformations change the structural behavior. Though the lateral components of the displacements of internal joints and the forces in internal bars show exponential-like decrease and increase, respectively, the change is not as fast as in the case of the inextensional members. In the investigated structure (with $EA = 10^4\,\text{kN}$), the sequence of the lateral components of nodal displacements has a ratio monotonically decreasing from 1.4426^{-1} to 3.2761^{-1}, and the sequence of bar forces has a ratio monotonically increasing from 1.2721 to 3.0352, that is, the ratio of change in both is approximately the same. If the force unit is Newtons instead of pounds, then the force in bar $\{9, 10\}$ of the infinitely rigid structure is $22{,}031\,\text{N}$ (Figure 6). As a consequence of the elastic properties of the structure, however, this value is decreased to $736\,\text{N}$ (Table 3), that is, *to approximately one thirtieth of its value*.

    With numerical calculation we found that, with a decrease in the magnitude of the load F on the elastic structure, while keeping monotonicity, the ratio in the sequence of forces in the internal bars increases, and converges to $3.042555\ldots$, the common ratio of the geometric progression of bar forces in the

infinitely rigid structure. For $F = 10^{-5}$ N, for instance, the ratio varies from $3.0368\ldots$ to $3.042553\ldots$. The nodal displacements are less regular: the ratio monotonically decreases from 3.0336^{-1} to 3.2891^{-1}.

The determination of the order of a chain-like infinitesimal mechanism was straightforward. This can give the impression that the determination of the order of an infinitesimal mechanism is an easy task. Unfortunately, this is not the case. Connelly called attention to difficulties [Tarnai 1989; Connelly and Servatius 1994] which partly come from the definitions of higher-order infinitesimal mechanisms. One problem can be if the elongation functions are not analytic at a point, for instance, where there is a cusp in the configuration space. This particular problem can be resolved by using series expansion with fractional exponents [Gáspár and Tarnai 1994]. Another, quite general mathematical problem is the parametrization. The displacement vector of any joint can be given as a function of parameter u, but we can define a new displacement vector as a function of parameter u^2. The result of this reparametrization is that the order of the infinitesimal mechanism will be twice as large as for the original parametrization. Consequently, the order can be arbitrarily large. This problem is not yet settled. A very promising approach was presented recently by [Stachel 2007] who suggested that the definition of an n-th order infinitesimal mechanism must be based on irreducible representations of flexes. Although the notion of a *higher-order infinitesimal mechanism* is a purely geometrical term, to simplify the analysis Koiter suggested the introduction of elasticity in the bars, that is, the use of the elastic energy function of the assembly and interpretation of the infinitesimal mechanism as the buckling mode of the assembly under zero loading. Koiter's idea overcomes the difficulties mentioned by Connelly, and, indeed, was successfully applied, for instance, by [Salerno 1992].

Our analysis in Section 5.2 has shown that the forces in Leonardo's structure show an exponential growth with ratio $3.0425\ldots$, which is different from Leonardo's own value of 3. However, it is possible to introduce a modification to the original structure to fit this number (see Figure 10). Now each cable
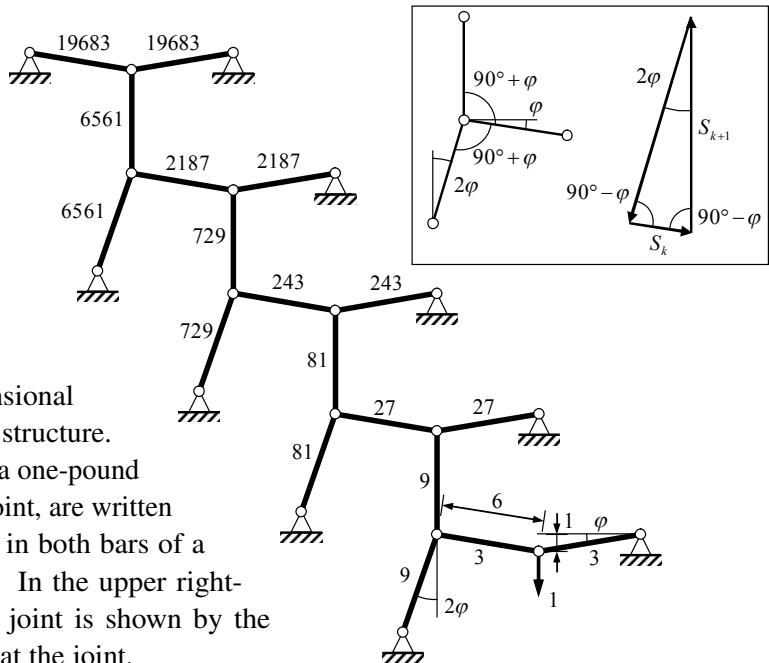
Figure 10. The modified inextensional bar-and-joint model of Leonardo's structure. The forces in the bars, arising from a one-pound vertical load at the lowest internal joint, are written on the respective bars. The forces in both bars of a three-pinned frame unit are equal. In the upper right-hand corner, the equilibrium of a joint is shown by the vector triangle of bar forces acting at the joint.

is connected to the middle point of the next one at an angle of $90° + \varphi$ in a symmetric way as shown in the inset in the upper right-hand corner.

References

[Connelly 1980] R. Connelly, "The rigidity of certain cabled frameworks and the second-order rigidity of arbitrarily triangulated convex surfaces", *Adv. Math.* **37**:3 (1980), 272–299.

[Connelly and Servatius 1994] R. Connelly and H. Servatius, "Higher-order rigidity: what is the proper definition?", *Discrete Comput. Geom.* **11**:1 (1994), 193–200.

[Gáspár and Tarnai 1994] Z. Gáspár and T. Tarnai, "Finite mechanisms have no higher-order rigidity", *Acta Tech. Hung.* **106**:3–4 (1994), 119–125.

[Hortobágyi 2000] Z. Hortobágyi, "Numerical analysis of inextensional, kinematically indeterminate assemblies", *Period. Polytech. Civ. Eng.* **44**:1 (2000), 43–55.

[Koiter 1984] W. T. Koiter, "On Tarnai's conjecture with reference to both statically and kinematically indeterminate structures", technical report 788, Laboratory for Engineering Mechanics, Delft University of Technology, 1984.

[Leonardo 1493] Leonardo da Vinci, *Tratado de estática y mecánica en italiano* (title added to manuscript 8937, Biblioteca Nacional de España), vol. 1, 1493. Facsimile edition published by McGraw-Hill, 1974.

[Nielsen 1998] J. Nielsen, "Leonardo da Vinci and the parallelogram of forces", *J. IASS* **39**:1 (1998), 47–53.

[Pellegrino and Calladine 1986] S. Pellegrino and C. R. Calladine, "Matrix analysis of statically and kinematically indeterminate frameworks", *Int. J. Solids Struct.* **22**:4 (1986), 409–428.

[Salerno 1992] G. Salerno, "How to recognize the order of infinitesimal mechanisms: a numerical approach", *Int. J. Numer. Methods Eng.* **35**:7 (1992), 1351–1395.

[Stachel 2007] H. Stachel, "A proposal for a proper definition of higher-order rigidity", in *Tensegrity Workshop 2007* (La Vacquerie, 2007), 2007.

[Tarnai 1984] T. Tarnai, "Comments on Koiter's classification of infinitesimal mechanisms", Technical report, Hungarian Institute for Building Science, Budapest, 1984.

[Tarnai 1989] T. Tarnai, "Higher-order infinitesimal mechanisms", *Acta Tech. Hung.* **102**:3–4 (1989), 363–378.

TIBOR TARNAI: tarnai@ep-mech.me.bme.hu
Department of Structural Mechanics, Budapest University of Technology and Economics, Müegyetem rkp. 3, Budapest, H-1521, Hungary

ANDRÁS LENGYEL: lengyel@eik.bme.hu
Department of Structural Mechanics, Budapest University of Technology and Economics, Müegyetem rkp. 3, Budapest, H-1521, Hungary

msp

TOPOLOGY OPTIMIZATION OF GEOMETRICALLY NONLINEAR STRUCTURES TRACING GIVEN LOAD-DISPLACEMENT CURVES

GIL HO YOON, JIN YEE NOH AND YOON YOUNG KIM

To design structures involving nonlinear structural responses by the topology optimization method is still a challenging problem. Here, the structural topology optimization tracing nonlinear load-displacement curves is investigated by employing the element connectivity parametrization formulation as it is expected to deal with low-density element related numerical instability more effectively than the element density based formulation. After the formulation is given in the setting of the element connectivity parametrization, the sensitivity analysis for load-displacement curve tracing problems implemented with a discretized finite element model is presented. Several numerical problems are considered to address issues occurring in the topology optimization of nonlinear structures. Finally, the findings from this investigation on the topology optimization tracing nonlinear load-displacement trajectories and future work are summarized as concluding remarks.

1. Introduction

Along with size and shape optimization methods, the topology optimization method [Bendsøe and Kikuchi 1988] has been used as an effective tool for optimizing structures and mechanical parts (see [Bendsøe and Sigmund 2003], for example). Compared with abundant investigations on linear structural problems, the topology optimization of nonlinear structural problems is relatively rare [Buhl et al. 2000; Bruns et al. 2002; Cho and Jung 2003; Yoon and Kim 2005b; Yoon et al. 2008]. In this study, we investigate the topology optimization of a continuum structure tracing a prescribed nonlinear load-displacement trajectory. Specifically, hardening or softening nonlinear responses without and without snap-through will be considered.

Structures exhibiting geometric hardening or softening behavior (see Figure 1) can be utilized for crash worthiness or energy absorption mechanism design. For example, one may wish to find a structure which can absorb a specific amount of impact energy in case of crash worthiness. There have been some interests in designing truss or beam structures to exhibit geometrical and/or material nonlinearities [Kamat and Ruangsilasingha 1985; Saxena 2005; Lu and Kota 2003; Ohsaki 2005; Huang and Xie 2008; Santer and Pellegrino 2008; Thai and Kim 2009] where the buckling loads of structures are maximized.

Figure 1. Geometric hardening or weakening model depicted with a displacement controlled Newton–Raphson scheme. (The prescribed displacement and the reaction force of the i-th load step are denoted by $^{t+\Delta t}U^i$ and $^{t+\Delta t}R^i$, respectively. The number of load steps is nl).

However, research on two- or three-dimensional continuum structures is relatively less active due to heavy computational cost and numerical instability [Buhl et al. 2000; Bruns et al. 2002; Yoon et al. 2007].

Perhaps, the first investigation on trajectory-following continuum topology optimization was done in [Sekimoto and Noguchi 2001] using the homogenization method. Trajectory problems were also solved in [Bruns et al. 2002] and [Bruns and Sigmund 2004] using the density based method. In these studies, an automated solution controlling method was employed to consider a snap-through phenomenon. To design compliant mechanisms generating given paths was also considered using the structural topology optimization algorithms [Lu and Kota 2003; Saxena 2005; Prasad and Diaz 2006]. In spite of these studies, some issues, such as low-density elements, still remain. Among others, this study is focused on the numerical instability of low-density elements and the adverse effects of postprocessing intermediate density elements to solid or void elements.

An optimized layout even at the converged state typically involves some low-density (and intermediate-density) elements. In this case, low-density continuum elements under a big load can inevitably lose the positive-definiteness of the system tangent stiffness matrix when nonlinear analysis is performed. Among others, Yoon and Kim [2005b] developed a method, called the element connectivity parametrization method (ECP) to avoid the loss of the positive definiteness of the system matrix. The concept of the ECP method may be briefly illustrated with Figure 2; see [Yoon and Kim 2005a; [2005b]; Langelaar et al. 2005; Yoon et al. 2007; 2008; Yoon 2010] for more details. In the ECP method, elements are connected through one-dimensional zero-length links having variable stiffness. Because the ECP method represents an element of low density by a solid element connected through one-dimensional low-stiffness links, the loss of the tangent stiffness of low-density continuum elements can be avoided. To avoid the aforementioned numerical instability, therefore, we will employ the ECP method for the present problem tracing a geometrically nonlinear load-displacement curve.
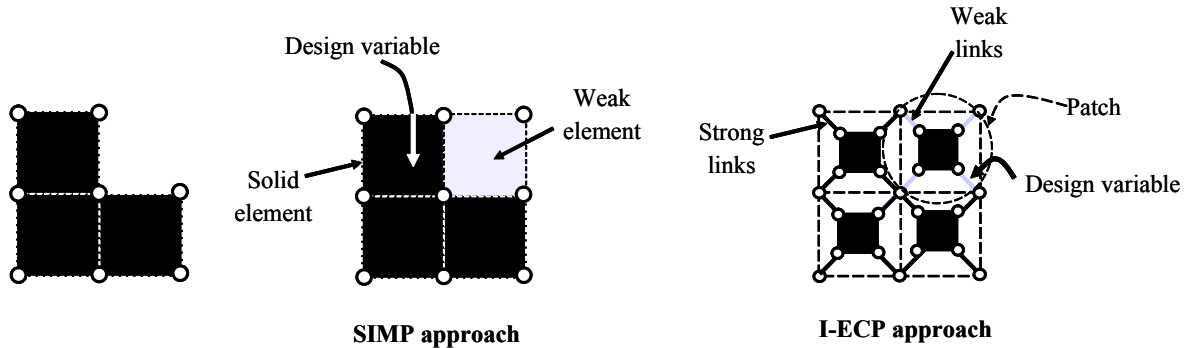
Figure 2. A structure (left) modeled with the SIMP (solid isotropic with penalization) method (middle) and the ECP-based method (right). All links have zero length.

Another important issue, especially in problems involving nonlinear structural responses, is related to intermediate-density elements appearing at the converged state because the postprocessing of intermediate elements to either solid or void elements can produce considerably different load-displacement curves. Although an implicit (and/or explicit) penalization technique to push discretizing elements to either void or solid states can be used, any topology optimization formulation using the notion of the density concept cannot completely suppress intermediate density-elements at the convergence state. (Integer-type topology optimization is not considered here because of impractically high computation cost.) For this reason, a special attention will be paid to the issue related to intermediate densities.

For the numerical analysis of geometrically nonlinear structures here, the displacement-controlled Newton–Raphson method will be used [Bathe 1996; Cho and Jung 2003; Huang and Xie 2008]. If a force control procedure is implemented, it is difficult to simulate complicated load-displacement trajectories such as snap-through or snap-back phenomena [Bathe 1996; Sekimoto and Noguchi 2001; Huang and Xie 2008]. Alternatively, one can use algorithms such as the arc-length method or the Broyden–Fletcher–Goldfarb–Shanno (BFGS) method (see [Bathe 1996], for example) when snap-back or snap-through phenomena need to be simulated. Since a monotonic softening or hardening load-displacement trajectory and a simple snap-through example are of the main interests in this study, the displacement-controlled Newton–Raphson method would be satisfactory here.

This paper is organized as follows. Section 2 explains the underlying governing equations for nonlinear analysis and the concept of the ECP modeling and Section 3 presents the topology optimization formulation and sensitivities analyses needed to solve load-displacement tracing problems. In Section 4, several numerical case studies are presented. Finally, the findings of this research are summarized and future research work is discussed.

2. Underlying equations and the ECP method

Before presenting an optimization formulation, this section describes underlying equations and notations needed for the finite element analysis and the ECP formulation.

2.1. *Geometrical nonlinear analysis.* For the analysis of geometrical nonlinear structures, the following Green–Lagrangian strain ($^{t+\Delta t}_0 \varepsilon_{ij}$) and the associated second PK (Piola–Kirchhoff) stress are used:

$$^{t+\Delta t}_0 \varepsilon_{ij} = \tfrac{1}{2}\left(^{t+\Delta t}_0 u_{i,j} + {}^{t+\Delta t}_0 u_{j,i} + {}^{t+\Delta t}_0 u_{l,i} \, {}^{t+\Delta t}_0 u_{l,j} \right) \quad (i = 1, 2), \tag{1}$$

where $^{t+\Delta t}_0 u_i$ is the displacement component in the x_i coordinate at time $t + \Delta t$ where the left subscript 0 indicates that the displacement is measured in the undeformed configuration. The comma denotes partial differentiation with respect to coordinate x_i. The repeated index implies summation. By denoting the displacement increment nodal vector and the displacement nodal vector at time $t + \Delta t$ of a generic point of a body occupying a domain V in equilibrium by ΔU and $^{t+\Delta t}U$, respectively, the following update rules for the Newton–Raphson method can be used (see [Bathe 1996; Cook et al. 2001], for instance):

$$\Re = {}^{t+\Delta t}F_{\text{Ext}} - {}^{t+\Delta t}F_{\text{Int}} = {}^{t+\Delta t}F_{\text{Ext}} - \int_V B^{\mathrm{T}} S \, dV = 0 \quad (S = CE), \tag{2}$$

$$^t K_{\mathrm{T}} = \frac{\partial \Re}{\partial^{t+\Delta t}U}, \tag{3}$$

$$^{t+\Delta t}U^{(k)} = {}^{t+\Delta t}U^{(k-1)} + \Delta U^{(k)}, \quad {}^{t+\Delta t}U^{(0)} = {}^t U, \tag{4}$$

$$^t K_{\mathrm{T}}^{(k-1)} \Delta U^{(k)} = \Re({}^{t+\Delta t}U^{(k-1)}), \tag{5}$$

where superscript k denotes the iteration step in the implemented Newton–Raphson method. The residual vector and the tangent stiffness matrix are denoted by $\Re$ and $^t K_{\mathrm{T}}$, respectively. The symbols $^{t+\Delta t}F_{\text{Ext}}$ and $^{t+\Delta t}F_{\text{Int}}$ denote the external load and internal load, respectively. The vector expressions of the second PK stress and the Green-strain are given by S and E and they are assumed to be related through a linear constitutive matrix C. The strain-displacement matrix is denoted by B.

2.2. *Geometrical nonlinear analysis for the ECP method.* In representing a layout in the framework of the ECP method, elements are not directly connected but through one-dimensional zero-length links with varying stiffness values. Therefore, the total number of nodes by ECP for two- or three-dimensional problems can be increased approximately by 5 or 9 times compared to those by the standard density approach [Yoon et al. 2007]. However, one can reduce the computation time and the size of the assembled system matrix by employing the patch model of [Yoon 2010; Yoon et al. 2007] (Figure 3) and using static condensation. Referring to Figure 3, the degrees of freedom of the inner nodes of the ECP patch can be statically condensed out so that only the degrees of freedom of the outer nodes can be kept in the system matrix. (For sample code in Matlab of the patch-model based ECP method, see [Yoon et al. 2008]).

Although only the degrees of freedom of the outer nodes will appear in the final system matrix, the displacement $u_{e,\text{in}}^{(k)}$ of the inner nodes and the displacement $u_{e,\text{out}}^{(k)}$ of the outer nodes should be simultaneously updated as

$$\begin{bmatrix} ^{t+\Delta t}u_{e,\text{out}}^{(k)} \\ ^{t+\Delta t}u_{e,\text{in}}^{(k)} \end{bmatrix} = \begin{bmatrix} ^{t+\Delta t}u_{e,\text{out}}^{(k-1)} \\ ^{t+\Delta t}u_{e,\text{in}}^{(k-1)} \end{bmatrix} + \begin{bmatrix} \Delta u_{e,\text{out}}^{(k)} \\ \Delta u_{e,\text{in}}^{(k)} \end{bmatrix} \tag{6}$$

The updated displacement increments for the outer nodes and the inner nodes are denoted by $\Delta \boldsymbol{u}_{e,\text{out}}^{(k)}$ and $\Delta \boldsymbol{u}_{e,\text{in}}^{(k)}$, respectively. The equation to find the increments is

$$\left\{ \begin{bmatrix} \boldsymbol{k}_{I,e} & -\boldsymbol{k}_{I,e} \\ -\boldsymbol{k}_{I,e} & \boldsymbol{k}_{I,e} \end{bmatrix} + \begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & {}^{t}\boldsymbol{k}_{\text{T},e}^{\text{struct},(k-1)} \end{bmatrix} \right\} \begin{bmatrix} \Delta \boldsymbol{u}_{e,\text{out}}^{(k)} \\ \Delta \boldsymbol{u}_{e,\text{in}}^{(k)} \end{bmatrix} = \begin{bmatrix} \mathfrak{R}_{e,\text{out}}^{(k-1)} \\ \mathfrak{R}_{e,\text{in}}^{(k-1)} \end{bmatrix}, \tag{7}$$

with

$$\boldsymbol{k}_{I,e} = l_e(\gamma_e) \boldsymbol{I}_{8\times 8}, \tag{8}$$

where $\boldsymbol{0}$ and $\boldsymbol{I}_{8\times 8}$ are the 8×8 zero matrix and identity matrix (in the two-dimensional case).

The link stiffness of the e-th patch, l_e, is a function of the design variable. The stiffness matrix, the residual force terms of the outer and the inner nodes of the e-th patch are denoted by ${}^{t}\boldsymbol{k}_{\text{T},e}^{\text{struct},(k-1)}$, $\mathfrak{R}_{e,\text{out}}^{(k-1)}$ and $\mathfrak{R}_{e,\text{in}}^{(k-1)}$, respectively. The force terms $\mathfrak{R}_{e,\text{out}}^{(k-1)}$ and $\mathfrak{R}_{e,\text{in}}^{(k-1)}$ are calculated as

$$\begin{bmatrix} \mathfrak{R}_{e,\text{out}}^{(k-1)} \\ \mathfrak{R}_{e,\text{in}}^{(k-1)} \end{bmatrix} = \begin{bmatrix} {}^{t+\Delta t}\boldsymbol{R}_e \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ {}^{t+\Delta t}_{0}\boldsymbol{f}_e^{\text{struct},(k-1)} \end{bmatrix} - \begin{bmatrix} {}^{t+\Delta t}_{0}\boldsymbol{f}_{e,\text{out}}^{\text{link},(k-1)} \\ {}^{t+\Delta t}_{0}\boldsymbol{f}_{e,\text{in}}^{\text{link},(k-1)} \end{bmatrix}, \tag{9}$$

$$\begin{bmatrix} {}^{t+\Delta t}_{0}\boldsymbol{f}_{e,\text{out}}^{\text{link},(k-1)} \\ {}^{t+\Delta t}_{0}\boldsymbol{f}_{e,\text{in}}^{\text{link},(k-1)} \end{bmatrix} = \begin{bmatrix} \boldsymbol{k}_{I,e} & -\boldsymbol{k}_{I,e} \\ -\boldsymbol{k}_{I,e} & \boldsymbol{k}_{I,e} \end{bmatrix} \begin{bmatrix} {}^{t+\Delta t}\boldsymbol{u}_{e,\text{out}}^{(k-1)} \\ {}^{t+\Delta t}\boldsymbol{u}_{e,\text{in}}^{(k-1)} \end{bmatrix}. \tag{10}$$

In (9), the externally applied force on the outer nodes is denoted by ${}^{t+\Delta t}\boldsymbol{R}_e$ and the internal force acting on the inner nodes by ${}^{t+\Delta t}_{0}\boldsymbol{f}_e^{\text{struct},(k-1)}$. By applying the static condensation strategy, the degrees related to the inner nodes are condensed out, leading to

$$ {}^{t}\boldsymbol{k}_{\text{Con},e}^{(k-1)} \Delta \boldsymbol{u}_{e,\text{out}}^{(k)} = \mathfrak{R}_{e,\text{out}}^{(k-1)} + \boldsymbol{k}_{I,e} (\boldsymbol{k}_{I,e} + {}^{t}\boldsymbol{k}_{\text{T},e}^{\text{struct},(k-1)})^{-1} \mathfrak{R}_{e,\text{in}}^{(k-1)}, \tag{11}$$

where

$$ {}^{t}\boldsymbol{k}_{\text{Con},e}^{(k-1)} = \left(\boldsymbol{k}_{I,e} - \boldsymbol{k}_{I,e} (\boldsymbol{k}_{I,e} + {}^{t+\Delta t}\boldsymbol{k}_{\text{T},e}^{\text{struct},(k-1)})^{-1} \boldsymbol{k}_{I,e} \right). \tag{12}$$
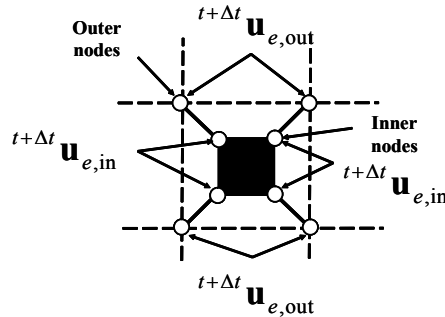
Figure 3. The e-th patch surrounding the e-th element for the ECP method. The inner nodes are the nodes defining the element while the outer nodes defining the e-th patch are used to connect patches sharing the same outer nodes. The solid lines connecting the inner and outer nodes denote zero-length one-dimensional elastic links.

Then the global tangent matrix is assembled as

$${}^{t}\boldsymbol{K}_{\text{Con}}^{(k-1)} = \sum_{e=1}^{N_p} {}^{t}\boldsymbol{k}_{\text{Con},e}^{(k-1)}, \tag{13}$$

where N_p is the total number of patches. Finally, the following system of equations is solved iteratively by a standard nonlinear solver such as the Newton–Raphson method or the arc-length method:

$${}^{t}\boldsymbol{K}_{\text{Con}}^{(k-1)} \Delta \boldsymbol{U}_{\text{out}}^{(k)} = \mathfrak{R}_{\text{Con}}^{(k-1)} \tag{14}$$

where

$$\mathfrak{R}_{\text{Con}}^{(k-1)} = \sum_{e=1}^{N_p} \left(\mathfrak{R}_{e,\text{out}}^{(k-1)} + \boldsymbol{k}_{I,e} (\boldsymbol{k}_{I,e} + {}^{t}\boldsymbol{k}_{\text{T},e}^{\text{struct},(k-1)})^{-1} \mathfrak{R}_{e,\text{in}}^{(k-1)} \right) \tag{15}$$

For efficient topology optimization procedures, it is crucial to use an appropriate interpolation function for the ECP method. In this paper, the following interpolation function is adopted for the link stiffness in (8):

$$\text{Solid: } l_e = l_{\max}, \quad \gamma_e = 1, \tag{16}$$

$$\text{Void: } l_e = l_{\min}, \quad \gamma_e = 0.001, \tag{17}$$

$$l_e = \alpha \frac{\gamma_e^n}{1 + (1 - \gamma_e^n)\tau} + \beta \left(\tau = \frac{\alpha \times s}{k_{\text{diagonal}}^{\text{struct}} \times k} \right), \tag{18}$$

$$\alpha = l_{\max} - l_{\min}, \quad \beta = l_{\min}, \tag{19}$$

$$\gamma_{\min} \leq \gamma_e \leq 1, \quad \gamma_{\min} = 0.001, \tag{20}$$

where k is the number of degrees of freedom per node and s and n are penalty parameters. A diagonal term of the linear stiffness matrix is denoted by $k_{\text{diagonal}}^{\text{struct}}$. The upper and lower bounds of the stiffness of links are denoted by $l_{\max}$ and $l_{\min}$, respectively. Here, $l_{\max}$ and $l_{\min}$ are set to $10^6 \times k_{\text{diagonal}}^{\text{struct}}$ and $10^{-6} \times k_{\text{diagonal}}^{\text{struct}}$, respectively. (See [Yoon et al. 2007; Yoon et al. 2008; Yoon 2010] for more details on the interpolation function adopted above and actual implementation on Matlab.)

3. Topology optimization formulation

3.1. *Optimization formulation.* To find a structural layout tracing a given load-displacement trajectory, the following optimization formulation is considered, where nl is the number of loadsteps and N_p is the number of discretizing patches:

$$\operatorname*{Min}_{\gamma} \Phi = \sum_{i=1}^{nl} \left\| \boldsymbol{L}_d^{\text{T}} {}^{t+\Delta t}\boldsymbol{U}^i - {}^{t+\Delta t}\boldsymbol{U}^{i,\text{ref}} \right\| \tag{21}$$

$$\text{s.t.} \quad \sum_{e=1}^{N_p} \rho_e(\gamma_e) v_e \leq V^*, \tag{22}$$

with $\gamma = \{\gamma_1, \gamma_2, \ldots, \gamma_{N_p}\}^T$. Here Φ is a scalar objective function measuring the sum of the differences between the actual displacements $^{t+\Delta t}U^i$ and the target displacement $^{t+\Delta t}U^{i,\text{ref}}$ for nl steps for a given load. The symbol L_d denotes a vector which has unity at the node of the displacement measurement point and all zeros for the rest. The density and element volume of the e-th patch are denoted by ρ_e and v_e, respectively. The prescribed volume limit is V^*. Instead of the objective function in (21), one may prescribe a displacement path and calculate the reaction force as the response. In this case, the objective function Φ is defined as

$$\Phi = \sum_{i=1}^{nl} \left\| L_r^{\text{T}} \, ^{t+\Delta t}F_{\text{Int}}^i - \, ^{t+\Delta t}R^{i,\text{ref}} \right\|, \tag{23}$$

where $^{t+\Delta t}R^{i,\text{ref}}$ and $^{t+\Delta t}F_{\text{Int}}^i$ are the target reaction force and the internal force vector at the i-th load step of the Newton–Raphson iteration. The symbol L_r is a vector which has unity at the node of the reaction force measurement point and all zeros for the rest.

The objective function in (23) was used by [Sekimoto and Noguchi 2001] to find optimal layouts satisfying a prescribed load-displacement trajectory. Theoretically, the use of either (21) or (23) would yield the same result (layout) if the same load-displacement trajectory is pursued. However, when the load-displacement trajectory becomes complex, the displacement control procedure is known to perform better in convergence. Although it can fail for a complex trajectory such as those involving snap-backs and/or snap-through phenomena, the displacement control procedure usually works better than the load control procedure using the objective function of (23) [Bathe 1996; Hellweg and Crisfield 1998; Liu and Nocedal 1989; Vila et al. 1997; Cook et al. 2001]. So it is implemented in the Newton–Raphson nonlinear solver for this investigation.

Although the functional form of (23) is used, there might be alternatives in selecting the specific norm. For instance, one can consider the Euclidean norm (2-norm)

$$\Phi = \sum_{i=1}^{nl} (L_r^{\text{T}} \, ^{t+\Delta t}F_{\text{Int}}^i - \, ^{t+\Delta t}R^{i,\text{ref}})^2, \tag{24}$$

the p-norm

$$\Phi = \left(\sum_{i=1}^{nl} (L_r^{\text{T}\,t+\Delta t}F_{\text{Int}}^i - \, ^{t+\Delta t}R^{i,\text{ref}})^p \right)^{1/p} \quad (p = 2, 4, 6, \ldots), \tag{25}$$

or the maximum norm

$$\Phi = \max\left(\left| L_r^{\text{T}} \, ^{t+\Delta t}F_{\text{Int}}^i - \, ^{t+\Delta t}R^{i,\text{ref}} \right| \right). \tag{26}$$

Obviously, the optimization convergence can be affected by the choice of the norm. The use of the Euclidean norm is to minimize the sum of every distance between the response reaction force and the reference (target) force for all incremental steps. On the other hand, the maximum norm minimizes the maximum distance. For example, let us consider an intermediate design for which the reaction follows the trajectory depicted in Figure 4. With the Euclidean norm, the average square of distances (the area marked by A) is to be minimized. With the p-norm with a sufficient large p or the maximum norm, the distance at load step 3 is minimized because the distance of the load step 3 (the vertical distance marked
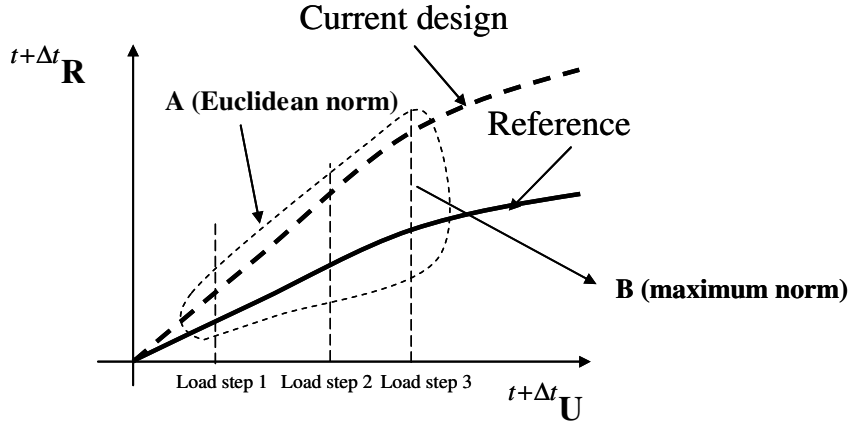
Figure 4. A typical load-displacement trajectory at an intermediate design step.
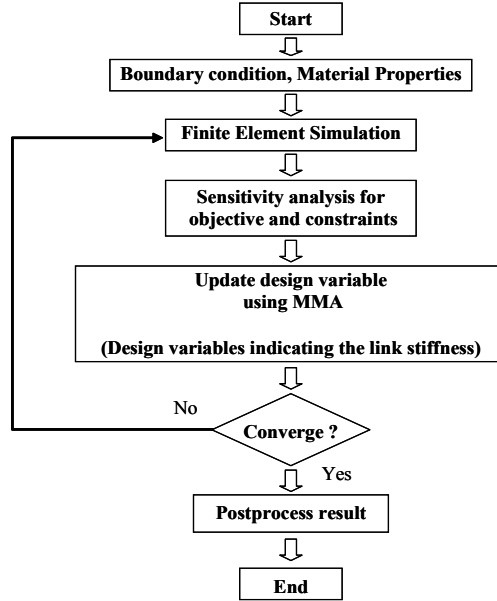
Figure 5. Topology design optimization procedure.

by B) is the largest. It is often useful to normalize the aforementioned norms as

$$\Phi = \sum_{i=1}^{nl} \left(\frac{\boldsymbol{L}_r^{\mathrm{T}\ t+\Delta t}\boldsymbol{F}_{\mathrm{Int}}^i}{t+\Delta t}\boldsymbol{R}^{i,\mathrm{ref}} -1 \right)^2, \quad \Phi = \left(\sum_{i=1}^{nl} \left(\frac{\boldsymbol{L}_r^{\mathrm{T}\ t+\Delta t}\boldsymbol{F}_{\mathrm{Int}}^i}{t+\Delta t}\boldsymbol{R}^{i,\mathrm{ref}} -1 \right)^p \right)^{1/p}, \quad \Phi = \max \left(\left| \frac{\boldsymbol{L}_r^{\mathrm{T}\ t+\Delta t}\boldsymbol{F}_{\mathrm{Int}}^i}{t+\Delta t}\boldsymbol{R}^{i,\mathrm{ref}} -1 \right| \right). \quad (27)$$

In general, it is difficult to judge in advance which norm would perform better for a given problem. Therefore, one may test various norms as given by (24)–(27) before finding optimal solutions. Figure 5 depicts an implemented topology optimization procedure to find a structure tracing a specific load-displacement trajectory.

3.2. Sensitivity analysis. To update the design variables ($\gamma = \{\gamma_1, \gamma_2, \ldots, \gamma_{Np-1}, \gamma_{Np}\}^T$), the optimization procedure as described in Figure 5 will be employed. To obtain the sensitivity of the objective function Φ with respect to the design variable, γ, it is necessary to derive the sensitivities of the displacement or the reaction force depending on the form of the selected objective function. Although the procedure to calculate the sensitivity for the present ECP method may be similar to that based on the element density method, it is worth giving it in some detail. For analysis, we assume that the size of the degrees of freedom of displacement is n_G and the displacement conditions (condensed degrees of freedom) are applied at $^{t+\Delta t}U_{kn}$ where $^{t+\Delta t}U_{un}$ denotes the unknown displacements (to be determined by a nonlinear solver). So the following definitions are introduced for the sensitivity analysis:

$$\Psi_d = L_d^{\mathrm{T}} {}^{t+\Delta t}U, \quad \Psi_r = L_r^{\mathrm{T}} {}^{t+\Delta t}F_{int,kn},$$

where L_d and L_r are $n_G \times 1$ and $n_{kn} \times 1$ vectors for the temporarily objective functions Ψ_d and Ψ_r of the displacement and the reaction force, respectively. By separating the known displacements from unknown displacements, one can express the governing equation by using the residual vector:

$$\mathfrak{R} = {}^{t+\Delta t}F_{\mathrm{Ext}} - {}^{t+\Delta t}F_{\mathrm{Int}} = \begin{bmatrix} {}^{t+\Delta t}F_{\mathrm{Ext},un} - {}^{t+\Delta t}F_{\mathrm{Int},un} \\ {}^{t+\Delta t}F_{\mathrm{Ext},kn} - {}^{t+\Delta t}F_{\mathrm{Int},kn} \end{bmatrix} = \begin{bmatrix} 0_{un \times 1} \\ 0_{kn \times 1} \end{bmatrix} \tag{28}$$

Because the boundary conditions are specified on $^{t+\Delta t}U_{kn}$, the second row of the above equation can be removed. Without the loss of generality, the load step index appearing as a right superscript is omitted.

3.3. Sensitivity of the displacement. The sensitivity of Ψ_d with respect to the design variable γ requires the sensitivity of the displacement $^{t+\Delta t}U$:

$$\Psi_d' = L_d^{\mathrm{T}} {}^{t+\Delta t}U' \tag{29}$$

where differentiation is with respect to γ. To calculate $^{t+\Delta t}U'$, the equilibrium equation, $\mathfrak{R} = 0_{n_G \times 1}$, is differentiated:

$$\frac{\partial \mathfrak{R}}{\partial {}^{t+\Delta t}U} \frac{\partial {}^{t+\Delta t}U}{\partial \gamma} + \mathfrak{R}' = 0_{n_G \times Np}. \tag{30}$$

If an adjoint variable λ is solved from

$$\frac{\partial \mathfrak{R}}{\partial {}^{t+\Delta t}U} \lambda = L_d, \tag{31}$$

the sensitivity Ψ_d' can be easily found from

$$\Psi_d' = -\lambda^{\mathrm{T}} \mathfrak{R}'. \tag{32}$$

3.4. Sensitivity of the reaction force. The sensitivity of Ψ_r with respect to the design variable γ can be obtained by a technique similar to that used to calculate Ψ_d'. For the displacement-controlled case, it is convenient to put $^{t+\Delta t}U$ as

$$^{t+\Delta t}U = [{}^{t+\Delta t}U_{un}, {}^{t+\Delta t}U_{kn}]^{\mathrm{T}} \tag{33}$$

In this case, the residual vector $\mathfrak{R}$ can be also put as

$$\mathfrak{R} = \begin{bmatrix} {}^{t+\Delta t}\boldsymbol{F}_{\text{Ext},un} \\ \boldsymbol{0}_{kn\times 1} \end{bmatrix} - \begin{bmatrix} {}^{t+\Delta t}\boldsymbol{F}_{\text{Int},un} \\ {}^{t+\Delta t}\boldsymbol{F}_{\text{Int},kn} \end{bmatrix} = \begin{bmatrix} \boldsymbol{0}_{un\times 1} \\ \boldsymbol{0}_{kn\times 1} \end{bmatrix}, \qquad \frac{\partial^{t+\Delta t}\boldsymbol{F}_{\text{Int}}}{\partial^{t+\Delta t}\boldsymbol{U}} = \begin{bmatrix} \dfrac{\partial^{t+\Delta t}\boldsymbol{F}_{\text{Int},un}}{\partial^{t+\Delta t}\boldsymbol{U}_{un}} & \dfrac{\partial^{t+\Delta t}\boldsymbol{F}_{\text{Int},un}}{\partial^{t+\Delta t}\boldsymbol{U}_{kn}} \\ \dfrac{\partial^{t+\Delta t}\boldsymbol{F}_{\text{Int},kn}}{\partial^{t+\Delta t}\boldsymbol{U}_{un}} & \dfrac{\partial^{t+\Delta t}\boldsymbol{F}_{\text{Int},kn}}{\partial^{t+\Delta t}\boldsymbol{U}_{kn}} \end{bmatrix}. \tag{34}$$

First, we differentiate Ψ_r:

$$\frac{d\Psi_r}{d\gamma} = \boldsymbol{L}_r^{\text{T}} \left(\frac{\partial^{t+\Delta t}\boldsymbol{F}_{\text{Int},kn}}{\partial^{t+\Delta t}\boldsymbol{U}_{un}} \frac{\partial^{t+\Delta t}\boldsymbol{U}_{un}}{\partial\gamma} + \frac{\partial^{t+\Delta t}\boldsymbol{F}_{\text{Int},kn}}{\partial^{t+\Delta t}\boldsymbol{U}_{kn}} \frac{\partial^{t+\Delta t}\boldsymbol{U}_{kn}}{\partial\gamma} + \frac{\partial^{t+\Delta t}\boldsymbol{F}_{\text{Int},kn}}{\partial\gamma} \right). \tag{35}$$

Because ${}^{t+\Delta t}\boldsymbol{U}_{un} = {}^{t+\Delta t}\boldsymbol{U}^*$ (condensed degrees of freedom), the second term in parenthesis in (35) vanishes: $\partial^{t+\Delta t}\boldsymbol{U}_{kn}/\partial\gamma = \boldsymbol{0}_{kn\times Np}$. To calculate the first term, the equilibrium equation (34) is differentiated:

$$\frac{\partial^{t+\Delta t}\boldsymbol{F}_{\text{Ext},un}}{\partial\gamma} - \frac{\partial^{t+\Delta t}\boldsymbol{F}_{\text{Int},un}}{\partial^{t+\Delta t}\boldsymbol{U}_{un}} \frac{\partial^{t+\Delta t}\boldsymbol{U}_{un}}{\partial\gamma} - \frac{\partial^{t+\Delta t}\boldsymbol{F}_{\text{Int},un}}{\partial\gamma} = \boldsymbol{0}_{un\times Np}. \tag{36}$$

By inserting (36) into (35), we obtain

$$\frac{\partial\Psi_r}{\partial\gamma} = \boldsymbol{L}_r^{\text{T}} \left(\frac{\partial^{t+\Delta t}\boldsymbol{F}_{\text{Int},kn}}{\partial^{t+\Delta t}\boldsymbol{U}_{un}} \left(\left(\frac{\partial^{t+\Delta t}\boldsymbol{F}_{\text{Int},un}}{\partial^{t+\Delta t}\boldsymbol{U}_{un}} \right)^{-1} \left(\frac{\partial^{t+\Delta t}\boldsymbol{F}_{\text{Ext},un}}{\partial\gamma} - \frac{\partial^{t+\Delta t}\boldsymbol{F}_{\text{Int},un}}{\partial\gamma} \right) \right) + \frac{\partial^{t+\Delta t}\boldsymbol{F}_{\text{Int},kn}}{\partial\gamma} \right)$$

$$= \boldsymbol{L}_r^{\text{T}} \begin{bmatrix} -\dfrac{\partial^{t+\Delta t}\boldsymbol{F}_{\text{Int},kn}}{\partial^{t+\Delta t}\boldsymbol{U}_{un}} \left(\dfrac{\partial^{t+\Delta t}\boldsymbol{F}_{\text{Int},un}}{\partial^{t+\Delta t}\boldsymbol{U}_{un}} \right)^{-1} & \boldsymbol{I}_{kn\times Np} \end{bmatrix} \begin{bmatrix} -\dfrac{\partial^{t+\Delta t}\boldsymbol{F}_{\text{Ext},un}}{\partial\gamma} + \dfrac{\partial^{t+\Delta t}\boldsymbol{F}_{\text{Int},un}}{\partial\gamma} \\ \dfrac{\partial^{t+\Delta t}\boldsymbol{F}_{\text{Int},kn}}{\partial\gamma} \end{bmatrix}. \tag{37}$$

We now introduce the adjoint variable $\boldsymbol{\lambda}$, with $\boldsymbol{\lambda}^{\text{T}} \equiv [\boldsymbol{\lambda}_{un} \ \boldsymbol{\lambda}_{kn}]^{\text{T}}$, where $\boldsymbol{\lambda}_{kn} = \boldsymbol{L}_r$, and simplify (37) to

$$\frac{\partial\Psi_r}{\partial\gamma} = \boldsymbol{\lambda}^{\text{T}} \begin{bmatrix} -\dfrac{\partial^{t+\Delta t}\boldsymbol{F}_{\text{Ext},un}}{\partial\gamma} + \dfrac{\partial^{t+\Delta t}\boldsymbol{F}_{\text{Int},un}}{\partial\gamma} \\ \dfrac{\partial^{t+\Delta t}\boldsymbol{F}_{\text{Int},kn}}{\partial\gamma} \end{bmatrix}, \tag{38}$$

$$\begin{bmatrix} \dfrac{\partial^{t+\Delta t}\boldsymbol{F}_{\text{Int},un}}{\partial^{t+\Delta t}\boldsymbol{U}_{un}} & \dfrac{\partial^{t+\Delta t}\boldsymbol{F}_{\text{Int},un}}{\partial^{t+\Delta t}\boldsymbol{U}_{kn}} \\ \dfrac{\partial^{t+\Delta t}\boldsymbol{F}_{\text{Int},kn}}{\partial^{t+\Delta t}\boldsymbol{U}_{un}} & \dfrac{\partial^{t+\Delta t}\boldsymbol{F}_{\text{Int},kn}}{\partial^{t+\Delta t}\boldsymbol{U}_{kn}} \end{bmatrix} \begin{bmatrix} \boldsymbol{\lambda}_{un} \\ \boldsymbol{\lambda}_{kn} \end{bmatrix} = \begin{bmatrix} \boldsymbol{0}_{un\times 1} \\ \boldsymbol{0}_{kn\times 1} \end{bmatrix}. \tag{39}$$

Since $\boldsymbol{\lambda}_{kn} = \boldsymbol{L}_r$, this leads to

$$\boldsymbol{\lambda} = \begin{bmatrix} -\left(\dfrac{\partial^{t+\Delta t}\boldsymbol{F}_{\text{Int},un}}{\partial^{t+\Delta t}\boldsymbol{U}_{un}} \right)^{-1} \dfrac{\partial^{t+\Delta t}\boldsymbol{F}_{\text{Int},kn}}{\partial^{t+\Delta t}\boldsymbol{U}_{kn}} \boldsymbol{L}_r \\ \boldsymbol{L}_r \end{bmatrix}. \tag{40}$$

The validity of the sensitivity analysis was checked against the direct numerical result by finite difference.

4. Examples of topology optimization

This paper considers problems defined only in two-dimensional domains, but the validity and efficiency of the ECP-based solution approach can be clearly demonstrated in the topology optimization tracing given load-displacement curves. It is noted that the developed ECP-based approach does not entail any special treatment or tuning for stabilizing possible instability in the Newton–Raphson iterations or in optimization processes. Because the ECP method is used here, all the remarks and findings are mainly applicable to an approach based on the ECP method, not on the element density method. One may see more detailed explanations of the advantages and drawbacks of the element density method for some nonlinear problems may be found in [Buhl et al. 2000; Bruns and Tortorelli 2003; Cho and Jung 2003; Yoon and Kim 2005b; Yoon et al. 2008]. Unless stated otherwise, uniform initial guesses of γ satisfying given mass constraints are used. The dimension, material properties and magnitude of the loads are so chosen as to demonstrate the potential of the present ECP based formulation. In this study, the method of moving asymptotes is used as an optimization algorithm [Svanberg 1987]; however other optimization algorithms can also be used.

Example 1: Controlling the load-displacement curve. As the first numerical example, we consider the problem depicted in Figure 6a. The design domain is discretized by 60×40 four-node quad elements. The displacement is prescribed at the top-right node while other boundary conditions are depicted in the figure. Young's modulus and Poisson's ratio are set to be $1 \, \text{N/m}^2$ and 0.4, respectively.

Before solving the optimization problem set up as (21) and (22), a nominal structure maximizing the reaction force for a specified displacement value where the structural response is found by using linear strains. The result is shown in Figure 6b, which should be the same as the result obtained by a standard compliance minimization using linear strains. To obtain the result in Figure 6b, the sensitivity filer of radius equal to 1.2 times the element size was used [Bendsøe and Sigmund 2003].

It will be interesting to investigate the load-displacement trajectory when the geometric nonlinearity using the Green–Lagrangian strain is used for the optimized layout of Figure 6b. The result is given in Figure 6c for which the maximum tip displacement $\delta_{\max}$ is set to be $0.12 \, \text{m}$. When the nonlinearity is considered, the drop in the reaction force is expected as the tip displacement (δ) reaches $\delta_{\max}$. This is due to a buckling phenomenon caused by the local bending indicated in Figure 6d.

Now let us try to find a structural layout following a prescribed load-displacement trajectory based on geometric nonlinear analysis. By using the trajectory in Figure 6c as a nominal trajectory, we set up a problem to find a structure that passes the following two points in the load-displacement trajectory:

$$\text{Target:} \begin{cases} {}^{t+\Delta t}\boldsymbol{R}^{1,\text{ref}} = 1.0 \times 10^{-3} \, \text{N at } \delta = 0.06 \, \text{m}, \\ {}^{t+\Delta t}\boldsymbol{R}^{2,\text{ref}} = 2.0 \times 10^{-3} \, \text{N at } \delta = 0.12 \, \text{m}. \end{cases} \tag{41}$$

The load-displacement relation required by (41) states that the target structure should not exhibit apparent weakening effects near $0.12 \, \text{m}$ for the input displacement because ${}^{t+\Delta t}\boldsymbol{R}^{2,\text{ref}}$ by (41) is larger than the corresponding value predicted in Figure 6c. In other words, the overall stiffness before $0.6 \, \text{m}$ of the input displacement should be smaller than that of the linear design but a sudden drop due to the geometric weakening in the reaction force observed at the linear design can be avoided at $\delta = 0.12 \, \text{m}$ if a structure satisfying (41) is indeed found.

By using the optimization formulation of (22) and (23) (with $V^* = 20\%$ as before), the optimized layouts tracing the target response of (41) are shown in Figure 7, left and middle. In defining (23), the Euclidean norm defined in (24) was used. The fixed filter implies the sensitivity filter using the fixed radius (r_{filter}) of 1.2 times the element size while the varying filter implies the sensitivity filter using the following filter radius (this can be viewed as a special case of the continuation method):

$$r_{\text{filter}} = \begin{cases} 1.2 \times \text{ element size } & \text{for loop} \leq 150 \text{ iterations,} \\ 1.0 \times \text{ element size } & \text{for loop} > 150 \text{ iterations.} \end{cases} \tag{42}$$

The main motivation to use the varying filter is to minimize the appearance of intermediate density elements because the existence of these elements can significantly affect structural response especially when nonlinear analysis is considered.
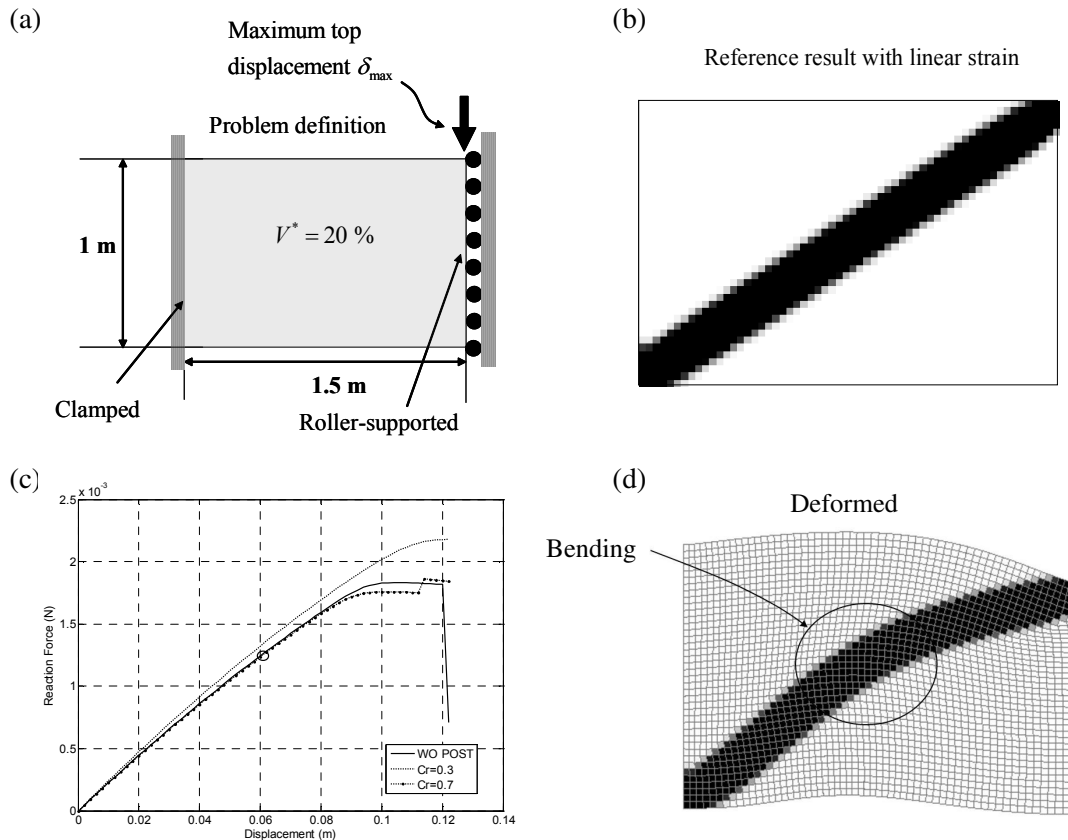
Figure 6. A test problem. (a) Problem definition (E=1 N/m^2, $\nu = 0.4$). (b) Optimized layout for $V^* = 20\%$ when linear analysis is used. (c) Load-displacement trajectory of the layout in (b) by geometrically nonlinear analysis; Cr is the threshold design variable value for hard-kill postprocessing. (d) Deformed shape by nonlinear analysis when the top displacement reaches 0.12 m.

To check the effects of postprocessing, for instance, Figure 6c compares the load-displacement curves before and after the hard-kill postprocessing by which the elements having design variables above or equal a threshold Cr are replaced by solid elements, and elements having the design variables lower than Cr by void elements. Two values of Cr were used: 0.3 and 0.7. As seen from Figure 6c, the differences among the responses can be neglected in the linear range but become significant in the nonlinear range. As a result, more effort should be made to avoid intermediate design values but the optimized layouts in existing works dealing with nonlinear load-displacement curve tracing seem to contain quite a large portion of intermediate variables. As shall be seen later with numerical results, the optimized layouts by the element connectivity method appear to have a smaller portion of intermediate variables.

Although the filter-size varying technique helped to reduce the occurrence of intermediate density elements, they were not completely removed. The load-displacement trajectories by the optimized layouts are plotted in Figure 7, right. Note that an additional member has appeared in the optimized layouts when the target trajectory is required to pass two reference points in (41). When δ exceeds $\delta_{\max} = 0.12$m, the reaction forces suddenly drop and the response pattern is significantly affected even by small differences in the layout configuration, which may be viewed as an imperfection.

Now let us find a structural layout to follow the load-displacement trajectories having slightly different reference points from those used to obtain the layouts in Figure 7. Among others, the value of the reaction force at $\delta = 0.06$ m is varied. The comparison of the obtained layouts in Figure 8a,b and that shown in Figure 8c indicates that the member staring at A should be more straight as a larger reaction force at $\delta = 0.06$ m is required. As a larger reaction force is required at $\delta = 0.06$ m, the reduction in the tangent stiffness of the obtained layout at $\delta = 0.12$ m is more significant. Another observation is that it may be not always possible to obtain a layout yielding a given load-displacement trajectory passing through arbitrarily specified load-displacement points because there may not exist such a layout. For instance, the trajectory marked by (a) in the graph of Figure 8 does not exactly pass through $({}^{t+\Delta t}\boldsymbol{R}^{2,\mathrm{ref}}, \delta) = (2.0 \times 10^{-3}\,\mathrm{N},\ 0.12\,\mathrm{m})$.

Furthermore, to prevent the sudden drop of the reaction force of the design in Figure 8c near $\delta = 0.144$ m, the one reference point at $\delta = 0.144$ m in Figure 9 is additionally inserted and the effect of this inclusion is tested. The changes of the details of the design of Figure 8c are observed in Figure 9.
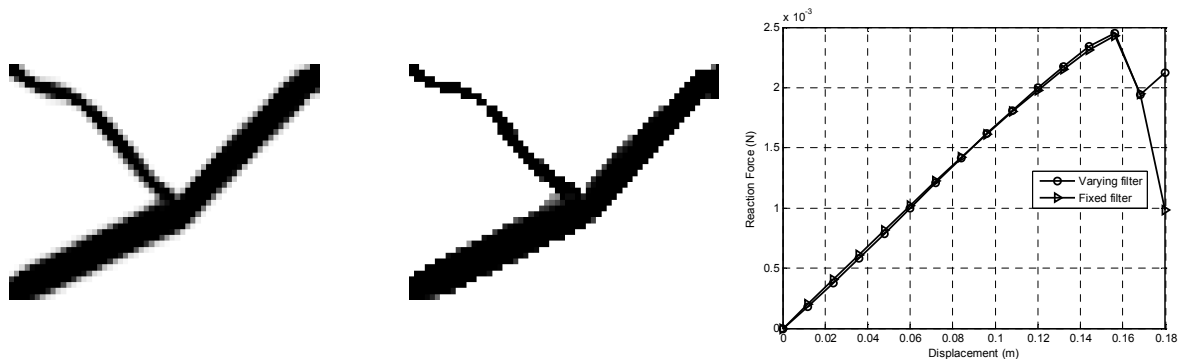
Figure 7. Layout optimization following the specific target reference points in the load-displacement trajectory: optimized layouts using fixed filter (left) and varying filter (middle); load-displacement trajectories for the optimized layouts (right).

Tracing $({}^{t+\Delta t}R^{1,ref},\delta)=(0.8\times10^{-3}\,N,0.06\,m)$
$({}^{t+\Delta t}R^{2,ref},\delta)=(2.0\times10^{-3}\,N,0.12\,m)$

Tracing $({}^{t+\Delta t}R^{1,ref},\delta)=(1.0\times10^{-3}\,N,0.06\,m)$
$({}^{t+\Delta t}R^{2,ref},\delta)=(2.0\times10^{-3}\,N,0.12\,m)$

Tracing $({}^{t+\Delta t}R^{1,ref},\delta)=(1.2\times10^{-3}\,N,0.06\,m)$
$({}^{t+\Delta t}R^{2,ref},\delta)=(2.0\times10^{-3}\,N,0.12\,m)$
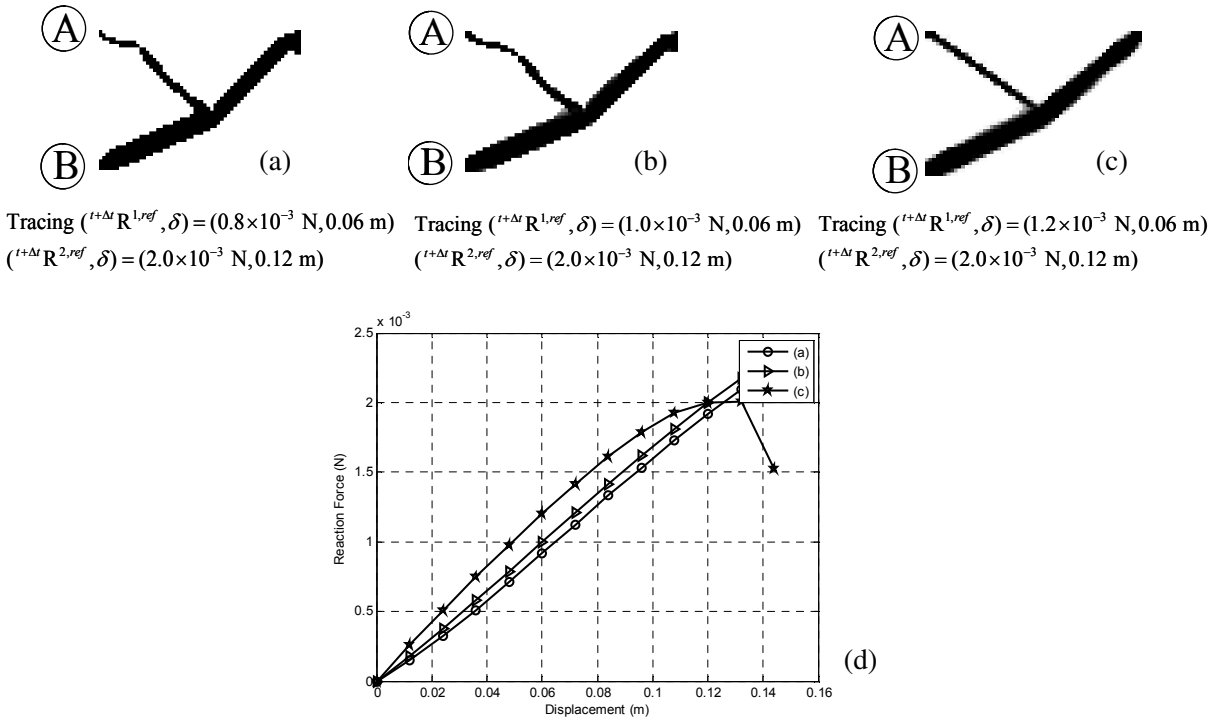
Figure 8. Top: optimized layouts following different target reference points in the load-displacement trajectory. Bottom: corresponding load-displacement trajectories.

The straight line near the point A becomes thicker and the straight line near the point C becomes narrow compared with the design in Figure 8c. This reveals that the ECP based topology optimization can change the details of the design to satisfy the given load and displacement curve. Furthermore, it proves that after the post processing of obtained layouts, the size or shape optimization process may be required to compensate the adverse effect of the postprocessing.

Tracing $({}^{t+\Delta t}R^{1,ref},\delta)=(1.2\times10^{-3}\,N,0.06\,m)$
$({}^{t+\Delta t}R^{2,ref},\delta)=(2.0\times10^{-3}\,N,0.12\,m)$
$({}^{t+\Delta t}R^{3,ref},\delta)=(2.0\times10^{-3}\,N,0.144\,m)$
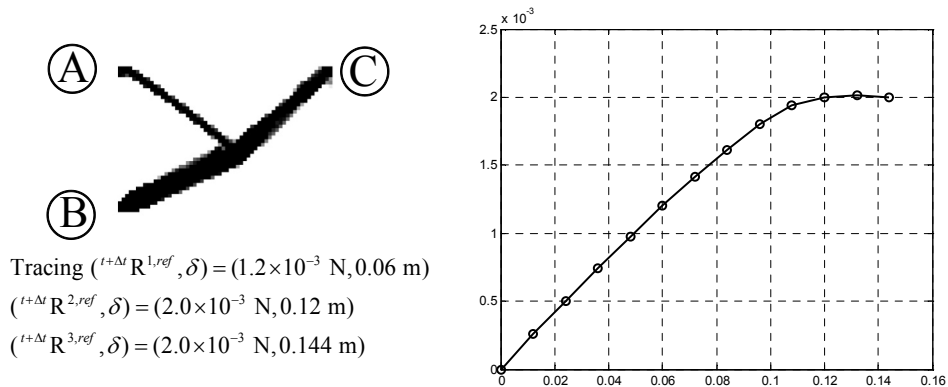
Figure 9. Optimization test by including an additional point for the design of Figure 8c: design (left) and load-displacement curve (right).
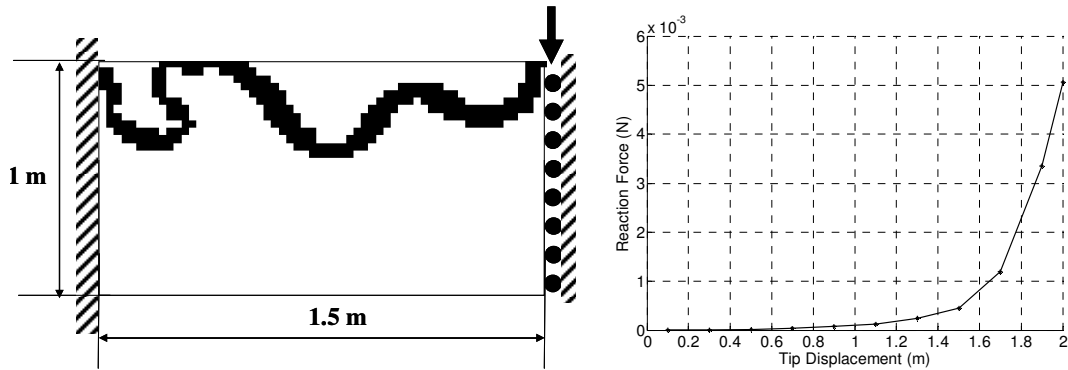
Figure 10. Left: reference layout. Right: load-displacement trajectory.

The normalized norm was also considered in the definition of the objective function. Although the optimized results are not shown, the layout configurations with and without the normalization were not much different. For this reason, the unnormalized Euclidean norm defined in (24) will be used in the subsequent examples.

Example 2: Design of a stiffening structure. To test if the proposed formulation can find a structural layout exhibiting geometrically stiffening behavior, a structure shown in Figure 10 is devised. When the structure is loaded at the tip marked by a vertical arrow, it exhibits geometrical stiffening. Note that the stiffening effect results from the constraint of the horizontal displacement along the roller support. The optimization problem considered is to recover the layout shown in Figure 10, left, when the curve in Figure 10, right, is given as the target trajectory. To solve this problem, the formulation given by (22) and (23) is adopted. Eleven points in the load-displacement curves marked by dots in Figure 10, right, are used. The intermediate and final layouts appearing during optimization iterations and the corresponding load-displacement trajectories are shown in Figure 11, which indicates that the proposed
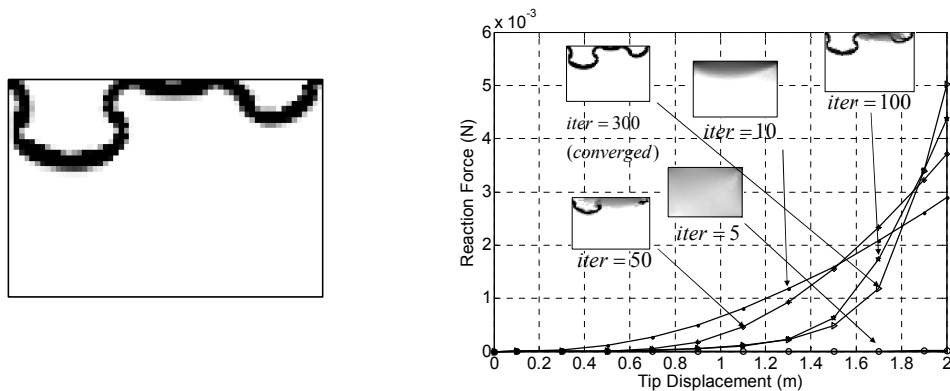
Figure 11. Optimized results by the developed method in recovering the reference layout exhibiting the nonlinear load-displacement curve in Figure 10. Left: the optimized layout. Right: load-displacement curves of the intermediate and final layouts.
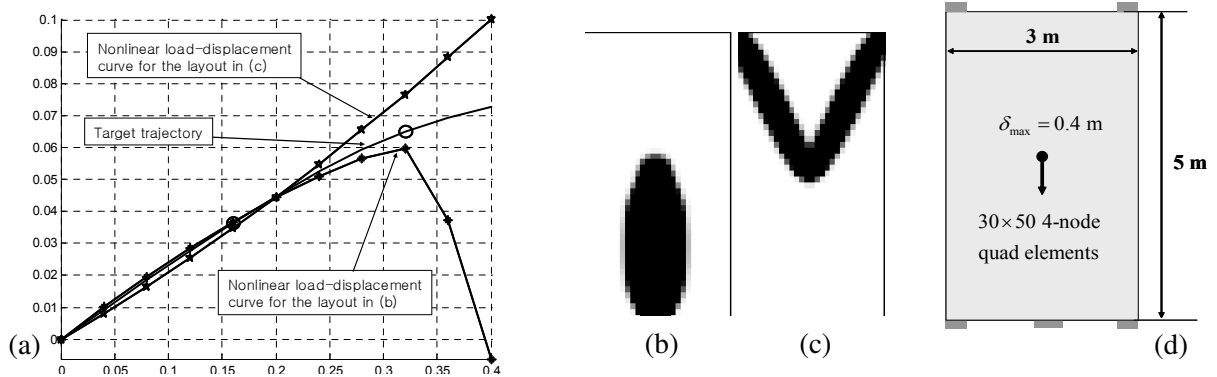
Figure 12. A problem to find a target nonlinear load-displacement trajectory. Left: target and limiting trajectories. Middle: reference layouts yielding linear and nonlinear limiting trajectories. Right: a compliance minimization problem used to obtain the given layouts with $E = 1\,\text{N/m}^2$, $\nu = 0.4$, volume $= 20\%$.

ECP-based formulation is capable of finding a layout very close to the one given in Figure 10, left, while the specified load-displacement curve is fairly accurately traced.

Example 3: Center-loaded structure with a specified structural response. Here, we look for a structural layout following a target load-displacement trajectory in Figure 12a, lying between two limiting trajectories. The limiting trajectories shown in the graph are the trajectories of the layouts (b) and (c) in Figure 12, optimized through compliance minimization (defined in Figure 12d) by using linear and geometrically nonlinear analyses. In obtaining the layouts, the center of the design domain was pulled down by $\delta_{\max} = 0.4\,\text{m}$ and a checkerboard filter of radius equal to 1.2 times the element size was used.

To solve this problem, $nl = 2$ was used and the specific reference points were $(^{t+\Delta t}\boldsymbol{R}^{1,\text{ref}}, \delta) = (0.0363\,\text{N}, 0.16\,\text{m})$ and $(^{t+\Delta t}\boldsymbol{R}^{2,\text{ref}}, \delta) = (0.065\,\text{N}, 0.32\,\text{m})$. Figure 13 shows the optimized result for this problem. The optimized layout (left) and the variation of the load-displacement trajectories are
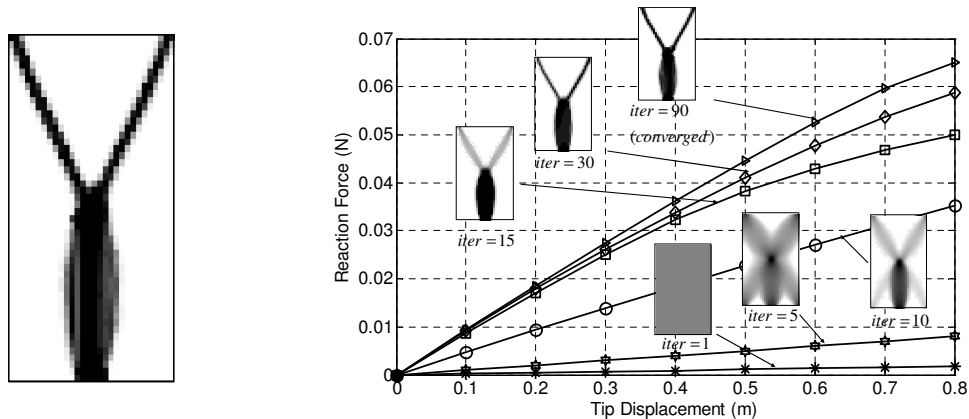
Figure 13. Iteration history of the load-displacement trajectories. The design objective is to find the target layout in Figure 12d.
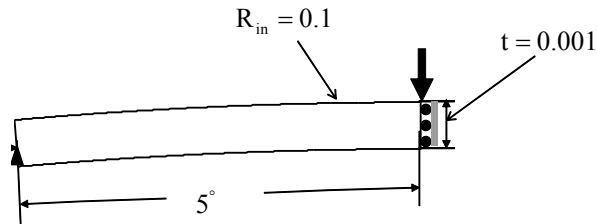
Figure 14. A problem to find a target snap-through trajectory ($E = 1\,\text{N/m}^2$, $v = 0.4$, volume $= 30\%$, input displacement $= 0.001$).

illustrated as the iteration proceeds. The target trajectory in Figure 12a is shown to be accurately traced by the trajectory of the optimized layout in Figure 13. Interestingly, the optimized layout in this case looks like a combination of the two layouts in parts (b) and (c) of Figure 12, with the size of every member roughly reduced by a factor of two. As remarked in earlier examples, it is difficult to obtain a completely distinct solid-void distribution without intermediate design variables.

Example 4: A design exhibiting snap-through. A problem involving snap-through is now considered. Since a relatively simple solver is employed, a structure without any snap-back is considered in Figure 14. The design domain is bounded by elliptical curves. If a slender symmetric structure is center-loaded, the design domain in Figure 14 can be viewed as its half. Then one can consider a roller-supported side which is loaded by a vertical displacement. The left side may be assumed to be fixed at its middle point. The radius and thickness of the ellipse are assumed to be 0.1 and 0.001, respectively. The domain is discretized by 120 by 20 elements and the allowed material usage is set to 30% of the total area. The input displacement is set to be 0.001 which is the same as the height of the design domain. This value would be sufficiently large to induce a snap-through phenomenon. The target points in the load-displacement curve for the topology optimization are marked in Figure 15, right. This curve is so selected as to be able to obtain a layout exhibiting snap-through. To obtain the optimized layout in Figure 15, left, the same ECP formulation as used in the previous examples is used. Solution convergence was stable. The trajectory by the optimized layout is marked by a solid line in Figure 15, right. This example demonstrates that the developed method can be used to solve problems involving snap-through phenomena. However, the present approach in its current form may not handle complex force-displacement curves without an advanced nonlinear solver.

Example 5: A column design exhibiting geometrically weakening. As the last case study, we consider an optimization problem to find a structure tracing the nonlinear load-displacement trajectory given in Figure 16, left. As shown by that figure, the target trajectory corresponds to the load-displacement curve of a slender column vertically loaded at the top left corner. Obviously, the solid straight column will be an optimized layout obtained as a solution to (22) and (23). To set up a nontrivial optimization problem, we take the design problem depicted in Figure 16, right; only the middle part of the column is assumed to be a design domain, but it is enlarged by 5 times. If the mass constraint of 20% is used, then the original straight column (middle diagram) is expected to be recovered. However, the actual optimized result is a rectangular box, quite different from the expected nominal rectangular solid. Seven points in
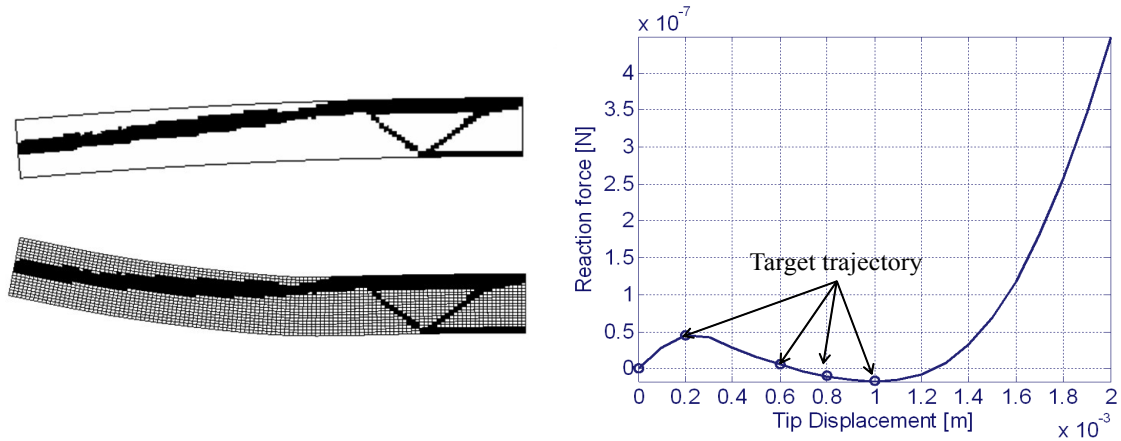
Figure 15. Example 4: a problem involving snap-through. Top left: optimized layout (undeformed). Bottom left: deformed shape (input displacement = 0.001). Right: load-displacement trajectory by the optimized layout; circles represent the target points.
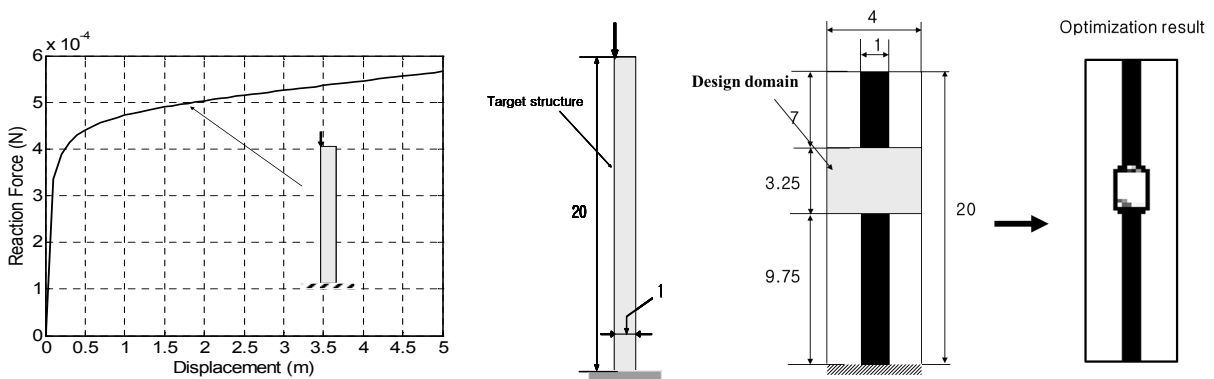
Figure 16. Example 5: a column design exhibiting geometrically weakening. Left: target trajectory. Middle: reference solid column structure ($E = 1 \, \text{N/m}^2$, $\nu = 0.3$). Right: an optimized layout in the expanded middle design domain.

the trajectory were used as the reference points to obtain the results in Figure 16, right, by the proposed topology optimization method. When a uniformly distributed density value of 0.2 (corresponding to 20%) was used, the solid column in the middle diagram was not recovered; one of local optima as given by the rightmost diagram was obtained. As Figure 17 shows, however, the optimized layout in Figure 16, right, obtained with 7 reference points, traces fairly closely the target trajectory. Figure 17 also shows the importance of using more reference points in order to trace the target trajectory accurately.

5. Conclusions

The structural topology problem to find optimal structural layouts exhibiting either geometrically hardening or weakening load-displacement trajectories was investigated by using the element connectivity
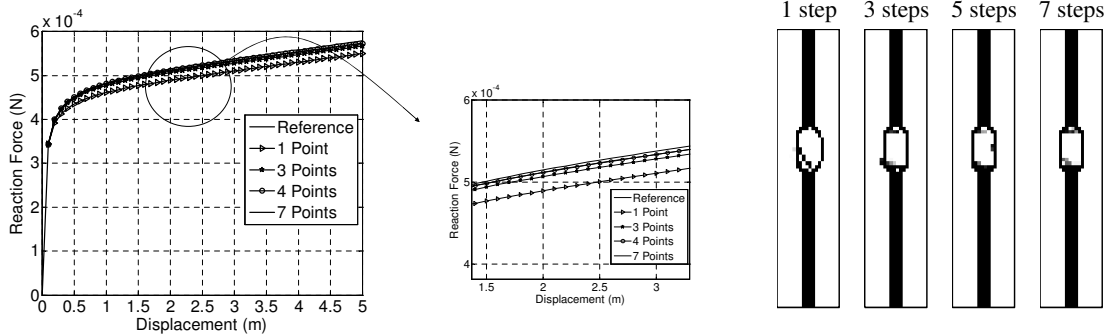
Figure 17. Comparison of the target load-displacement trajectory and the trajectories of the optimized structures obtained with varying numbers of reference points in the load-displacement trajectory. The layouts on the right correspond to, respectively, 1, 3, 4, and 7 points.

parametrization formulation. The following findings are made through this study. First, the ECP method effectively avoided the numerical instability or the erratic convergence related to low-density elements during the topology optimization tracing geometrically nonlinear load-displacement trajectories without any additional stabilization treatment in the Newton–Raphson scheme or prior tuning in optimization. Second, the issues of local optima and gray elements due to the design variable relaxation from the binary variables to the continuous variables appeared in the ECP based method, but the gray element problems appeared to be less severe than those with the element based approach. A load-displacement trajectory sufficiently close to the target one was found while the optimized layout was completely different from the nominal layout used to produce the target trajectory. Therefore, much attention should be paid in interpreting optimized results. In several cases, the postprocessed layouts yielded different load-displacement trajectories from those of the unprocessed layouts. Although a gray-element-free topology optimization formulation that can handle arbitrarily shaped load-displacement curve requires further research, the findings and numerical results obtained with the ECP formulation suggest that it can be very useful to advance the research in complete handling of general nonlinear problems.

Personal notes in memory of MarieLu Steele, by Yoon Young Kim

MarieLu was one of the most cheerful and lively persons I have met. She was always positive, and, I was revitalized every time I was in her company. I strongly believe that her positive energy enabled the dramatic success of the *International Journal of Solids and Structures* for which she was a dedicated Associate Editor for many years. Recently, her enthusiasm and devotion ensured the successful launching of the *Journal of Mechanics of Materials and Structures*, a new high-quality journal in the field of mechanics. I will dearly miss her energizing power and dedication to the mechanics community.

References

[Bathe 1996] K.-J. Bathe, *Finite element procedures*, Prentice Hall, Englewood Cliffs, NJ, 1996.

[Bendsøe and Kikuchi 1988] M. P. Bendsøe and N. Kikuchi, *Generating optimal topologies in structural design using a homogenization method*, Matematisk Institut, Danmarks Tekniske Højskole, Lyngby, 1988.

[Bendsøe and Sigmund 2003] M. P. Bendsøe and O. Sigmund, *Topology optimization: theory, methods and applications*, Springer, Berlin, 2003.

[Bruns and Sigmund 2004] T. E. Bruns and O. Sigmund, "Toward the topology design of mechanisms that exhibit snap-through behavior", *Comput. Methods Appl. Mech. Eng.* **193**:36-38 (2004), 3973–4000.

[Bruns and Tortorelli 2003] T. E. Bruns and D. A. Tortorelli, "An element removal and reintroduction strategy for the topology optimization of structures and compliant mechanisms", *Int. J. Numer. Methods Eng.* **57**:10 (2003), 1413–1430.

[Bruns et al. 2002] T. E. Bruns, O. Sigmund, and D. A. Tortorelli, "Numerical methods for the topology optimization of structures that exhibit snap-through", *Int. J. Numer. Methods Eng.* **55**:10 (2002), 1215–1237.

[Buhl et al. 2000] T. Buhl, C. B. W. Pedersen, and O. Sigmund, "Stiffness design of geometrically nonlinear structures using topology optimization", *Struct. Multidiscip. Optim.* **19**:2 (2000), 93–104.

[Cho and Jung 2003] S. Cho and H.-S. Jung, "Design sensitivity analysis and topology optimization of displacement-loaded non-linear structures", *Comput. Methods Appl. Mech. Eng.* **192**:22–24 (2003), 2539–2553.

[Cook et al. 2001] R. D. Cook et al., *Concepts and applications of finite element analysis*, 4th ed., Wiley, New York, 2001.

[Hellweg and Crisfield 1998] H.-B. Hellweg and M. A. Crisfield, "A new arc-length method for handling sharp snap-backs", *Comput. Struct.* **66**:5 (1998), 704–709.

[Huang and Xie 2008] X. Huang and Y. M. Xie, "Topology optimization of nonlinear structures under displacement loading", *Eng. Struct.* **30**:7 (2008), 2057–2068.

[Kamat and Ruangsilasingha 1985] M. P. Kamat and P. Ruangsilasingha, "Optimization of space trusses against instability using design sensitivity derivatives", *Eng. Optim.* **8**:3 (1985), 177–188.

[Langelaar et al. 2005] M. Langelaar, G. H. Yoon, Y. Y. Kim, and F. van Keulen, "Topology optimization of shape memory alloy actuators", in *WCSMO6: Proceedings of the 6th World Congress on Structural and Multidisciplinary Optimization* (Rio de Janeiro, 2005), edited by J. Herskovits et al., International Society for Structural and Multidisciplinary Optimization/COPPE, Rio de Janeiro, 2005.

[Liu and Nocedal 1989] D. C. Liu and J. Nocedal, "On the limited memory BFGS method for large scale optimization", *Math. Program.* **45**:1–3 (1989), 503–528.

[Lu and Kota 2003] K.-J. Lu and S. Kota, "Synthesis of shape morphing compliant mechanisms using a load path representation method", pp. 337–348 in *Smart structures and materials 2003: Modeling, signal processing, and control* (San Diego, CA, 2003), edited by R. C. Smith, SPIE Proceedings **5049**, SPIE, Bellingham, WA, 2003.

[Ohsaki 2005] M. Ohsaki, "Design sensitivity analysis and optimization for nonlinear buckling of finite-dimensional elastic conservative structures", *Comput. Methods Appl. Mech. Eng.* **194**:30-33 (2005), 3331–3358.

[Prasad and Diaz 2006] J. Prasad and A. R. Diaz, "Synthesis of bistable periodic structures using topology optimization and a genetic algorithm", *J. Mech. Des.* (*ASME*) **128**:6 (2006), 1298–1306.

[Santer and Pellegrino 2008] M. Santer and S. Pellegrino, "Compliant multistable structural elements", *Int. J. Solids Struct.* **45**:24 (2008), 6190–6204.

[Saxena 2005] A. Saxena, "Synthesis of compliant mechanisms for path generation using genetic algorithm", *J. Mech. Des.* (*ASME*) **127**:4 (2005), 745–752.

[Sekimoto and Noguchi 2001] T. Sekimoto and H. Noguchi, "Homologous topology optimization in large displacement and buckling problems", *JSME Int. J. A Mech. M.* **44**:4 (2001), 616–622.

[Svanberg 1987] K. Svanberg, "The method of moving asymptotes: a new method for structural optimization", *Int. J. Numer. Methods Eng.* **24**:2 (1987), 359–373.

[Thai and Kim 2009] H.-T. Thai and S.-E. Kim, "Large deflection inelastic analysis of space trusses using generalized displacement control method", *J. Constr. Steel Res.* **65**:10-11 (2009), 1987–1994.

[Vila et al. 1997] A. Vila, A. Rodríguez-Ferran, and A. Huerta, "A note on a numerical benchmark test: an axisymmetric shell under ring loads", *Commun. Numer. Methods Eng.* **13**:3 (1997), 181–192.

[Yoon 2010] G. H. Yoon, "Maximizing the fundamental eigenfrequency of geometrically nonlinear structures by topology optimization based on element connectivity parameterization", *Comput. Struct.* **88**:1–2 (2010), 120–133.

[Yoon and Kim 2005a] G. H. Yoon and Y. Y. Kim, "Element connectivity parameterization for topology optimization of geo-metrically nonlinear structures", *Int. J. Solids Struct.* **42**:7 (2005), 1983–2009.

[Yoon and Kim 2005b] G. H. Yoon and Y. Y. Kim, "The element connectivity parameterization formulation for the topology design optimization of multiphysics systems", *Int. J. Numer. Methods Eng.* **64**:12 (2005), 1649–1677.

[Yoon et al. 2007] G. H. Yoon, J. S. Jensen, and O. Sigmund, "Topology optimization of acoustic-structure interaction problems using a mixed finite element formulation", *Int. J. Numer. Methods Eng.* **70**:9 (2007), 1049–1075.

[Yoon et al. 2008] G. H. Yoon, Y. Y. Kim, M. Langelaar, and F. van Keulen, "Theoretical aspects of the internal element connectivity parameterization approach for topology optimization", *Int. J. Numer. Methods Eng.* **76**:6 (2008), 775–797.

GIL HO YOON: ghy@hanyang.ac.kr
Department of Mechanical Engineering, Hanyang University, 222 Wangsimni-ro, Seongdong-gu, Seoul 133-791, South Korea

JIN YEE NOH: jyn@knu.ac.kr
School of Mechanical Engineering, Kyungpook National University, South Korea

YOON YOUNG KIM: yykim@snu.ac.kr
Multiscale Design Center and Integrated Design and Analysis of Structures Laboratory, School of Mechanical and Aerospace Engineering, Seoul National University, Kwanak-Gu San 56-1, Seoul 151-742, South Korea

msp

■
■msp

ANALYSIS OF HARD COATINGS ON A SUBSTRATE CONTAINING INHOMOGENEITIES

KUN ZHOU, LEON M. KEER AND Q. JANE WANG

We investigate the effect stiff inhomogeneities in the substrate beneath a hard coating have on the elastic field. The solution of multiple interacting three-dimensional inhomogeneities in a half-space is utilized by modeling a coating layer as an inhomogeneity of finite size. The study shows that stiff inhomogeneities in the substrate do not worsen the cracking and debonding of hard coatings, but are still detrimental to the yielding behavior of the substrate even though it is under the protection of a hard coating.

1. Introduction

Inhomogeneities with elastic moduli different from those of their surrounding matrix are ubiquitous in solid materials, and their presence significantly affects the physical and mechanical properties of materials at the local and the global scale [Mortensen 2007]. For a given matrix material, inhomogeneities are categorized into two types: stiff ones which have larger Young's modulus than the matrix and compliant ones with smaller Young's modulus. Generally, stiff inhomogeneities can cause stress to increase in the matrix material, compared with compliant inhomogeneities. Nonmetallic inhomogeneities such as oxides and nitrides are typical stiff inhomogeneities in steels.

Hard coatings have been used to protect substrate matrix materials from wear and yielding damage; however, they are subject to interfacial cracking and debonding and surface cracking failure under combined heavy normal and tangential loading. Such failure may be more likely when inhomogeneities in the substrate are closer to the interface of the coating-substrate system. To improve reliability and functionality, it would be useful to understand how and to what extent inhomogeneities degrade the performance of coatings. However, almost no theoretical or experimental studies on such problems exist, due to the complex nature of the solution for the elastic field of inhomogeneities in a coating-substrate system. Such a solution would be a foundation for further studies but is not yet available.

Eshelby [1957] pioneered the work on inhomogeneities by solving the elastic field of an ellipsoidal inhomogeneity embedded in an infinite matrix with the equivalent inclusion method (EIM). The EIM assumes that an inhomogeneity can be modeled by a homogenous inclusion with unknown equivalent eigenstrain to be determined. A homogeneous inclusion has the same material properties as the matrix but contains eigenstrain, a generic term that refers to inelastic strain such as plastic strain, misfit strain, thermal expansion and phase transformation [Mura 1987]. Since then, the EIM has been used extensively to investigate a single inhomogeneity [Johnson et al. 1980a; 1980b; Nakasone et al. 2000; Dong et al. 2002; Kirilyuk and Levchuk 2005], two interacting inhomogeneities [Moschovidis and Mura 1975; Fond

---

et al. 2001; Shodja and Sarvestani 2001; Shodja et al. 2003], and multiple inhomogeneities [Nemat-Nasser et al. 1982; 1993; Luo and Weng 1987; Hori and Nemat-Nasser 1993; Shodja and Roumi 2005; Duan et al. 2006] in an infinite matrix. However, the interactions of all inhomogeneities were not fully taken into account in the referred studies of multiple inhomogeneities. Other methods have also been used to study a single or two inhomogeneities in an infinite matrix, e.g., the complex function method [Kushch et al. 2005], the boundary element method [Tan et al. 1992], and the volume integral equation method [Dong et al. 2003].

So far, few investigations have been conducted on inhomogeneities in a half-space. These studies include a circular rigid disc [Hunter and Gamblen 1974; Selvadurai 2001], an ellipsoidal inhomogeneity [Tsuchida and Mura 1983], a hemispheroidal inhomogeneity [Kouris and Mura 1989], a spheroidal in-homogeneity [Tsuchida et al. 2000], two concentric spherical inhomogeneities [Molchanov et al. 2002], and one or multiple two-dimensional arbitrarily shaped inhomogeneities [Kuo 2007; 2008] The bound-ary element method was used in these last two works, while the Papkovich–Neuber or the Boussinesq displacement potential was used to solve the problems in the other referred works. Studies were also performed on a single or multiple inhomogeneities in one of two joining half-surfaces of dissimilar materials, e.g., [Meguid and Zhu 1995; Yu and Kuang 2003; Brusselaars et al. 2007]. The results of those studies can be applicable to the problems of inhomogeneities in a half-space when one of the two joining half-spaces is set free.

Recently, Zhou et al. [2010] solved multiple interacting three-dimensional inhomogeneities of arbitrary shape in a half-space using the EIM. In this paper, we apply this solution to study the effect of inhomo-geneities embedded in the substrate of a coating-substrate system on the elastic field of the coating by modeling the coating as an inhomogeneity with respect to the substrate.

2. Methodology

Consider a coating-substrate system that contains multiple inhomogeneities in the substrate and is subject to external loading on the surface of the coating. The loading is prescribed as a known function and assumed to be located within a circular area of radius R. The solution strategy is to model the coating by a cuboidal inhomogeneity of size $L_x \times L_y \times H$ with H equal to the thickness of the coating. The dimensions L_x and L_y are set to be much larger than both R and H to simulate the infinite dimensions of the coating in the x-y plane. In this way, the original coating-substrate problem with inhomogeneities is converted into a half-space problem with inhomogeneities.

Let us now turn to a half-space with elastic moduli C_{ijkl} ($i, j, k, l = 1, 2, 3$) that contains n inhomo-geneities Ω_ψ ($\psi = 1, 2, \ldots, n$) with elastic moduli C^ψ_{ijkl}. Since there is no inelastic strain within Ω_ψ, the total strain ε_{ij} at any point within Ω_ψ is the elastic strain ε^e_{ij}. According to Hooke's law, the stress σ_{ij} within Ω_ψ is given by $\sigma_{ij} = C^\psi_{ijkl}\varepsilon_{kl}$. Using the EIM, each Ω_ψ is simulated by a homogeneous inclusion Ω^I_ψ that has the same elastic moduli as the matrix but contains equivalent eigenstrain ε^*_{ij} to be determined. The equivalent eigenstrains ε^*_{ij} are introduced to represent material differences of the inhomogeneities, the interactions among them, and their responses to the external load (applied stresses). The total strain ε_{ij} within Ω^I_ψ contains elastic strain ε^e_{ij} and equivalent eigenstrain ε^*_{ij}, and according to Hooke's law, the stress σ_{ij} within Ω^I_ψ is given by $\sigma_{ij} = C_{ijkl}(\varepsilon_{kl} - \varepsilon^*_{kl})$. Furthermore, σ_{ij} can be decomposed into $\sigma_{ij} = \sigma^*_{ij} + \sigma^0_{ij}$, where σ^*_{ij} is the eigenstress caused by the equivalent eigenstrains ε^*_{ij} in all Ω^I_ψ and σ^0_{ij}

the applied stress due to the external load. From these three stress expressions, the governing equation is established as

$$C^{\Psi}_{ijkl}C^{-1}_{klmq}\sigma^*_{mq} - \sigma^*_{ij} + C^{\Psi}_{ijkl}\varepsilon^*_{kl} = \sigma^0_{ij} - C^{\Psi}_{ijkl}C^{-1}_{lkmq}\sigma^0_{mq} \quad (\psi = 1, 2, \ldots, n; \; i, j, k, l, m, q = 1, 2, 3). \quad (1)$$

where the equivalent eigenstrains ε^*_{ij} are unknowns. This equation cannot be solved until the relationship between σ^*_{ij} and ε^*_{ij} is determined.

Figure 1 illustrates a cuboidal computational domain D of size $L_x \times L_y \times L_z$ that contains n inhomogeneities Ω_ψ in a half-space, which is bounded by the surface plane $z = 0$ in the x-y-z coordinate system. The cuboidal inhomogeneity Ω_1 of size $L_x \times L_y \times H$ at the surface is used to model a layer of coating of thickness H. The domain D is discretized into $N_x \times N_y \times N_z$ small cuboidal elements of the same size, and each Ω_ψ is then geometrically approximated by a collection of such cuboidal elements. The finer the discretization, the more accurate the approximation would be. If the size of the cuboidal element is small enough, each cuboidal inhomogeneity within Ω_ψ can be treated as containing uniform eigenstrain. It is noted that the eigenstrain within Ω_ψ is still nonuniform. For convenience, each cuboidal element is indexed by a sequence of three integers $[\alpha, \beta, \gamma]$ $(0 \le \alpha \le N_x - 1, 0 \le \beta \le N_y - 1, 0 \le \gamma \le N_z - 1)$. Using the EIM, each cuboidal inhomogeneity is simulated as a cuboidal homogenous inclusion with uniform equivalent eigenstrain. Thus, (1) can be recast into

$$(\mathbf{C}_{\alpha,\beta,\gamma}\mathbf{C}^{-1} - \mathbf{I})\boldsymbol{\sigma}^*_{\alpha,\beta,\gamma} + \mathbf{C}_{\alpha,\beta,\gamma}\boldsymbol{\varepsilon}^*_{\alpha,\beta,\gamma} = (\mathbf{I} - \mathbf{C}_{\alpha,\beta,\gamma}\mathbf{C}^{-1})\boldsymbol{\sigma}^0_{\alpha,\beta,\gamma},$$
$$(\mathbf{C}_{\alpha,\beta,\gamma} \in \mathbf{C}_\psi \; (\psi = 1, 2, \ldots, n), \; 0 \le \alpha \le N_x - 1, \; 0 \le \beta \le N_y - 1, \; 0 \le \gamma \le N_z - 1), \quad (2)$$

where $\mathbf{C}$, $\boldsymbol{\varepsilon}^*$, $\boldsymbol{\sigma}^*$ and $\boldsymbol{\sigma}^0$ are written in matrix form and $\mathbf{I}$ is a unit matrix. The symbol $\mathbf{C}$ denotes the elastic moduli of the substrate matrix, $\mathbf{C}_{\alpha,\beta,\gamma}$ denotes the elastic moduli of the cuboid $[\alpha, \beta, \gamma]$ within Ω_ψ, $\boldsymbol{\varepsilon}^*_{\alpha,\beta,\gamma}$ the uniform equivalent eigenstrain within $[\alpha, \beta, \gamma]$, $\boldsymbol{\sigma}^*_{\alpha,\beta,\gamma}$ the stress at an observation point inside $[\alpha, \beta, \gamma]$ caused by the equivalent eigenstrains in all the cuboidal inclusions, and $\boldsymbol{\sigma}^0_{\alpha,\beta,\gamma}$ the applied stress at the observation point. All the observation points are chosen to be at the center of each cuboid.
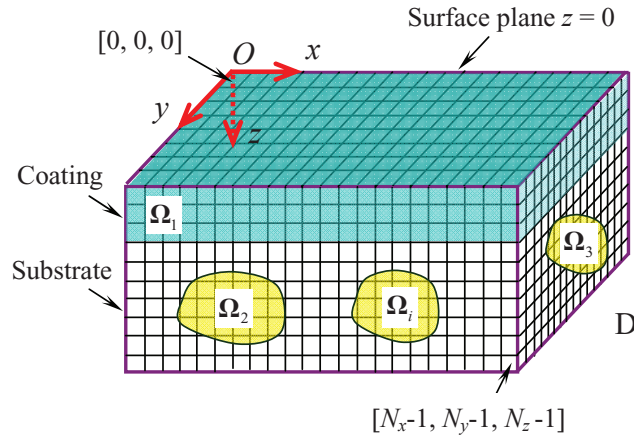
Figure 1. Discretization of a computational domain D into $N_x \times N_y \times N_z$ cuboids. The domain D contains n arbitrarily shaped inhomogeneities in an isotropic half-space, where Ω_1 is used to model a layer of coating on the substrate matrix.

The final solvable form for (2) is given in [Zhou et al. 2010] as

$$(\mathbf{C}_{\alpha,\beta,\gamma}\mathbf{C}^{-1} - \mathbf{I}) \sum_{\phi=0}^{N_z-1} \sum_{\zeta=0}^{N_y-1} \sum_{\xi=0}^{N_x-1} \mathbf{A}_{\alpha-\xi,\beta-\zeta,\gamma-\phi} \boldsymbol{\varepsilon}^*_{\xi,\zeta,\phi} + \mathbf{C}_{\alpha,\beta,\gamma} \boldsymbol{\varepsilon}^*_{\alpha,\beta,\gamma} = (\mathbf{I} - \mathbf{C}_{\alpha,\beta,\gamma}\mathbf{C}^{-1})\boldsymbol{\sigma}^0_{\alpha,\beta,\gamma}$$

$$(\mathbf{C}_{\alpha,\beta,\gamma} \in \mathbf{C}_\psi \,(\psi = 1, 2, \ldots, n), \ 0 \le \alpha \le N_x - 1, \ 0 \le \beta \le N_y - 1, \ 0 \le \gamma \le N_z - 1). \quad (3)$$

where $\mathbf{A}_{\alpha-\xi,\beta-\zeta,\gamma-\varphi}$ relates the eigenstress $\boldsymbol{\sigma}^*_{\alpha,\beta,\gamma}$ at the observation point in the cuboid $[\alpha, \beta, \gamma]$ to the uniform eigenstrain $\boldsymbol{\varepsilon}^*_{\xi,\zeta,\varphi}$ in the cuboid $[\xi, \zeta, \varphi]$. The expression for $\mathbf{A}_{\alpha-\xi,\beta-\zeta,\gamma-\phi}$ is given in the paper by Chiu [1978] who solved the elastic field of a cuboid containing uniform eigenstrain in a half-space. Chiu's solution shows that there are stress singularities at certain edges and corners of the cuboid, depending upon the types of eigenstrain in the cuboid (also see [Chiu 1977; Mura 1987]). Nevertheless, the stress singularities would not affect the solution of (3) because the observation points where the stresses are concerned are taken only at the center of each cuboid.

Equation (3) can be solved by the conjugated gradient method [Shewchuk 1994], a well-established method for solving linear equations by iteration. Furthermore, the fast Fourier transform technique can be used for rapid calculation of the summations in (3) to achieve computational efficiency [Zhou et al. 2009]. Once all the eigenstrains $\boldsymbol{\varepsilon}^*$ are determined, the eigenstress $\boldsymbol{\sigma}^*$ due to $\boldsymbol{\varepsilon}^*$ is known and then the overall stress field $\boldsymbol{\sigma}$ is obtained by $\boldsymbol{\sigma} = \boldsymbol{\sigma}^* + \boldsymbol{\sigma}^0$.

3. Results and discussion

The study is performed on the elastic analysis of a layer of WC hard coating on a steel substrate in which a stiff cuboidal Al_2O_3 inhomogeneity is formed near the substrate surface prior to coating deposition, as illustrated in Figure 2. The steel matrix has Young's modulus $E = 210\,\text{GPa}$ and Poisson's ratio $\nu = 0.28$,
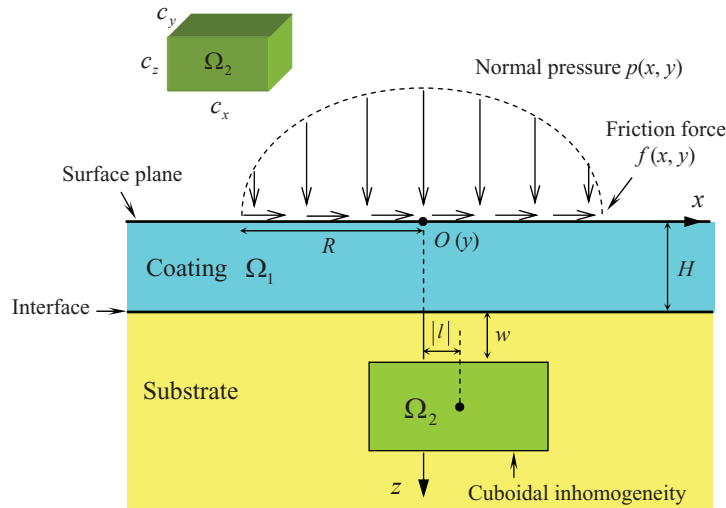
Figure 2. Schematic of a cuboidal inhomogeneity located beneath the interface of a coating-substrate system. The surface of the coating is subject to pressure and friction force.

the WC coating has $E_1 = 670\,\text{GPa}$ and $\nu_1 = 0.28$, and the Al_2O_3 inhomogeneity has $E_2 = 344\,\text{GPa}$ and $\nu_2 = 0.25$. The coating surface, bounded by the plane $z = 0$ in the x-y-z Cartesian coordinate system, is subjected to the prescribed normal pressure $p(x, y) = p_0\left(1 - (x/R)^2 - (y/R)^2\right)^{1/2}$ and friction force $f(x, y) = \mu p(x, y)$ in a circle of radius R. The coating thickness is denoted by H, and the cases of $h \leq 0.5R$ are mainly investigated. The cuboidal inhomogeneity has dimensions $c_x = c_y = R$ and $c_z = 0.5R$; its depth from the interface is denoted by w. The center of the cuboid is located in the plane $y = 0$ and distanced from the plane $x = 0$ by $|l|$ with l being the x-coordinate value of the center.

In the calculation, a computational domain of size $16R \times 16R \times 1.5R$ is used and the coating is simulated by the cuboidal inhomogeneity Ω_1 of $16R \times 16R \times H$. The cuboid Ω_1 has much larger dimension in the length and width directions than in the thickness direction in order to approximate the effect of infinite dimensions that the coating has in the x-y plane. The dimension $16R$ is also much larger than the radius R of the pressure area on the coating surface. The inhomogeneity, as mentioned, may have singularities at its edges and corners, but these singularities would not affect the stress calculation.

The case of $H = 0.5R$, $w = 0.25R$, $l = 0$, and $\mu = 0.3$ is first studied and shows that both the surface and interfacial stresses of the coating approach zero approximately at $x = \pm 4R$. Therefore, it is justifiable to model the infinitely extended coating by an inhomogeneity of finite size embedded at $x = \pm 8R$. Figure 3 presents the surface stress component σ_x^{surf} and the interfacial stress components σ_x^{int}, σ_z^{int} and σ_{xz}^{int} within the coating in the central plane $y = 0$ and along the x direction. All the stress components are normalized by the peak pressure p_0 at $(0, 0, 0)$. Results show that the surface normal stress σ_x^{surf} is tensile in the area behind the friction direction ($x < -R$), but compressive both in the pressure area and in the area ahead of the friction direction ($x > -R$). The stress σ_x^{surf} reaches its maximum tensile value around the edge of the pressure area at $x = -R$, and therefore is likely to cause crack nucleation there. The interfacial normal stress σ_x^{int} within the coating is tensile in the area under the surface pressure and its maximum value is comparable to that of σ_x^{surf}. This tensile stress may cause cracks to nucleate at the interface and then grow into the film along the direction perpendicular to the interface or cause the kinking of
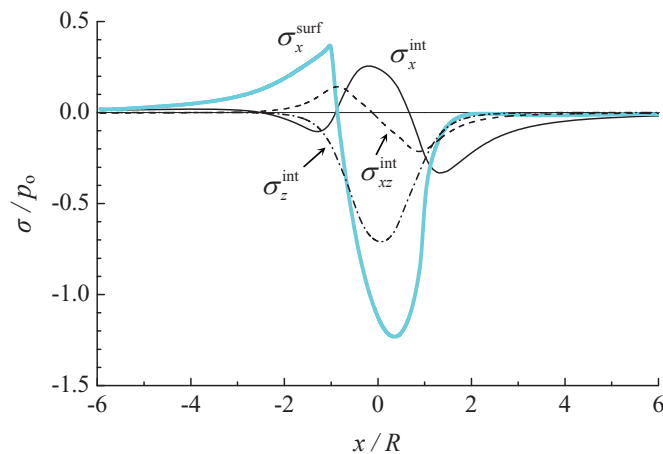
Figure 3. The component σ_x^{surf} of the surface stress and the components σ_x^{int}, σ_z^{int}, and σ_{xz}^{int} of the interfacial stress in the coating along the x-axis direction in the computational domain of size $16R \times 16R \times 1.5R$.
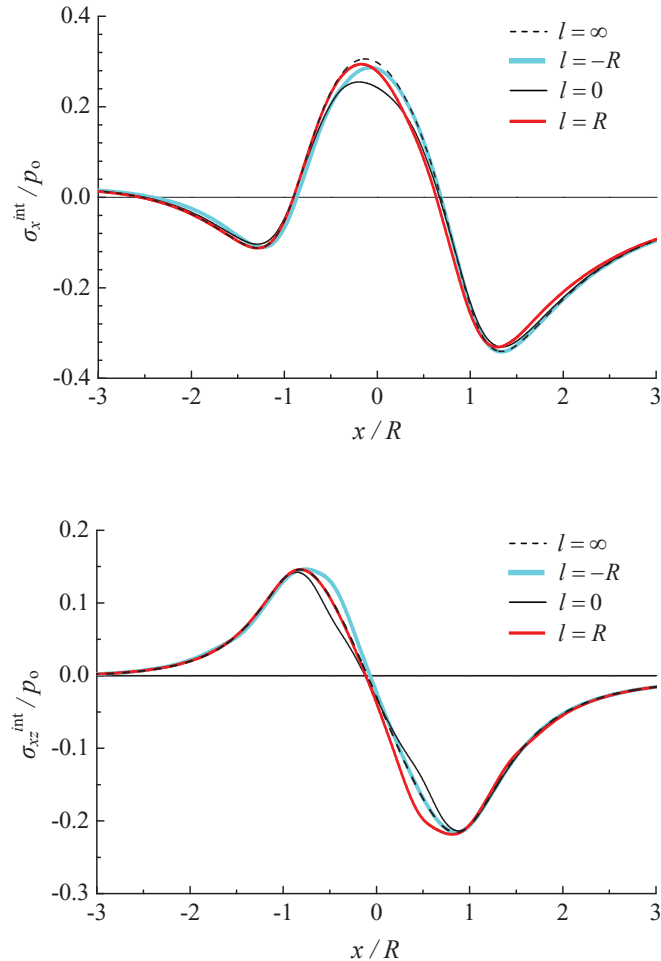
Figure 4. Effect of the horizontal position of the inhomogeneity Ω_2 on the interfacial normal stress σ_x^{int} (top) and shear stress σ_{xz}^{int} (bottom) within the coating.

interfacial cracks. Besides cracking within the coating, interfacial debonding of the coating is another damage phenomenon that needs to be considered. The interfacial shear stresses σ_{xz}^{int} affect the adhesion of the coating onto the substrate; their low magnitude would generally lead to good adhesion, thereby preventing the occurrence of interfacial debonding. Figure 3 shows that σ_{xz}^{int} reaches positive and negative maximum values in the locations approximately beneath the edges of surface pressure area ($x = \pm R$). It is predicted that interfacial debonding is likely to start at such locations.

We then study the effect of the inhomogeneity Ω_2 on cracking and debonding of the coating. Figure 4 compares the interfacial stress components σ_x^{int} and σ_{xz}^{int} among the cases for different horizontal locations of the inhomogeneity Ω_2. The results show that the presence of Ω_2 decreases the tensile values of σ_x^{int}, compared with when Ω_2 is absent ($l = \infty$). The maximum tensile value of σ_x^{int} is decreased by about 17% when Ω_2 is directly beneath the peak surface pressure ($l = 0$). The decrease becomes smaller when Ω_2 is horizontally displaced from $l = 0$ to $l = -R$ or $l = R$. On the other hand, the change in both the
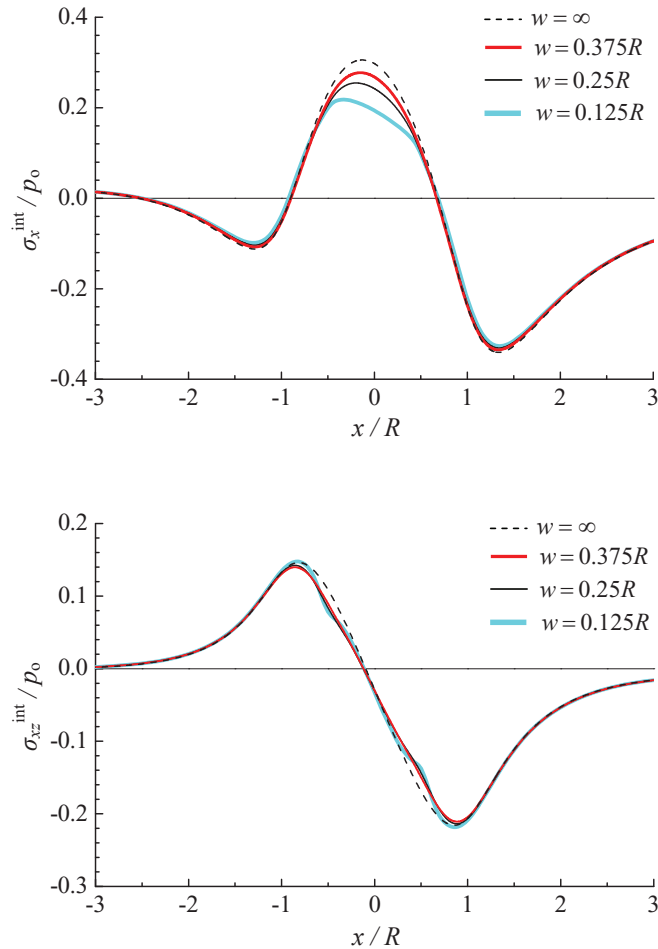
Figure 5. Effect of the depths of the inhomogeneity Ω_2 on the interfacial normal stress σ_x^{int} (top) and shear stress σ_{xz}^{int} (bottom) within the coating.

maximum positive and negative magnitudes of σ_{xz}^{int} is almost inappreciable. At $l = 0$, σ_{xz}^{int} decreases in the areas near the points of maximum negative and positive magnitudes, while at $l = -R$ ($l = R$), σ_{xz}^{int} increases in the area near the point of the maximum positive (negative) magnitude. Nevertheless, the overall change in σ_{xz}^{int} is not significant.

Figure 5 compares the interfacial stress components σ_x^{int} and σ_{xz}^{int} among the cases for different depths of the inhomogeneity Ω_2 at $l = 0$ and shows that as Ω_2 approaches the interface, the maximum tensile value of σ_x^{int} decreases and the reduction reaches about 27% at the depth $w = 0.125R$. The overall change in σ_{xz}^{int} is not significant except that the peak and valley shapes of σ_{xz}^{int} are narrower in the presence of Ω_2 than in the absence of Ω_2 ($w = \infty$).

Figures 4 and 5 demonstrate that the presence of stiff inhomogeneities in the substrate would not worsen the cracking and debonding of a hard coating and instead may likely play a role to decrease the possibility of their occurrence.
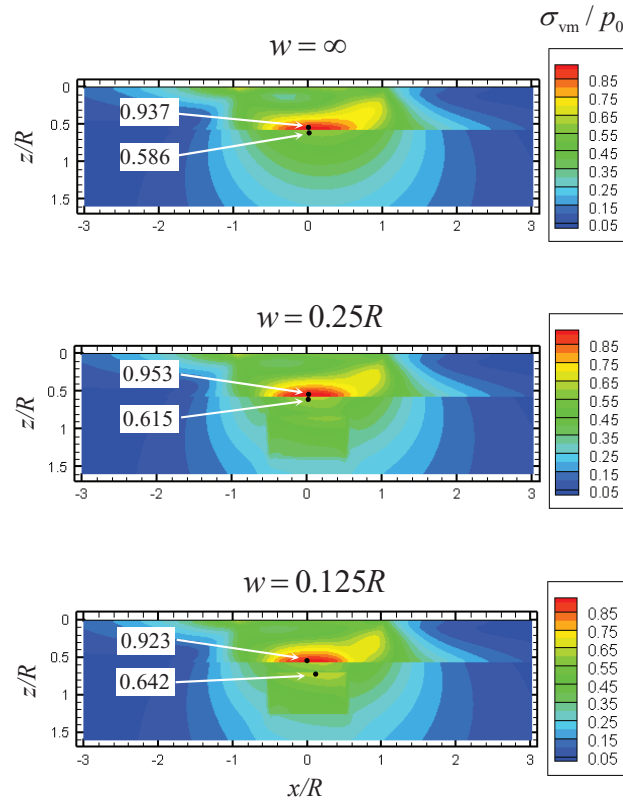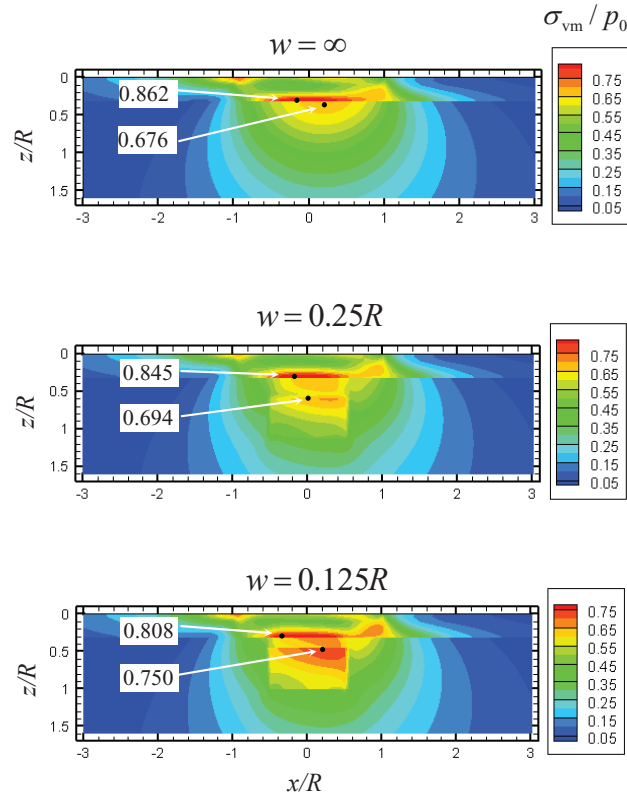
Figure 6. Effect of the depths of the inhomogeneity Ω_2 on the von Mises stress field in the coating-substrate system. The coating has a thickness of $0.5R$.

Figure 6 further presents the normalized von Mises stresses $\bar{\sigma}_{vm}$ in both the coating and the substrate for different depths of the inhomogeneity Ω_2 at $l = 0$. In the absence of Ω_2 ($w = \infty$), the maximum von Mises stress $\bar{\sigma}_{vm}^{max}$ in the coating reaches 0.937 at the interface location approximately beneath the peak surface pressure. As Ω_2 approaches the interface, $\bar{\sigma}_{vm}^{max}$ in the coating slightly increases to 0.953 and then decreases to 0.923, and its location remains the same. The increase in $\bar{\sigma}_{vm}^{max}$ is almost negligible (about 1.7%). In contrast, $\bar{\sigma}_{vm}^{max}$ in the substrate monotonically increases from 0.586 to 0.642, and its location changes from the interface location to the edge plane of the inhomogeneity Ω_2. The increase in $\bar{\sigma}_{vm}^{max}$ is significant (about 8.7%) and means yielding is more likely to occur in the substrate. Therefore, stiff inhomogeneities are still detrimental to the yielding behavior of the substrate even though it is protected by a layer of hard coating.

The thickness of a coating is a key parameter that affects its performance. The previous examples consider coatings of thickness $H = 0.5R$. The coatings with a thinner thickness $H = 0.25R$ are also studied. In Figure 7, the solid lines plot the surface stress component σ_x^{surf} and the interfacial stress components σ_x^{int} and σ_{xz}^{int} in the coating for the case of $H = 0.25R$, $w = 0.25R$, $l = 0$, and $\mu = 0.3$ along the x-axis, while the dashed lines represent the corresponding stress components for the case in the absence of the inhomogeneity Ω_2. Their comparison shows that the presence of Ω_2 does not increase σ_x^{surf}, σ_x^{int} and σ_{xz}^{int} or cause significant change to them. Thus, for a thinner hard coating, it is still valid
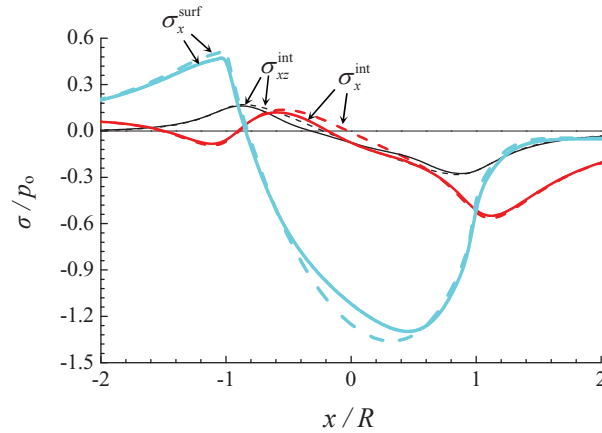
Figure 7. The surface stress component σ_x^{surf} and the interfacial stress components σ_x^{int} and σ_{xz}^{int} in the coating of thickness $H = 0.25R$ along the x-axis direction.

to draw the conclusion that stiff inhomogeneities in the substrate would not worsen the cracking and debonding of the coating.

Figure 7 shows that the maximum surface tensile stress σ_x^{surf} is much larger than the maximum interfacial tensile stress σ_x^{int}, which leads to the conclusion that surface cracking may start much earlier than cracking nucleated at the interface or dominate the damage mechanisms. An optimal design of hard coatings should balance the two cracking mechanisms to avoid the early occurrence of either one of them. The proper selection of coating thickness is a way to achieve this balance. Figure 3 demonstrates that when the coating is selected to have the thickness of $H = 0.5R$, the maximum tensile value of σ_x^{surf} is close to that of σ_x^{int}.

The normalized von Mises stresses $\bar{\sigma}_{\text{vm}}$ in both the coating and the substrate are presented in Figure 8 and show that as the inhomogeneity is located closer to the coating-substrate interface, $\bar{\sigma}_{\text{vm}}$ in the substrate increases and the increase reaches about 11%, which demonstrates again that detrimental effect of the stiff inhomogeneity on the yielding behavior of the substrate must be accounted for. Furthermore, the comparison between Figures 6 and 8 shows that a thicker coating ($X = 0.5R$) can provide better yielding protection to the substrate than a thinner coating ($H = 0.25R$).

Inhomogeneities can become stringers or clusters during their formation process. We investigate a stringer of Al_2O_3 inhomogeneities in the WC-coated steel substrate. The stringer is modeled as three identical cuboidal Al_2O_3 inhomogeneities, which have the same dimension as the previously studied single inhomogeneity and are equally distanced by $0.25R$. Figure 9 compares the von Mises stress field of the stringer between the two cases in which the coating has different thicknesses of $H = 0.5R$ and $h = 0.25R$. The stringer is located beneath the interface at the depth $w = 0.125R$ and shows that the interactions among the inhomogeneities distort the stress field around them and the distortion becomes stronger as the coating becomes thinner.

Although this study focuses on inhomogeneities of cuboidal shape, it should be noted that the present method is also capable of handling inhomogeneities of any arbitrary shapes. An arbitrarily shaped geometry can be approximated by the appropriate arrangement of multiple infinitesimal cuboidal elements.

Figure 8. Effect of the depths of the inhomogeneity Ω_2 on the von Mises stress field in the coating-substrate system. The coating has a thickness of $0.25R$.
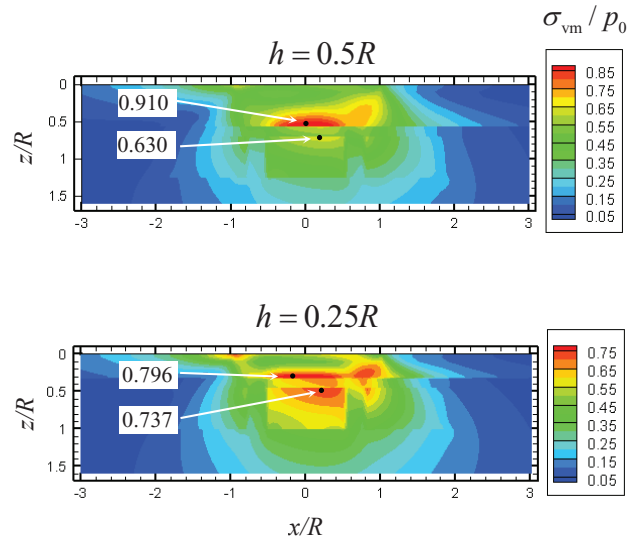
Figure 9. Effect of the depths of the inhomogeneity Ω_2 on the von Mises stress field in the coating-substrate system. The coating has a thickness of $0.25R$.

4. Conclusions

The effect of stiff inhomogeneities on the elastic field of hard coatings is studied by using the solution of multiple interacting three-dimensional inhomogeneities of arbitrary shape in a half-space. In the utilization of the inhomogeneity solution, a layer of coating is modeled by an inhomogeneity of finite size embedded in the half-space, which has much larger dimensions in the horizontal directions than in the depth direction to simulate the effect of the finitely extended dimensions of the coating. The study shows:

(1) Stiff inhomogeneities formed in the substrate would not worsen the cracking and debonding of a hard coating.

(2) Stiff inhomogeneities are still detrimental to the yielding behavior of the substrate even though coated by a layer of hard coating.

These considerations may provide guidance to minimize the potential damage induced by inhomogeneities to coating-substrate systems.

This study also suggests that the effect of inhomogeneities on yielding and plastic zone expansion cannot be neglected when a film-substrate system containing inhomogeneities is subject to elastic-plastic indentation. Based on the present method, a further study can be conducted to develop an elastic-plastic indentation model that takes into account inhomogeneities for indentation tests.

Acknowledgements

References

[Brusselaars et al. 2007] N. Brusselaars, S. G. Mogilevskaya, and S. L. Crouch, "A semi-analytical solution for multiple circular inhomogeneities in one of two joined isotropic elastic half-planes", *Eng. Anal. Bound. Elem.* **31**:8 (2007), 692–705.

[Chiu 1977] Y. P. Chiu, "On the stress field due to initial strains in a cuboid surrounded by an infinite elastic space", *J. Appl. Mech.* (*ASME*) **44**:4 (1977), 587–590.

[Chiu 1978] Y. P. Chiu, "On the stress field and surface deformation in a half space with a cuboidal zone in which initial strains are uniform", *J. Appl. Mech.* (*ASME*) **45**:2 (1978), 302–306.

[Dong et al. 2002] C. Y. Dong, Y. K. Cheung, and S. H. Lo, "A regularized domain integral formulation for inclusion problems of various shapes by equivalent inclusion method", *Comput. Methods Appl. Mech. Eng.* **191**:31 (2002), 3411–3421.

[Dong et al. 2003] C. Y. Dong, S. H. Lo, and Y. K. Cheung, "Numerical solution of 3D elastostatic inclusion problems using the volume integral equation method", *Comput. Methods Appl. Mech. Eng.* **192**:1-2 (2003), 95–106.

[Duan et al. 2006] H. L. Duan, Y. Jiao, X. Yi, Z. P. Huang, and J. Wang, "Solutions of inhomogeneity problems with graded shells and application to core-shell nanoparticles and composites", *J. Mech. Phys. Solids* **54**:7 (2006), 1401–1425.

[Eshelby 1957] J. D. Eshelby, "The determination of the elastic field of an ellipsoidal inclusion, and related problems", *Proc. R. Soc. Lond. A* **241**:1226 (1957), 376–396.

[Fond et al. 2001] C. Fond, A. Riccardi, R. Schirrer, and F. Montheillet, "Mechanical interaction between spherical inhomogeneities: an assessment of a method based on the equivalent inclusion", *Eur. J. Mech. A Solids* **20**:1 (2001), 59–75.

[Hori and Nemat-Nasser 1993] M. Hori and S. Nemat-Nasser, "Double-inclusion model and overall moduli of multi-phase composites", *Mech. Mater.* **14**:3 (1993), 189–206.

[Hunter and Gamblen 1974] S. C. Hunter and D. Gamblen, "The theory of a rigid circular disc ground anchor buried in an elastic soil either with adhesion or without adhesion", *J. Mech. Phys. Solids* **22**:5 (1974), 371–399.

[Johnson et al. 1980a] W. C. Johnson, Y. Y. Earmme, and J. K. Lee, "Approximation of the strain field associated with an inhomogeneous precipitate, 1: Theory", *J. Appl. Mech.* (*ASME*) **47**:4 (1980), 775–780.

[Johnson et al. 1980b] W. C. Johnson, Y. Y. Earmme, and J. K. Lee, "Approximation of the strain field associated with an inhomogeneous precipitate, 2: The cuboidal inhomogeneity", *J. Appl. Mech.* (*ASME*) **47**:4 (1980), 781–788.

[Kirilyuk and Levchuk 2005] V. S. Kirilyuk and O. I. Levchuk, "Stress state of a transversely isotropic medium with an arbitrarily oriented spheroidal inclusion", *Prik. Mekh.* **41**:2 (2005), 33–40. In Russian: translated in *Int. Appl. Mech.* **41**:2 (2005), 137–143.

[Kouris and Mura 1989] D. A. Kouris and T. Mura, "The elastic field of a hemispherical inhomogeneity at the free surface of an elastic half space", *J. Mech. Phys. Solids* **37**:3 (1989), 365–379.

[Kuo 2007] C.-H. Kuo, "Stress disturbances caused by the inhomogeneity in an elastic half-space subjected to contact loading", *Int. J. Solids Struct.* **44**:3-4 (2007), 860–873.

[Kuo 2008] C.-H. Kuo, "Contact stress analysis of an elastic half-plane containing multiple inclusions", *Int. J. Solids Struct.* **45**:16 (2008), 4562–4573.

[Kushch et al. 2005] V. I. Kushch, S. V. Shmegera, and V. A. Buryachenko, "Interacting elliptic inclusions by the method of complex potentials", *Int. J. Solids Struct.* **42**:20 (2005), 5491–5512.

[Luo and Weng 1987] H. A. Luo and G. J. Weng, "On Eshelby's inclusion problem in a three-phase spherically concentric solid, and a modification of Mori–Tanaka's method", *Mech. Mater.* **6**:4 (1987), 347–361.

[Meguid and Zhu 1995] S. A. Meguid and Z. H. Zhu, "Stress distribution in dissimilar materials containing inhomogeneities near the interface using a novel finite element", *Finite Elem. Anal. Des.* **20**:4 (1995), 283–298.

[Molchanov et al. 2002] I. N. Molchanov, I. S. Levchenko, N. N. Fedonyuk, A. N. Khimich, and T. V. Chistyakova, "Numerical simulation of the stress concentration in an elastic half-space with a two-layer inclusion", *Prik. Mekh.* **38**:3 (2002), 65–71. In Russian: translated in *Int. Appl. Mech.* **38**:3 (2002), 308–314.

[Mortensen 2007] A. Mortensen, *Concise encyclopedia of composite materials*, 2nd ed., Elsevier, Amsterdam, 2007.

[Moschovidis and Mura 1975] Z. A. Moschovidis and T. Mura, "Two-ellipsoidal inhomogeneities by the equivalent inclusion method", *J. Appl. Mech.* (*ASME*) **42**:4 (1975), 847–852.

[Mura 1987] T. Mura, *Micromechanics of defects in solids*, Martinus Nijhoff, The Hague, 1987.

[Nakasone et al. 2000] Y. Nakasone, H. Nishiyama, and T. Nojiri, "Numerical equivalent inclusion method: a new computational method for analyzing stress fields in and around inclusions of various shapes", *Mater. Sci. Eng. A* **285**:1-2 (2000), 229–238.

[Nemat-Nasser et al. 1982] S. Nemat-Nasser, T. Iwakuma, and M. Hejazi, "On composites with periodic structure", *Mech. Mater.* **1**:3 (1982), 239–267.

[Nemat-Nasser et al. 1993] S. Nemat-Nasser, M. Yu, and M. Hori, "Solids with periodically distributed cracks", *Int. J. Solids Struct.* **30**:15 (1993), 2071–2095.

[Selvadurai 2001] A. P. S. Selvadurai, "On the displacements of an elastic half-space containing a rigid inhomogeneity", *Int. J. Geomech.* (*ASCE*) **1**:2 (2001), 149–174.

[Shewchuk 1994] J. R. Shewchuk, *An introduction to the conjugate gradient method without the agonizing pain*, Carnegie Mellon University, Pittsburgh, PA, 1994.

[Shodja and Roumi 2005] H. M. Shodja and F. Roumi, "Overall behavior of composites with periodic multi-inhomogeneities", *Mech. Mater.* **37**:2-3 (2005), 343–353.

[Shodja and Sarvestani 2001] H. M. Shodja and A. S. Sarvestani, "Elastic fields in double inhomogeneity by the equivalent inclusion method", *J. Appl. Mech.* (*ASME*) **68**:1 (2001), 3–10.

[Shodja et al. 2003] H. M. Shodja, I. Z. Rad, and R. Soheilifard, "Interacting cracks and ellipsoidal inhomogeneities by the equivalent inclusion method", *J. Mech. Phys. Solids* **51**:5 (2003), 945–960.

[Tan et al. 1992] C. L. Tan, Y. L. Gao, and F. F. Afagh, "Anisotropic stress analysis of inclusion problems using the boundary integral equation method", *J. Strain Anal. Eng. Des.* **27**:2 (1992), 67–76.

[Tsuchida and Mura 1983] E. Tsuchida and T. Mura, "The stress field in an elastic half space having a spheroidal inhomogeneity under all-around tension parallel to the plane boundary", *J. Appl. Mech.* (*ASME*) **50**:4a (1983), 807–816.

[Tsuchida et al. 2000] E. Tsuchida, Y. Arai, K. Nakazawa, and I. Jasiuk, "The elastic stress field in a half-space containing a prolate spheroidal inhomogeneity subject to pure shear eigenstrain", *Mater. Sci. Eng. A* **285**:1-2 (2000), 339–345.

[Yu and Kuang 2003] J. H. Yu and Z. B. Kuang, "The stress analysis of an ellipsoidal inhomogeneity in dissimilar media", *Compos. Sci. Technol.* **63**:7 (2003), 955–966.

[Zhou et al. 2009] K. Zhou, W. W. Chen, L. M. Keer, and Q. J. Wang, "A fast method for solving three-dimensional arbitrarily shaped inclusions in a half space", *Comput. Methods Appl. Mech. Eng.* **198**:9-12 (2009), 885–892.

[Zhou et al. 2010] K. Zhou, L. M. Keer, Q. J. Wang, X. Ai, K. Sawamiphakdi, P. Glaws, and M. Paire, "Interactions of multiple arbitrarily-shaped 3D inhomogeneous inclusions in a half space", 2010. Submitted manuscript.

KUN ZHOU: kzhou@ntu.edu.sg
Department of Mechanical Engineering, Northwestern University, 2145 Sheridan Road, Evanston, IL 60208-3100, United States
Current address: *School of Mechanical and Aerospace Engineering, Nanyang Technological University, 50 Nanyang Avenue, Singapore 639798, Singapore*

LEON M. KEER: l-keer@northwestern.edu
Department of Mechanical Engineering, Northwestern University, 2145 Sheridan Road, Evanston, IL 60208-3100, United States

Q. JANE WANG: qwang@northwestern.edu
Department of Mechanical Engineering, Northwestern University, 2145 Sheridan Road, Evanston, IL 60208-3100, United States