

Editor
ROBERT M. McMEEKING

Assistant to the Editor
LIZ MONTANA

APPLIED MECHANICS DIVISION

Executive Committee
(Chair) **K. RAVI-CHANDAR**
D. J. INMAN
T. N. FARRIS
A. J. ROSAKIS
Z. SUO
T. E. TEZDUYAR

Associate Editors
Y. N. ABOUSLEIMAN (2008)
E. M. ARRUDA (2007)
M. R. BEGLEY (2008)
J. CAO (2008)
E. CORONA (2008)
H. ESPINOSA (2007)
N. GHADDAR (2009)
S. GOVINDJEE (2009)
Y. Y. HUANG (2008)
S. KRISHNASWAMY (2008)
K. M. LIECHT (2009)
A. M. MANIATY (2007)
A. MASUD (2009)
I. MEZIC (2009)
M. P. MIGNOLET (2009)
S. MUKHERJEE (2009)
O. M. O'REILLY (2007)
M. OSTOJA-STARZEWSKI (2009)
A. RAMAN (2010)
T. W. SHIELD (2008)
N. S. NAMACHCHIVAYA (2009)
Z. SUO (2009)
T. E. TEZDUYAR (2007)
N. TRIANTAFYLIDIS (2007)
W.-C. WIE (2010)
B. A. YOUNIS (2009)

PUBLICATIONS COMMITTEE

Chair, **BAHRAM RAVANI**

OFFICERS OF THE ASME

President, **SAM Y. ZAMRIK**
Executive Director, **V. R. CARTER**
Treasurer, **T. PESTORIUS**

PUBLISHING STAFF

Managing Director, Publishing
PHILIP DI VIETRO
Manager, Journals
COLIN MCATEER
Production Coordinator
JUDITH SIERANT
Production Assistant
MARISOL ANDINO

Transactions of the ASME, Journal of Applied Mechanics (ISSN 0021-8936) is published bimonthly (Jan., Mar., May, July, Sept., Nov.) by The American Society of Mechanical Engineers, Three Park Avenue, New York, NY 10016. Periodicals postage paid at New York, NY and additional mailing offices. POSTMASTER: Send address changes to Transactions of the ASME, Journal of Applied Mechanics, c/o THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS, 22 Law Drive, Box 2300, Fairfield, NJ 07007-2300. CHANGES OF ADDRESS must be received at Society headquarters seven weeks before they are to be effective. Please send old label and new address.

STATEMENT from By-Laws. The Society shall not be responsible for statements or opinions advanced in papers or printed in its publications (B7.1, Para. 3).

COPYRIGHT © 2007 by The American Society of Mechanical Engineers. For authorization to photocopy material for internal or personal use under those circumstances not falling within the fair use provisions of the Copyright Act, contact the Copyright Clearance Center (CCC), 222 Rosewood Drive, Danvers, MA 01923, tel: 978-750-8400, www.copyright.com. Request for special permission or bulk copying should be addressed to Reprints/Permission Department, Canadian Goods & Services Tax Registration #126148048.

TECHNICAL PAPERS

- 833** Solution for a Semi-Permeable Interface Crack Between Two Dissimilar Piezoelectric Material
Q. Li and Y. H. Chen
- 845** Energy Analysis and Decoupling in Three-Dimensional Impacts of Multibody Systems
Seyed Ali Modarres Najafabadi, József Kövecses, and Jorge Angeles
- 852** Nonlinear Behavior and Critical State of a Penny-Shaped Dielectric Crack in a Piezoelectric Solid
Chun-Ron Chiang and George J. Weng
- 861** Designing Optimal Volume Fractions For Functionally Graded Materials With Temperature-Dependent Material Properties
Florin Bobaru
- 875** Thermoelastic Dynamic Instability (TEDI) in Frictional Sliding of a Half-Space Against a Rigid Non-Conducting Wall
L. Afferrante and M. Ciavarella
- 885** The Girsanov Linearization Method for Stochastically Driven Nonlinear Oscillators
Nilanjan Saha and D. Roy
- 898** Stress-Strain Relationship for Metal Hollow Sphere Materials as a Function of Their Relative Density
D. Karagiozova, T. X. Yu, and Z. Y. Gao
- 908** On the Normal Component of Centralized Frictionless Collision Sequences
Pieter J. Mosterman
- 916** Dynamic Stress of a Circular Cavity Buried in a Semi-Infinite Functionally Graded Material Subjected to Shear Waves
Xue-qian Fang, Chao Hu, and Shan-yi Du
- 923** A Note on the Principle of Maximum Dissipation Rate
F. D. Fischer and J. Svoboda
- 927** Analysis of One-Dimensional and Two-Dimensional Thin Film "Pull-in" Phenomena Under the Influence of an Electrostatic Potential
Gang Duan and Kai-Tak Wan
- 935** Parametric Formulation of the Finite-Volume Theory for Functionally Graded Materials—Part I: Analysis
Marcio A. A. Cavalcante, Severino P. C. Marques, and Marek-Jerzy Pindera
- 946** Parametric Formulation of the Finite-Volume Theory for Functionally Graded Materials—Part II: Numerical Results
Marcio A. A. Cavalcante, Severino P. C. Marques, and Marek-Jerzy Pindera
- 958** The Parametric Resonance Instability in a Drilling Process
Bo-Wun Huang and Jao-Hwa Kuang
- 965** Normality Structures With Thermodynamic Equilibrium Points
Q. Yang, R. K. Wang, and L. J. Xue
- 972** On the One-Parameter Lorentzian Spherical Motions and Euler-Savary Formula
M. Tosun, M. A. Gungor, and I. Okur

(Contents continued on inside back cover)

This journal is printed on acid-free paper, which exceeds the ANSI Z39.48-1992 specification for permanence of paper and library materials. ©™

♻️ 85% recycled content, including 10% post-consumer fibers.

- 978 Computational Examination of the Effect of Material Inhomogeneity on the Necking of Stent Struts Under Tensile Loading
J. P. McGarry, B. P. O'Donnell, P. E. McHugh, E. O'Cearbhaill, and R. M. McMeeking
- 990 Dynamic Failure Behavior of Polycrystalline Alumina Under Impact Loading
Guowen Yao and Zhanfang Liu
- 996 Magnetic Force and Thermal Expansion as Failure Mechanisms of Electrothermal MEMS Actuators Under Electrostatic Discharge Testing
Jonathan D. Weiss
- 1006 A One-Dimensional Theory for the Analysis of Strip Surface Acoustic Waveguides
Shi Chen, Tiantong Tang, and Zhaohong Wang
- 1011 Series Solution of Three-Dimensional Unsteady Laminar Viscous Flow Due to a Stretching Surface in a Rotating Fluid
Yue Tan and Shi-Jun Liao
- 1019 Stoneley Wave Generation in Joined Materials With and Without Thermal Relaxation Due to Thermal Mismatch
L. M. Brock

TECHNICAL BRIEFS

- 1026 Modal Analysis of Nonviscously Damped Beams
S. Adhikari, M. I. Friswell, and Y. Lei
- 1031 Alternative Approaches for the Derivation of Discontinuous Galerkin Methods for Nonlinear Mechanics
L. Noels and R. Radovitzky
- 1037 Effects of Temperature Changes on the Mechanical and Ballistic Responses in Biathlon Shooting
Christelle Grebot and Alain Burtheret
- 1042 Fluid-Structure Interaction Effects in the Dynamic Response of Free-Standing Plates to Uniform Shock Loading
Nayden Kambouchev, Raul Radovitzky, and Ludovic Noels
- 1046 Integral Representation of Energy Release Rate in Graded Materials
Z.-H. Jin and C. T. Sun
- 1049 Elastic Fields Around the Cohesive Zone of a Mode III Crack Perpendicular to a Bimaterial Interface
W. Zhang and X. Deng
- 1053 Lateral Motion of an Axially Moving Tape on a Cylindrical Guide Surface
Bart Raeymaekers and Frank E. Talke

DISCUSSION

- 1057 Discussion: "A Mathematical Model for Frictional Elastic-Plastic Sphere-on-Flat Contacts at Sliding Incipient"
(Chang, L., and Zhang, H., 2007, ASME J. Appl. Mech., 74, pp. 100–106)
Izhak Etsion

The ASME Journal of Applied Mechanics is abstracted and indexed in the following:

Alloys Index, Aluminum Industry Abstracts, Applied Science & Technology Index, Ceramic Abstracts, Chemical Abstracts, Civil Engineering Abstracts, Compendex (The electronic equivalent of Engineering Index), Computer & Information Systems Abstracts, Corrosion Abstracts, Current Contents, EEA (Earthquake Engineering Abstracts Database), Electronics & Communications Abstracts Journal, Engineered Materials Abstracts, Engineering Index, Environmental Engineering Abstracts, Environmental Science and Pollution Management, Fluidex, Fuel & Energy Abstracts, GeoRef, Geotechnical Abstracts, INSPEC, International Aerospace Abstracts, Journal of Ferrocement, Materials Science Citation Index, Mechanical Engineering Abstracts, METADEX (The electronic equivalent of Metals Abstracts and Alloys Index), Metals Abstracts, Nonferrous Metals Alert, Polymers Ceramics Composites Alert, Referativnyi Zhurnal, Science Citation Index, SciSearch (Electronic equivalent of Science Citation Index), Shock and Vibration Digest, Solid State and Superconductivity Abstracts, Steels Alert, Zentralblatt MATH

Solution for a Semi-Permeable Interface Crack Between Two Dissimilar Piezoelectric Material

Q. Li

Y. H. Chen¹

e-mail: yhchen2@mail.xjtu.edu.cn

School of Aerospace,
Xian Jiao-Tong University,
Xian City 710049, P.R. China

A semi-permeable interface crack in dissimilar piezoelectric materials is studied in detail. Attention is focused on the influence induced from the permittivity of the medium inside the crack gap on the near-tip singularity and the crack tip energy release rate (ERR). The Stroh complex variable theory (Stroh, A. N., 1958, Philos. Mag. 3, pp. 625–646; Ting, T. C. T., Int. J. Solids Struct., 22, pp. 965–983) is used to obtain the solution, from which some useful numerical results for 21 kinds of dissimilar piezoelectric materials are calculated. They are combined from seven kinds of commercial piezoelectric ceramics. The distribution of the normal electric displacement component (NEDC) along the interface crack is assumed to be uniform and the corresponding problem is then deduced to a Hilbert problem with an unknown NEDC. Solving the Hilbert problem and determining the near-tip field for each of the 21 bimaterials, we determine the crack tip singularities and find that the crack-tip singularity for a certain combination of two dissimilar piezoelectric materials can be either oscillatory or nonoscillatory when the poling axes of both piezoelectric materials are perpendicular to the interface crack. Energy analyses for PZT-4/BaTiO₃ as a typical nonoscillatory class bimaterial and those for PZT-5H/BaTiO₃ as a typical oscillatory class bimaterial are specially studied in detail under four different conditions: (i) the crack gap is filled with air or vacuum; (ii) the crack gap is filled with silicon oil to avoid discharge; (iii) the crack gap is conducting; and (iv) the electrically impermeable crack. Detailed comparisons are performed among the four cases. We conclude that the different values of the permittivity have no influence on the crack tip singularity but have significant influences on the crack tip ERR under the combined electromechanical loading. We also conclude that the previous investigations under the insulating crack model are incorrect or misleading since the model overestimates the effect of the electric field on the ERR very much and the results of the ERR for the impermeable crack show significant discrepancies from those for the semi-permeable crack. Whereas the previous investigations under the conducting crack model may be accepted in a tolerant way, the results of the ERR show very small discrepancies from those for the semi-permeable crack model, especially when it filled with silicon oil. [DOI: 10.1115/1.2711232]

Keywords: piezoelectric material, semi-permeable interface crack, permittivity, the crack tip ERR, insulating crack, conducting crack

1 Introduction

Interface crack problems in dissimilar anisotropic piezoelectric materials have received considerable attention in the past 20 years [3–16]. Research in this area is motivated by strong engineering demands to design new composite materials with mechanical–electric coupling properties. However, all previous efforts were either under the impermeable crack assumption (charge free) or under the permeable crack assumption. To the present authors' knowledge, there are no attempts in the literature, which account for the permittivity of medium inside an interface crack gap, although the semi-permeable crack (or finite permeable crack) has already been proposed by Parton and Kudryavtsev [17] and Hao and Shen [18] and widely used by many researchers in homogeneous piezoelectric materials with cracks [19–27]. The key feature of the semi-permeable crack is to account for the permittivity of the medium inside the crack gap, from which the electric boundary condition along the crack faces should involve both mechanical quantities and electric quantities and in this way it becomes a

nonlinear model as shown in McMeeking's figures [22–24] and as pointed out by Chen and Hasebe [27]. On one hand, under the impermeable crack assumption, efforts have been done by a number of investigators in the aspect of the interface fracture for piezoelectric bimaterials. For example, Kuo and Barnett [5] studied the interfacial impermeable cracks in bonded piezoelectric half-spaces and Suo et al. [6] solved the problem of an interface impermeable crack between two dissimilar anisotropic piezoelectric media. Qin and Yu [9] studied an arbitrary-oriented plane crack terminating at the interface between dissimilar piezoelectric materials. Beom and Atlury [8] studied the near-tip field and Ru et al. [10] researched electric-field induced interfacial cracking in multilayer electrostrictive for two typical cases; (1) an interface crack lying between an electrode layer and ceramic; and (2) an interface crack with one tip at the embedded electrode edge. However, these results might be misleading as emphasized by McMeeking [22–24] since the unphysical electric boundary condition: the impermeable crack assumption or the charge-free condition was imposed on the crack surfaces without any consideration of the permittivity of the medium inside the crack gap. Dunn [19] and Sosa and Khutoryansky [20] and many others have already found that the impermeable crack assumption would lead to erroneous results and the permittivity of medium inside the crack gap could not be entirely neglected. More recently, Mc-

¹Corresponding author.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received April 28, 2006; final manuscript received September 27, 2006. Review conducted by Zhigang Suo.

Meeking [22–24] pointed out that, in fact, cracks in the engineering are usually filled with air/vacuum, and thus the permittivity of the medium inside the crack is about 1000 times smaller than those of commercial piezoelectric materials, but it is never zero and then the normal electric displacement inside the crack gap must not vanish. On the other hand, some previous researchers used the permeable crack model to study the interface crack. However, this model is only valid in such a way that the crack in piezoelectric ceramics is slit enough for the electric field to penetrate into it. By considering the permittivity of air or vacuum inside the crack gap, the semi-permeable or finite permeable electric boundary condition has been proposed by Parton and Kudryavtsev [17] and Hao and Shen [18]. The key feature of this condition is the correct connection condition across the interface between air or vacuum and piezoelectric ceramic [22–24]. By comparing it with both impermeable and permeable models, the semi-permeable electric boundary condition on the crack faces appears to be a physically reasonable model [21–24] as reviewed by Chen and Hasebe [27]. Although the semi-permeable electric boundary condition has been successfully used to solve the crack problems in homogeneous piezoelectric media, to the authors' knowledge no solutions have been presented in the literature for the interface crack problems between two dissimilar piezoelectric materials.

The goal of this research is to obtain the fundamental solution accounting for the permittivity of the medium inside an interface crack gap between two dissimilar piezoelectric materials. The poling axes of the two dissimilar piezoelectrics are assumed to be perpendicular to the interface. The generalized Stroh formalism proposed by Suo et al. [6] for an impermeable interface crack is used to satisfy the connection condition across the interface between the medium inside the crack gap and each of the two dissimilar piezoelectric materials. In general, for purely elastic material, it is well known that the case of remote loading can be readily reduced to the case of crack-face loading by the superposition principle. The solution can be divided into two parts: the homogeneous solution without the crack and the disturbed solution including the crack subjected to surface tractions. However, when the piezoelectric bimaterial with an interface crack is loaded at infinity, the superposition principle will lead to some confusion because the electric field exists inside the crack and the electric displacement on the crack surface cannot be known beforehand. In the present analysis, we successfully avoid using the superposition principle and directly consider the piezoelectric bimaterials loaded at infinity. After performing some manipulations, the semi-permeable interface crack problem is deduced to a Hilbert problem with an unknown normal electric displacement component on the crack surfaces. After solving the Hilbert problem, the full near-tip fields are determined in the explicit forms and numerical values of the field oscillatory index ε and another index κ are given for 21 types of piezoelectric bimaterials combined, respectively, by seven kinds of piezoelectric materials PZT-4, PZT-5H, PZT-6B, PZT-7A, P-7, BaTiO₃, and PZT-PIC whose material constants were reported in the literature. It is concluded that the singularity at a semi-permeable interface crack tip for a certain combination of dissimilar piezoelectric materials can either be oscillatory or nonoscillatory when the poling axes of both piezoelectric materials are perpendicular to the interface crack. That is, the crack tip singularity may either be governed by $-1/2 \pm i\varepsilon$ (called the ε class bimaterial) or by $-1/2 \pm \kappa$ (called the κ class bimaterial), depending on the combination of the two dissimilar piezoelectric materials. It is concluded that the following 14 kinds of dissimilar piezoelectric materials: PZT-4/PZT-5H, PZT-4/PZT-6B, PZT-4/PZT-7A, PZT-4/P-7, PZT-4/BaTiO₃, PZT-5H/PZT-7, PZT-5H/PZT-PIC 151, PZT-6B/BaTiO₃, PZT-7A/P-7, PZT-7A/BaTiO₃, PZT-7A/PZT-PIC 151, P-7/BaTiO₃, P-7/PZT-PIC 151, and BaTiO₃/PZT-PIC 151 belong to the nonoscillatory singularity or the κ class bimaterial, whereas the following seven kinds: PZT-4/PZT-PIC 151, PZT-5H/PZT-6B, PZT-5H/PZT-7A,

PZT-5H/BaTiO₃, PZT-6B/PZT-7A, PZT-6B/P-7, and PZT-6B/PZT-PIC 151 belong to the oscillatory singularity or the ε class bimaterial. These two kinds of crack tip singularity are different from the classical inverse square root singularity in two different ways. Those in the first only change the magnitude of the crack tip singularity. Those in the second yield the well-known oscillatory singularity [27].

Subsequently, the stress intensity factors and the electric displacement intensity factor (EDIF) at the interface crack tip are obtained, which uniquely characterize the near-tip singular stress fields and singular electric field. By using the crack closure integral, the crack tip energy release rate (ERR) is also given in the explicit forms. In order to provide some useful numerical results, special attention is focused on PZT-4/BaTiO₃ as a typical κ class bimaterial and PZT-5H/BaTiO₃ as a typical ε class bimaterial, respectively. Four different crack models are considered: (i) the crack gap is filled with air or vacuum; (ii) the crack gap is filled with silicon oil to avoid discharge [29,30]; (iii) the crack gap is assumed to be fully permeable, i.e., the so-called permeable crack with infinite large permittivity (e.g., the gap is filled with NaCl solution [31]); and (iv) the electrically impermeable crack. Detailed numerical results show that the permittivity of the medium inside the crack gap has a significant influence on the effect of the applied electric field on the crack tip ERR. As those in homogeneous piezoelectric materials, the impermeable assumption with infinitesimal or negligible permittivity significantly overestimates the effect of the electric field on the crack tip ERR significantly, which yields significant discrepancies of the ERR from those for the semi-permeable crack. Whereas the permeable crack with infinite large permittivity underestimates the effect of the electric field but in a tolerant way, it yields small discrepancies of the ERR from those for the semi-permeable crack, especially when it is filled with silicon oil. What's more, it is seen that the semi-permeable crack filled with silicon oil yields much smaller values of the crack tip ERR than those for the semi-permeable crack filled with air or vacuum under the same magnitude of the mechanical–electric loading. It is concluded that under the electromechanical loading at infinity, the permittivity of the medium inside an interface crack gap plays an important role in evaluating the interface crack stability since it influences the crack-tip ERR significantly, although the permittivity has no influence on the interface crack tip singularity. It should be emphasized that the analysis presented in this paper is a linear piezoelectric analysis, in which the near-tip nonlinearity has been entirely neglected. Obviously, the fracture criterion in dissimilar piezoelectric materials should be concerned with some material microstructure such as the near-tip domain switching or microcracks and the mesoscopic failure mechanism. Subsequent investigations into these topics are absolutely needed.

2 Analytical Solutions and Near-Tip Singularity for an Interface Crack

Consider an interface crack of length $2a$ between two dissimilar piezoelectric materials with the poling direction perpendicular to the interface (see Fig. 1). Assume that the remote uniform mechanical–electric loads, respectively denoted by σ_{21}^∞ , σ_{22}^∞ , σ_{23}^∞ , and E_2^∞ , are acting on the infinitely large piezoelectric bimaterial and along the direction perpendicular to the crack surfaces. All the basic formulations of the generalized Stroh theory [6] are presented in Appendix A. As Parton and Kudryavtsev [17] and Hao and Shen [18] proposed, the permittivity of the medium inside the crack gap, denoted by ε_0 , would play an important role in piezoelectric fracture. In this section, as an initial attempt, the distribution of the normal electric displacement component along the interface crack induced from the permittivity is assumed to no longer be zero as the impermeable crack model; rather it is assumed to be uniform and denoted by D_2^0 . This assumption is valid when the poling directions of both materials are perpendicular to

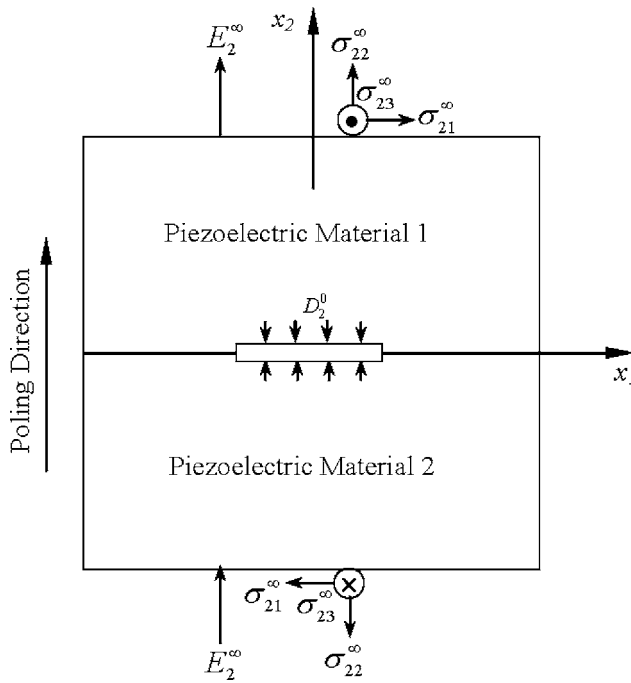


Fig. 1 A semi-permeable interface crack of length $2a$ between two dissimilar piezoelectric materials subjected to the far field mechanical and electric loading

the interface crack. The finite permeable crack or the semi-permeable crack can be formulated as follows (sometimes we abbreviated it as the PKHS crack model [27])

$$D_2^0(u_2^+ - u_2^-) = -\varepsilon_v(\phi^+ - \phi^-) \quad |x_1| < a \quad (1a)$$

$$D_2^+ = D_2^-, \quad \sigma_{2j}^+ = \sigma_{2j}^- = 0 \quad |x_1| < a \quad (1b)$$

where the superscripts $+$ and $-$ refer to the upper and lower crack surface, respectively; u_2 is the displacement component normal to the crack face; ϕ is the electrical potential along the crack surfaces; and ε_v is the permittivity of the medium inside the interface crack gap.

As pointed out by Suo et al. [6], a crack may be thought of as a low-capacitance medium carrying a potential drop. Thus, the electric field inside the gap could be simply the potential difference divided by the gap width. It is well known that the electrically permeable crack and the impermeable crack are two limitations of Eq. (1) when setting the permittivity ε_v to be infinitely large so that $\phi^+ = \phi^-$, or infinitely small so that $D_2^+ = D_2^- = 0$, respectively.

For the problem shown in Fig. 1, the continuity of the generalized stresses $\Sigma(x_1)$ across the x_1 axis on both the bonded and cracked segments requires that

$$\mathbf{B}_1 \mathbf{f}_1^+(x_1) + \bar{\mathbf{B}}_1 \bar{\mathbf{f}}_1^-(x_1) = \mathbf{B}_2 \mathbf{f}_2^-(x_1) + \bar{\mathbf{B}}_2 \bar{\mathbf{f}}_2^+(x_1), \quad |x_1| < \infty \quad (2)$$

where the subscript 1 and 2 separately indicate that the quantities are defined in materials 1 and 2; the constant complex vectors $\mathbf{B}_1$ and $\mathbf{B}_2$ and unknown complex vector functions; and $\mathbf{f}_1(z)$ and $\mathbf{f}_2(z)$ are given in Eqs. (A9) and (A4) of Appendix A, respectively.

Thus, the above formulation can be rearranged in the following form

$$\mathbf{B}_1 \mathbf{f}_1^+(x_1) - \bar{\mathbf{B}}_2 \bar{\mathbf{f}}_2^+(x_1) = \mathbf{B}_2 \mathbf{f}_2^-(x_1) - \bar{\mathbf{B}}_1 \bar{\mathbf{f}}_1^-(x_1) \quad |x_1| < \infty \quad (3)$$

Define a new analytical function as

$$\Psi(z) = \begin{cases} \mathbf{B}_1 \mathbf{f}_1'(z) - \bar{\mathbf{B}}_2 \bar{\mathbf{f}}_2'(z), & x_2 > 0 \\ \mathbf{B}_2 \mathbf{f}_2'(z) - \bar{\mathbf{B}}_1 \bar{\mathbf{f}}_1'(z), & x_2 < 0 \end{cases} \quad (4)$$

and reduce Eq. (3) to

$$\Psi^+(x_1) - \Psi^-(x_1) = 0 \quad |x_1| < \infty \quad (5)$$

The solution of Eq. (5) is then given as follows [32]

$$\Psi(z) = \Psi(\infty) \quad (6)$$

where

$$\Psi(\infty) = \mathbf{B}_1 \mathbf{f}_1'(\infty) - \bar{\mathbf{B}}_2 \bar{\mathbf{f}}_2'(\infty) = \mathbf{B}_2 \mathbf{f}_2'(\infty) - \bar{\mathbf{B}}_1 \bar{\mathbf{f}}_1'(\infty) \quad (7)$$

Define the jump of the generalized displacements across the interface

$$\Delta \mathbf{u}(x_1) = \mathbf{u}_1^+(x_1) - \mathbf{u}_2^-(x_1) \quad (8)$$

where $\mathbf{u} = (u_1, u_2, u_3, \phi)^T$ denotes the generalized displacements for piezoelectrics under generalized plane strain (see Eq. (A2) in Appendix A).

Combining Eqs. (A2), (4), and (6) yields

$$i \Delta \mathbf{u}'(x_1) = \mathbf{H} \mathbf{B}_1 \mathbf{f}_1'^+(x_1) - \bar{\mathbf{H}} \mathbf{B}_2 \mathbf{f}_2'^-(x_1) - (\bar{\mathbf{Y}}_2 - \bar{\mathbf{Y}}_1) \Psi(\infty) \quad (9)$$

where $\mathbf{H}$, $\mathbf{Y}_2$, and $\mathbf{Y}_1$ are constant complex vectors (see Eq. (A10) in Appendix A). For piezoelectric bimetals as shown in Fig. 1, the continuity of displacements across the bonded interface requires that a function defined as

$$\mathbf{g}(z) = \begin{cases} \mathbf{B}_1 \mathbf{f}_1'(z) & x_2 > 0 \\ \mathbf{H}^{-1} \bar{\mathbf{H}} \mathbf{B}_2 \mathbf{f}_2'(z) + \mathbf{H}^{-1} (\bar{\mathbf{Y}}_2 - \bar{\mathbf{Y}}_1) \Psi(\infty) & x_2 < 0 \end{cases} \quad (10)$$

where $\mathbf{g}(z)$ is analytic in the whole plane except on the cracking line.

Using Eqs. (A3), (9), and (10), we can obtain

$$i \Delta \mathbf{u}'(x_1) = \mathbf{H} [\mathbf{g}^+(x_1) - \mathbf{g}^-(x_1)] \quad (11)$$

$$\Sigma(x_1) = \mathbf{g}^+(x_1) + \bar{\mathbf{H}}^{-1} \mathbf{H} \mathbf{g}^-(x_1) - \mathbf{K}^0 \quad (12)$$

where $\mathbf{K}^0 = \bar{\mathbf{H}}^{-1} (\mathbf{Y}_2 + \bar{\mathbf{Y}}_2) \Psi(\infty)$. On the interface crack faces, the generalized stresses $\Sigma(x_1) = \mathbf{T}^0 = (0, 0, 0, D_2^0)^T$ refer the traction-free conditions and the uniform normal electric displacement along the semi-permeable interface crack as shown in Fig. 1. Equation (12) yields a nonhomogeneous Riemann-Hilbert problem as expressed in the following form

$$\mathbf{g}^+(x_1) + \bar{\mathbf{H}}^{-1} \mathbf{H} \mathbf{g}^-(x_1) = \mathbf{K}^0 + \mathbf{T}^0 \quad |x_1| < a \quad (13)$$

As a special case, when the piezoelectric bimaterial continuum degenerates to a homogeneous one, the matrix $\mathbf{H}$ in Eq. (A10) can be real. Thus, Eq. (13) can be degenerated into

$$\mathbf{g}^+(x_1) + \mathbf{g}^-(x_1) = \Psi(\infty) + \mathbf{T}^0 \quad |x_1| < a \quad (14)$$

The solving procedure follows those developed by England [33] and Suo et al. [6], we can then obtain

$$\mathbf{g}(z) = \frac{\Psi(\infty) + \mathbf{T}^0}{2} \left(1 - \frac{z}{\sqrt{z^2 - a^2}} \right) + \frac{\mathbf{c}_0 + \mathbf{c}_1 z + \cdots + \mathbf{c}_n z^n}{\sqrt{z^2 - a^2}} \quad (15)$$

where $\mathbf{c}_n = [c_n^{(1)}, c_n^{(2)}, c_n^{(3)}, c_n^{(4)}]$; whereas D_2^0 and the constant vectors $\mathbf{c}_n$ are unknown. To find $\mathbf{c}_n$, we let $z \rightarrow \infty$ in Eq. (15) and find that the condition of the stresses being bounded at infinity leads to

$$\mathbf{c}_1 = \mathbf{g}(\infty) = \mathbf{B} \mathbf{f}'(\infty); \quad \mathbf{c}_n = 0 \quad (n > 1) \quad (16)$$

The remaining unknown constants $\mathbf{c}_0$ can be determined by using the single-valued conditions of mechanical displacements and electric potential. Considering Eq. (11), these conditions require

$$\int_{L_c} \mathbf{g}(z) dz = 0 \quad (17)$$

where L_c is a clockwise closed-contour encircled the interface crack. Substituting Eq. (15) into Eq. (17) results in $c_0=0$.

Thus, the complex potential vector $\mathbf{g}(z)$ in Eq. (15) can be finally expressed as

$$\mathbf{g}(z) = \frac{\mathbf{\Psi}(\infty) + \mathbf{T}^0}{2} + \frac{\mathbf{T}^\infty - \mathbf{T}^0}{2} \frac{z}{\sqrt{z^2 - a^2}} \quad (18)$$

where $\mathbf{T}^\infty = \mathbf{B}\mathbf{f}'(\infty) + \bar{\mathbf{B}}\bar{\mathbf{f}}'(\infty) = (\sigma_{21}^\infty, \sigma_{22}^\infty, \sigma_{23}^\infty, D_2^\infty)^T$.

Considering the singular parts of the generalized stress fields at the crack tip as the polar coordinates $(r, \theta) \rightarrow 0$, we obtain from Eqs. (12) and (18)

$$\Sigma(r) = \frac{\sqrt{\pi a}}{\sqrt{2\pi r}} (\mathbf{T}^\infty - \mathbf{T}^0) \quad (19)$$

The jump of the mechanical displacements and the drop of electric potential along the crack gap $(-a \leq x_1 \leq a)$ can be obtained from Eqs. (11) and (18)

$$\Delta \mathbf{u} = \mathbf{u}^+ - \mathbf{u}^- = \mathbf{H}(\mathbf{T}^\infty - \mathbf{T}^0) \sqrt{a^2 - x_1^2} \quad (20)$$

In order to determine the normal electric displacement component D_2^0 along the crack surfaces induced from the permittivity of the medium inside the crack gap as Parton and Kudryavtsev [17] and Hao and Shen [18] discussed, the semi-permeable electric boundary condition (1a) should be used at present. By substituting Eq. (20) into Eq. (1a), we can obtain

$$D_2^0 = -\varepsilon_v \frac{\phi^+ - \phi^-}{u_2^+ - u_2^-} = \frac{H_{41}\sigma_{21}^\infty + H_{42}\sigma_{22}^\infty + H_{43}\sigma_{23}^\infty + H_{44}(D_2^\infty - D_2^0)}{H_{21}\sigma_{21}^\infty + H_{22}\sigma_{22}^\infty + H_{23}\sigma_{23}^\infty + H_{24}(D_2^\infty - D_2^0)} \quad (21)$$

where H_{ij} ($i, j=1, 2, 3, 4$) are elements of the 4×4 matrix $\mathbf{H}$ whose definition has been given in Eq. (A10) of Appendix A.

On one hand, for a homogenous piezoelectric material with real material matrix $\mathbf{H}$, the singularity at the semi-permeable crack tip is the well-known classical *inverse square root singularity* and the stress intensity factors K_I , K_{II} , and K_{III} and the electric displacement intensity factor K_D can be expressed as

$$K_I = \sqrt{\pi a} \sigma_{22}^\infty, \quad K_{II} = \sqrt{\pi a} \sigma_{21}^\infty, \quad K_{III} = \sqrt{\pi a} \sigma_{23}^\infty, \quad K_D = \sqrt{\pi a} (D_2^\infty - D_2^0) \quad (22)$$

At this moment, the crack tip ERR in a cracked homogenous piezoelectric material can be obtained by using the crack closure integral. It should be emphasized that for a semi-permeable crack the crack tip ERR and the total potential ERR are different as pointed out by McMeeking [24]. With the right crack tip extending by a small amount δa , the crack tip ERR can be expressed as

$$G = \lim_{\delta a \rightarrow 0} \frac{1}{2\delta a} \int_0^{\delta a} \Sigma^T(r) \Delta \mathbf{u}(\delta a - r) dr \quad (23)$$

By substituting Eqs. (19) and (20) into Eq. (23), together with Eq. (22), the crack tip ERR in terms of the stress intensity factors is obtained for the homogenous piezoelectric material

$$G = \frac{1}{4} \mathbf{K}^T \mathbf{H} \mathbf{K} \quad (24)$$

where $\mathbf{K} = (K_{II}, K_I, K_{III}, K_D)^T$. In deriving Eq. (24), the identity

$$\int_0^1 t^q (1-t)^{-q} dt = q \pi / \sin q \pi \quad |\operatorname{Re}(q)| < 1 \quad (25)$$

has been used with $q = -1/2$.

On the other hand, for the dissimilar piezoelectric materials with complex material matrix $\mathbf{H}$, the homogeneous Hilbert problem of Eq. (13) should be considered in more detail, which can be expressed in the following form

$$\mathbf{g}^+(x_1) + \bar{\mathbf{H}}^{-1} \mathbf{H} \mathbf{g}^-(x_1) = 0 \quad |x_1| < a \quad (26)$$

A general solution can be sought in the form [6]

$$\mathbf{h}(z) = \mathbf{w} z^{-1/2 + i\varepsilon_\alpha} \quad (27)$$

where $\mathbf{w}$ is now a four-element column vector and ε_α an arbitrary number, both are to be determined. Substituting Eq. (27) into Eq. (26) gives an eigenvalue problem

$$\bar{\mathbf{H}} \mathbf{w} = e^{2\pi i \varepsilon_\alpha} \mathbf{H} \mathbf{w} \quad (28)$$

By separating the matrix $\mathbf{H}$ into a real part $\mathbf{D}$ and an imaginary $\mathbf{W}$, Eq. (28) can be rewritten as

$$(\mathbf{D}^{-1} \mathbf{W} + i \eta \mathbf{I}) \mathbf{w} = 0 \quad (29)$$

where

$$\eta = -\tanh(\pi \varepsilon_\alpha), \quad \text{or} \quad \varepsilon_\alpha = -\frac{1}{\pi} \tanh^{-1}(\eta) = \frac{1}{2\pi} \ln \frac{1-\eta}{1+\eta} \quad (30)$$

It is well known that the characteristic value problem Eq. (29) leads to the following characteristic equation

$$\|\mathbf{D}^{-1} \mathbf{W} + i \eta \mathbf{I}\| = \eta^4 + 2b\eta^2 + c = 0 \quad (31)$$

where $\|\cdot\|$ denotes the determinant of a matrix and

$$b = \frac{1}{4} \operatorname{tr}[(\mathbf{D}^{-1} \mathbf{W})^2], \quad c = \|\mathbf{D}^{-1} \mathbf{W}\| \quad (32)$$

By solving the characteristic equation Eq. (31), the roots of Eq. (31) can then be expressed as

$$\eta_{1,2} = \pm [(b^2 - c)^{1/2} - b]^{1/2} \quad (33a)$$

$$\eta_{3,4} = \pm i[(b^2 - c)^{1/2} + b]^{1/2} \quad (33b)$$

It is seen from Eqs. (33a) and (33b) that one pair of roots is real and the other is purely imaginary. Thus, we can denote the four roots of ε_α corresponding to η_α ($\alpha=1, 2, 3, 4$) as ε , $-\varepsilon$, $i\kappa$, and $-i\kappa$, respectively, where both ε and κ are real numbers. After doing this, we get

$$\varepsilon = \frac{1}{\pi} \tanh^{-1}[(b^2 - c)^{1/2} - b]^{1/2} \quad (34a)$$

$$\kappa = \frac{1}{\pi} \tanh^{-1}[(b^2 - c)^{1/2} + b]^{1/2} \quad (34b)$$

Therefore, the associated linear independent eigenvectors $\mathbf{w}_\alpha$ ($\alpha=1, 2, 3, 4$) can then be determined by the characteristic value problem Eq. (29), and the four eigenpairs should have the following structures

$$(\varepsilon, \mathbf{w}_1), \quad (-\varepsilon, \mathbf{w}_2), \quad (i\kappa, \mathbf{w}_3), \quad (-i\kappa, \mathbf{w}_4) \quad (35)$$

Here, we should note the recent work by Ou and Wu [34] who studied an impermeable interface crack in dissimilar piezoelectric materials and have shown an interesting singularity theorem. They found that the two singularity parameters ε and κ cannot be non-zero simultaneously for all transversely isotropic piezoelectric bimaterial combinations selected by them. This leads to the classification of piezoelectric bimaterials, i.e., the bimaterials can be divided into two groups: the κ -class bimaterials with vanishing ε and nonoscillating singularity, and the ε -class bimaterials with vanishing κ and oscillating singularity. However, no one in the literature has accounted for the fact that whether the interesting singularity theorem [34,35] holds in the semi-permeable interface crack. In other words, whether and how the permittivity of the

Table 1 Material constants of typical piezoelectric ceramics

Material constants	PZT-4	PZT-5H	PZT-6B	PZT-7A	P-7	BaTiO ₃	PZT-PIC 151
c_{11} (10^{10} N·m ⁻²)	13.9	12.6	16.8	14.8	13.0	15.0	11.0
c_{12} (10^{10} N·m ⁻²)	7.78	5.50	6.00	7.62	8.30	6.60	6.3
c_{13} (10^{10} N·m ⁻²)	7.43	5.30	6.00	7.42	8.30	6.60	6.4
c_{33} (10^{10} N·m ⁻²)	11.3	11.7	16.3	13.1	11.9	14.6	10.0
c_{44} (10^{10} N·m ⁻²)	2.56	3.53	2.71	2.54	2.50	4.4	2.0
e_{13} (C·m ⁻²)	-6.98	-6.50	-0.90	-2.10	-10.3	-4.35	-9.6
e_{33} (C·m ⁻²)	13.8	23.3	7.10	9.50	14.7	17.5	15.1
e_{15} (C·m ⁻²)	13.4	17.0	4.60	9.70	13.5	11.4	12.0
χ_{11} (10^{-10} C·(V·m) ⁻¹)	60.0	151	36.0	81.1	171	98.7	98.2
χ_{33} (10^{-10} C·(V·m) ⁻¹)	54.7	130	34.0	73.5	186	112	75.4

medium inside the interface crack gap influences the two singular parameters ε and κ for a certain combination of two dissimilar piezoelectric materials are still unclear to date.

Motivated by strong engineering demands to design new piezoelectric composite materials which are always made of two phase piezoelectric materials, the singularity analyses at a semi-permeable interface crack tip with the singularity index ε or κ and the associated linear independent eigenvector matrices $\mathbf{w}$ are absolutely necessary. In this section, we use seven piezoelectric materials: PZT-4, PZT-5H, PZT-6B, PZT-7A, P-7, BaTiO₃, and PZT-PIC 151 to construct 21 kinds of dissimilar piezoelectric materials: PZT-4/PZT-5H, PZT-4/PZT-6B, PZT-4/PZT-7A, PZT-4/P-7, PZT-4/BaTiO₃, PZT-4/PZT-PIC 151, PZT-5H/PZT-6B, PZT-5H/PZT-7A, PZT-5H/PZT-7, PZT-5H/BaTiO₃, PZT-5H/PZT-PIC 151, PZT-6B/PZT-7A, PZT-6B/P-7, PZT-6B/BaTiO₃, PZT-6B/PZT-PIC 151, PZT-7A/P-7, PZT-7A/BaTiO₃, PZT-7A/PZT-PIC 151, P-7/BaTiO₃, P-7/PZT-PIC 151, and BaTiO₃/PZT-PIC 151. The material constants of the seven kinds of piezoelectric materials are listed in Table 1. Numerical values of the singularity index ε and κ for the 21 kinds of bimaterials are computed by Eqs. (34a) and (34b) and listed in Appendix B, whereas those for the matrices $\mathbf{w}$ are determined by Eq. (29) and also listed in Appendix B. It is found that when the poling axes of the two dissimilar piezoelectric materials are perpendicular to the interface crack, there is no coexistence of the singular parameters ε and κ at the semi-permeable interface crack tip for all 21 kinds of dissimilar piezoelectric combinations. This means that the singularity for a certain combination of dissimilar piezoelectric materials can either be oscillatory or nonoscillatory. That is, the crack tip singularity may either be governed by $-1/2 \pm i\varepsilon$ or by $-1/2 \pm \kappa$, depending on the combination of the two dissimilar piezoelectric materials. As a result, the singularity theorem derived by Ou and Wu [34] for an impermeable interface crack in dissimilar piezoelectric materials has been proven to still be valid in the present semi-permeable interface crack. This means that the permittivity has no influence on the interface crack singularity. Indeed, piezoelectric bimaterials with a semi-permeable interface crack should be divided into two groups: the κ -class bimaterials with $\varepsilon=0$ that show nonoscillating singularity and the ε -class bimaterials with $\kappa=0$ that show well-known oscillating singularity. It is seen from Appendix B that the 14 kinds of dissimilar piezoelectric materials: PZT-4/PZT-5H, PZT-4/PZT-6B, PZT-4/PZT-7A, PZT-4/P-7, PZT-4/BaTiO₃, PZT-5H/PZT-7, PZT-5H/PZT-PIC 151, PZT-6B/BaTiO₃, PZT-7A/P-7, PZT-7A/BaTiO₃, PZT-7A/PZT-PIC 151, P-7/BaTiO₃, P-7/PZT-PIC 151, and BaTiO₃/PZT-PIC 151 belong to the nonoscillatory singularity, whereas the seven kinds: PZT-4/PZT-PIC 151, PZT-5H/PZT-6B, PZT-5H/PZT-7A, PZT-5H/BaTiO₃, PZT-6B/PZT-7A, PZT-6B/P-7, and PZT-6B/PZT-PIC 151 belong to the oscillatory singularity.

Furthermore, we should consider the inherent relations among eigenvectors in the present case. When both ε and κ are nonzero,

Ting [2] and Suo et al. [6] had already concluded that the eigenvectors $\mathbf{w}_\alpha$ satisfy certain orthogonal relations based on the four distinct singularity parameters ε_α . However, at present ε and κ are not nonzero simultaneously and the singularity analyses will lead to the repeated eigenvalue either for the κ -class or the ε -class bimaterials. Hence some new inherent relations should be introduced.

After some manipulations, we find that for the ε -class piezoelectric bimaterials, the associated eigenvectors satisfy

$$\bar{\mathbf{H}}\mathbf{w}_1 = e^{2\pi\varepsilon}\mathbf{H}\mathbf{w}_1, \quad \bar{\mathbf{H}}\mathbf{w}_2 = e^{-2\pi\varepsilon}\mathbf{H}\mathbf{w}_2, \quad \bar{\mathbf{H}}\mathbf{w}_3 = \mathbf{H}\mathbf{w}_3, \quad \bar{\mathbf{H}}\mathbf{w}_4 = \mathbf{H}\mathbf{w}_4 \quad (36)$$

from which it can be verified that $\mathbf{w}_\alpha$ satisfy the following inherent relations

$$\begin{bmatrix} \mathbf{w}_1^T \\ \mathbf{w}_2^T \\ \mathbf{w}_3^T \\ \mathbf{w}_4^T \end{bmatrix} \mathbf{H}[\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4] = \begin{bmatrix} 0 & \mathbf{w}_1^T \mathbf{H} \mathbf{w}_2 & 0 & 0 \\ \mathbf{w}_2^T \mathbf{H} \mathbf{w}_1 & 0 & 0 & 0 \\ 0 & 0 & \mathbf{w}_3^T \mathbf{H} \mathbf{w}_3 & 0 \\ 0 & 0 & 0 & \mathbf{w}_4^T \mathbf{H} \mathbf{w}_4 \end{bmatrix} \quad (37)$$

However, for the κ -class piezoelectric bimaterials, the associated eigenvectors satisfy

$$\bar{\mathbf{H}}\mathbf{w}_1 = \mathbf{H}\mathbf{w}_1, \quad \bar{\mathbf{H}}\mathbf{w}_2 = \mathbf{H}\mathbf{w}_2, \quad \bar{\mathbf{H}}\mathbf{w}_3 = e^{2\pi i\kappa}\mathbf{H}\mathbf{w}_3, \quad \bar{\mathbf{H}}\mathbf{w}_4 = e^{-2\pi i\kappa}\mathbf{H}\mathbf{w}_4 \quad (38)$$

from which it can be proved that there are the following inherent relations between the eigenvectors $\mathbf{w}_\alpha$

$$\begin{bmatrix} \mathbf{w}_1^T \\ \mathbf{w}_2^T \\ \mathbf{w}_3^T \\ \mathbf{w}_4^T \end{bmatrix} \mathbf{H}[\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4] = \begin{bmatrix} \mathbf{w}_1^T \mathbf{H} \mathbf{w}_2 & 0 & 0 & 0 \\ 0 & \mathbf{w}_2^T \mathbf{H} \mathbf{w}_1 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{w}_3^T \mathbf{H} \mathbf{w}_4 \\ 0 & 0 & \mathbf{w}_4^T \mathbf{H} \mathbf{w}_3 & 0 \end{bmatrix} \quad (39)$$

Obviously, the above inherent relations obtained here are different from those obtained by Tong [2] and Suo et al. [6].

Now, we study the present problems in more detail to obtain the full-field solution. In order to solve the nonhomogeneous Hilbert Eq. (13), it is convenient to display physical quantities in an ei-

genvector representation. The potential functions $\mathbf{g}(z)$ and constant vector $\mathbf{T}$ can be expressed by the eigenvectors $\mathbf{w}$ as

$$\mathbf{g}(z) = h_1(z)\mathbf{w}_1 + h_2(z)\mathbf{w}_2 + h_3(z)\mathbf{w}_3 + h_4(z)\mathbf{w}_4 \quad (40)$$

$$\mathbf{T} = \mathbf{K}^0 + \mathbf{T}^0 = t_1\mathbf{w}_1 + t_2\mathbf{w}_2 + t_3\mathbf{w}_3 + t_4\mathbf{w}_4 \quad (41)$$

Substituting Eqs. (40) and (41) into Eq. (13), we obtain the following decoupled Hilbert equations

$$h_\alpha^+(x_1) + e^{2\pi\epsilon_\alpha} h_\alpha^-(x_1) = t_\alpha \quad (\alpha = 1, 2, 3, 4) \quad (42)$$

The solving procedure of Eq. (42) follows those developed by England [33] and Suo et al. [6], which yields

$$h_\alpha(z) = \frac{t_\alpha}{1 + e^{2\pi\epsilon_\alpha}} \left[1 - \left(\frac{z-a}{z+a} \right)^{-i\epsilon_\alpha} \frac{z - 2i\epsilon_\alpha a}{\sqrt{z^2 - a^2}} \right] + \left(\frac{z-a}{z+a} \right)^{-i\epsilon_\alpha} \frac{c_0^\alpha + c_1^\alpha z + \dots + c_n^\alpha z^n}{\sqrt{z^2 - a^2}} \quad (43)$$

By considering Eqs. (41) and (43), Eq. (40) gives

$$\mathbf{g}(z) = \mathbf{w} \left\langle \left\langle \frac{1}{1 + e^{2\pi\epsilon_\alpha}} \left[1 - \left(\frac{z-a}{z+a} \right)^{-i\epsilon_\alpha} \frac{z - 2i\epsilon_\alpha a}{\sqrt{z^2 - a^2}} \right] \right\rangle \right\rangle \mathbf{w}^{-1} (\mathbf{K}^0 + \mathbf{T}^0) + \mathbf{w} \left\langle \left\langle \left(\frac{z-a}{z+a} \right)^{-i\epsilon_\alpha} \right\rangle \right\rangle \frac{\mathbf{c}_0 + \mathbf{c}_1 z + \dots + \mathbf{c}_n z^n}{\sqrt{z^2 - a^2}} \quad (44)$$

where $\langle\langle\cdot\rangle\rangle$ indicates the diagonal matrix in which each component is varied according the Greek index α ; $\mathbf{c}_n = (c_n^1, c_n^2, c_n^3, c_n^4)^T$; whereas the normal component of electric displacement along the interface crack D_2^0 and the constant vectors $\mathbf{c}_n$ in Eq. (44) are unknown. To find $\mathbf{c}_n$ it is convenient to use the asymptotic features of Eq. (44) by taking the limit as $z \rightarrow \infty$ and using the far-field uniform finite loading condition, we obtain

$$\mathbf{c}_1 = \left\langle \left\langle \frac{1}{1 + e^{2\pi\epsilon_\alpha}} \right\rangle \right\rangle \mathbf{w}^{-1} (\mathbf{T}^\infty + \mathbf{K}^0); \quad \mathbf{c}_n = 0 \quad (n > 1) \quad (45)$$

In deriving Eq. (45), the following relation has been used

$$\mathbf{w}^{-1} (\mathbf{I} + \bar{\mathbf{H}}^{-1} \mathbf{H}) \mathbf{w} = \langle\langle 1 + e^{2\pi\epsilon_\alpha} \rangle\rangle \quad (46)$$

It should be noted that the determination of the remaining unknown constants $\mathbf{c}_0$ in Eq. (44) needs the use of the single-valued condition of the generalized displacement

$$\int_{L_c} \Delta u(z) dz = 0 \quad (47)$$

Substituting Eqs. (11) and (44) into Eq. (47), we obtain

$$\mathbf{c}_0 = \left\langle \left\langle \frac{-2i\epsilon_\alpha a}{1 + e^{2\pi\epsilon_\alpha}} \right\rangle \right\rangle \mathbf{w}^{-1} (\mathbf{T}^\infty + \mathbf{K}^0) \quad (48)$$

Thus, the full mechanical-electric field for the semi-permeable crack between two dissimilar piezoelectric materials can be expressed as follows

$$\Sigma(z) = \mathbf{T}^0 + \mathbf{w} \left\langle \left\langle \left(\frac{z-a}{z+a} \right)^{-i\epsilon_\alpha} \frac{z - 2i\epsilon_\alpha a}{\sqrt{z^2 - a^2}} \right\rangle \right\rangle \mathbf{w}^{-1} (\mathbf{T}^\infty - \mathbf{T}^0) \quad (49)$$

The jump of the generalized displacement (u_1 , u_2 , u_3 , and ϕ) across the interface crack surfaces ($-a \leq x_1 \leq a$) is then given by

$$\Delta \mathbf{u}(x_1) = \mathbf{H} \mathbf{w} \langle\langle \exp(-\pi\epsilon_\alpha)(a-x_1)^{1/2-i\epsilon_\alpha}(a+x_1)^{1/2+i\epsilon_\alpha} \rangle\rangle \mathbf{w}^{-1} (\mathbf{T}^\infty - \mathbf{T}^0) \quad (50)$$

It should be noted that, in general, for all most practical engineering piezoelectric bimetals, the absolute values of ϵ_α are much smaller than 1/2, varying in the range between 10^{-3} and 10^{-2} (see Appendix B). Thus, the values of ϵ_α comparing 1/2 are

negligible and the jump of the mechanical displacement and the drop of the electric potential across the crack surfaces can be approximately determined from Eq. (50)

$$u_2^+ - u_2^- \approx \mathbf{H}_2 \mathbf{w} \langle\langle \exp(-\pi\epsilon_\alpha) \rangle\rangle \mathbf{w}^{-1} (\mathbf{T}^\infty - \mathbf{T}^0) \sqrt{a^2 - (x_1)^2} \quad (51)$$

$$\phi^+ - \phi^- \approx \mathbf{H}_4 \mathbf{w} \langle\langle \exp(-\pi\epsilon_\alpha) \rangle\rangle \mathbf{w}^{-1} (\mathbf{T}^\infty - \mathbf{T}^0) \sqrt{a^2 - (x_1)^2} \quad (52)$$

where $\mathbf{H}_2$ and $\mathbf{H}_4$ are the second and forth rows of $\mathbf{H}$, respectively.

By substituting Eqs. (51) and (52) into the semi-permeable electric boundary condition Eq. (1a), the normal component of electric displacement inside the crack D_2^0 can be derived by

$$D_2^0 = -\epsilon_v \frac{\phi^+ - \phi^-}{u_2^+ - u_2^-} = -\epsilon_v \frac{\mathbf{H}_2 \mathbf{w} \langle\langle \exp(-\pi\epsilon_\alpha) \rangle\rangle \mathbf{w}^{-1} (\mathbf{T}^\infty - \mathbf{T}^0)}{\mathbf{H}_4 \mathbf{w} \langle\langle \exp(-\pi\epsilon_\alpha) \rangle\rangle \mathbf{w}^{-1} (\mathbf{T}^\infty - \mathbf{T}^0)} \quad (53)$$

The singular stress fields along the bonded interface near the semi-permeable interface crack tip can be derived as the polar coordinate system $(r, \theta) \rightarrow 0$

$$\Sigma(r) = \frac{1}{\sqrt{2\pi r}} \mathbf{Y}(r^{-i\epsilon_\alpha}) (\mathbf{T}^\infty - \mathbf{T}^0) \quad (54)$$

where

$$\mathbf{Y}(r^{-i\epsilon_\alpha}) \mathbf{w} \left\langle \left\langle \sqrt{\pi a} \left(\frac{r}{2a} \right)^{-i\epsilon_\alpha} (1 - 2i\epsilon_\alpha) \right\rangle \right\rangle \mathbf{w}^{-1} \quad (55)$$

Since there is no coexistence of the singular parameters ϵ and κ in the semi-permeable interface crack tip for all dissimilar piezoelectric combinations (see Appendix B), it can be concluded from Eq. (54) that the singularity of near-tip fields for a certain combination of dissimilar piezoelectric materials can either be oscillatory or nonoscillatory when the poling axes of both piezoelectric materials are perpendicular to the interface crack. That is, the crack tip singularity may either be governed by $-1/2 \pm i\epsilon$ or by $-1/2 \pm \kappa$, depending on the combination of the two dissimilar piezoelectric materials. Thus the vector of real-valued stress and electric displacement intensity factors, which uniquely characterize the singular fields at the semi-permeable interface crack tip between two dissimilar piezoelectric materials with complex material matrix $\mathbf{H}$, can be defined as [8]

$$\mathbf{K} = \lim_{r \rightarrow 0} \sqrt{2\pi r} \mathbf{Y}(r^{i\epsilon_\alpha}) \Sigma(r) \quad (56)$$

where $\mathbf{K} = (K_{II}, K_I, K_{III}, K_D)^T$. The intensity factor $\mathbf{K}$ may be considered as an extension of the elastic version proposed by Wu [35] and Qu and Li [36].

By substituting the above near-tip stress fields and the jump of generalized displacement into the crack closure integral Eq. (23), the crack tip ERR for the semi-permeable interface crack can be obtained as

$$G = \frac{a}{2} (\mathbf{T}^\infty - \mathbf{T}^0)^T \mathbf{H} \mathbf{w} \langle\langle \chi_\alpha \rangle\rangle \mathbf{w}^{-1} (\mathbf{T}^\infty - \mathbf{T}^0) \quad (57)$$

where

$$\chi_\alpha = \frac{(-\frac{1}{2} + i\epsilon_\alpha)\pi}{\sin[(-\frac{1}{2} + i\epsilon_\alpha)\pi]} e^{-\pi\epsilon_\alpha} (1 + 2i\epsilon_\alpha)$$

In deriving this, we have used the inherent relations described in Eq. (37) and (39) and the identity Eq. (25) with $q = (-1/2 + i\epsilon_\alpha)$.

3 The Influence of the Permittivity on the Crack Tip ERR

In order to portray the influence of the permittivity of the medium inside the interface crack gap on the crack tip ERR, numerical results for an interface crack of length $2a$ ($a = 100 \mu\text{m}$) in a typical κ class bimaterial PZT-4/BaTiO₃ and those in a typical

Table 2 The crack tip ERR corresponding to two different roots under purely mechanical loading for case i

σ_{22}^{∞} (MPa)	D_2^0 (root 1)	G/a (root 1)	D_2^0 (root 2)	G/a (root 2)
1.0	4.258×10^{-2}	-2.347×10^5	-2.078×10^{-4}	3.062×10^1
3.0	4.478×10^{-2}	-2.642×10^5	-5.929×10^{-4}	2.753×10^2
5.0	4.694×10^{-2}	-2.948×10^5	-9.425×10^{-4}	7.637×10^2
7.0	4.908×10^{-2}	-3.265×10^5	-1.262×10^{-3}	1.494×10^3
9.0	5.119×10^{-2}	-3.592×10^5	-1.556×10^{-3}	2.463×10^3

ϵ -class bimaterial PZT-5H/BaTiO₃ are presented, respectively, with PZT-4 or PZT-5H located on the upper half-space and BaTiO₃ located on the lower half-space. Here, the poling direction of each material is always parallel to the x_2 axis, as shown schematically in Fig. 1. The material properties of PZT-4, PZT-5H, and BaTiO₃ can be seen in Table 1. Four different crack cases are considered: (i) the crack gap is filled with air or vacuum; (ii) the crack gap is filled with silicon oil to avoid discharge [29,30]; (iii) the crack gap is assumed to be fully permeable or filled with NaCl solution [31]; and (iv) the electrically impermeable crack [10]. For case (i), the permittivity ϵ_v is equal to that of air or vacuum $\epsilon_0 = 8.85 \times 10^{-12} \text{ C}^2/\text{N m}^2$, whereas ϵ_v should be $2.5\epsilon_0$ when silicon oil is filled in the crack gap. It is well known that, physically, an impermeable crack model can be reached when the permittivity ϵ_v approaches zero [17]. We use $\epsilon_v = 10^{-5}\epsilon_0$ to approximately describe this charge-free condition, whereas a permeable crack model can be reached when the permittivity ϵ_a approaches infinity [17]. We use $\epsilon_a = 10^8\epsilon_0$ to approximately describe the permeable crack condition.

Firstly, numerical results of the electric displacement solution D_2^0 along the interface crack obtained from Eq. (53) for case (i) with air or vacuum inside the crack gap are discussed. It is noted that Eq. (53) is a quadratic of D_2^0 and thus two distinct roots may exist. Numerical studies show that the discriminant of Eq. (53) is positive for all cases considered. Hence, it yields two real roots for D_2^0 , in which only one real root is physically admissible for a given bimaterial subjected to given loadings. In order to judge which root should be accepted as the admissible root, the crack tip ERR formulated by Eq. (57) is calculated under purely mechanical loading $\sigma_{22}^{\infty} = 1 \text{ MPa}, 3 \text{ MPa}, 5 \text{ MPa}, 7 \text{ MPa}, 9 \text{ MPa}$, respectively, and presented in Table 2. It is found that the crack tip ERR, respectively, induced from two roots under the far-field tension loading yield different results: the negative G corresponding to root 1 and the positive G corresponding to root 2. Obviously, only root 2 will be admissible since it can yield the positive crack tip ERR. Similar discussions hold true for the other three crack cases. In particular, it can be shown from Eq. (53) that if the medium inside the crack is ideally insulated, i.e., $\epsilon_v \rightarrow 0$, then $D_2^0 = 0$, then the commonly used impermeable assumption is reached, whereas if the medium is ideally conducting, i.e., $\epsilon_v \rightarrow \infty$, then we can obtain $\phi^+ = \phi^-$, which is the permeable crack solution. Hence, the traditional impermeable and permeable crack boundary conditions are actually the two limits of the semi-permeable boundary condition.

Calculated numerical results of the crack tip ERR for an interface crack of length $2a$ ($a = 100 \mu\text{m}$) in a typical κ -class bimaterial PZT-4/BaTiO₃ and in a typical ϵ -class bimaterial PZT-5H/BaTiO₃ are plotted in Figs. 2(a)–2(d) and Figs. 3(a)–3(d), respectively. The electric field loading varies from -0.5 MV/m to 0.5 MV/m and four different values of mechanical loading: $0 \text{ MPa}, 2 \text{ MPa}, 5 \text{ MPa}$, and 10 MPa are considered. For each magnitude of the mechanical load, all of four crack cases are considered. It can be seen from Fig. 2(a), that when the mechanical loading vanishes, the crack tip ERR of the semi-permeable crack approaches the values induced from the permeable crack. It is approximately equal to zero with the electric field

varying from -0.5 MV/m to 0.5 MV/m . This indicates that the electric field has a negligible effect on the crack tip ERR in the absence of mechanical loading. A reasonable physical interpretation is that when the mechanical loading is zero, there is no crack opening and then the continuity of the electric potential on the upper and lower crack faces will be satisfied automatically. In this way, the semi-permeable crack model will degenerate into the permeable crack. It is consistent with the assertion, as pointed out by McMeeking [22] for homogeneous piezoelectric materials, that the crack is “invisible” to an applied electric field and there is no crack tip ERR in the presence of an applied electric field without a mechanical loading, whereas it can be found from Figs. 2(a) and 2(b) that for an impermeable crack, the ERR curve is always negative and far apart from those of other three curves. Indeed, in the present study for an interface crack, the impermeable crack model is physically unrealistic. All previous solutions for an interface crack under such an assumption might be incorrect or misleading [22,23], whereas under a purely electric field, the permeable crack model will yield reasonable results since Fig. 2(a) shows a clear feature that the numerical results calculated from the semi-permeable crack (whether the medium inside the crack gap is air or silicon oil) coincide well with those from the permeable crack model. However, this feature is no longer valid in combined mechanical electric loadings. Figure 2(b) with 2 MPa mechanical loading shows that when the κ -class bimaterial PZT-4/BaTiO₃ is under combined mechanical–electric loading, discrepancies between the curves calculated from the semi-permeable crack and those calculated from the permeable crack model are remarkable, especially when the applied electric field becomes larger. For example, when the applied electric field in PZT-4/BaTiO₃ under 2 MPa is -0.5 MV/m and 0.5 MV/m , respectively, the crack tip ERR for air or vacuum is about 0.0109 N m and 0.0099 N m , respectively, far apart from 0.0123 N m for the permeable crack (over 10%); Figure 2(c) with 5 MPa mechanical loading shows that the corresponding values for air or vacuum are 0.0688 N m and 0.0652 N m , respectively, far apart from 0.0766 N m for the permeable crack (over 10%); Figure 2(d) with 10 MPa mechanical loading shows that the corresponding values of the ERR are 0.2790 N m and 0.2751 N m , respectively, far apart from 0.3063 N m for a permeable crack (over 10%). Moreover, it is seen that the positive electric field yields larger discrepancies than those induced from the negative electric field. This is due to the nonlinear feature of the semi-permeable crack model as pointed out by McMeeking [22–24] and reviewed by Chen and Hasebe [27] in homogenous piezoelectrics. It is interesting to see that when the crack gap is filled with silicon oil, the discrepancies become much smaller than those for air or vacuum. These discrepancies are always within the range of 3% when the electric field varies from -0.5 MV/m to 0.5 MV/m . For example, when the mechanical loading is 10 MPa and the electric field is -0.5 MV/m and 0.5 MV/m , respectively, Fig. 2(d) shows that the ERR for the semi-permeable with silicon oil is 0.3002 N m and 0.2999 N m , respectively, very close to 0.3063 N m for the permeable crack. It is concluded that, by comparing it to the semi-permeable crack filled with silicon oil, the impermeable crack model greatly overestimates the effect of the applied electric field whether a mechanical loading is imposed or not, whereas the permeable crack model underestimates the effect but in a tolerant way even though a larger mechanical loading is imposed.

Similar conclusions can be seen in the typical ϵ -class bimaterial PZT-5H/BaTiO₃ as shown in Figs. 3(a)–3(d), although the crack tip singularity in PZT-5H/BaTiO₃ is quite different from that in PZT-4/BaTiO₃. Indeed, the previous results for an interface crack in dissimilar piezoelectric materials should be suspected if they were obtained from the impermeable crack model, whereas those obtained from the permeable crack model may be accepted in a tolerant way if the medium inside the crack gap is silicon oil rather than air or vacuum.

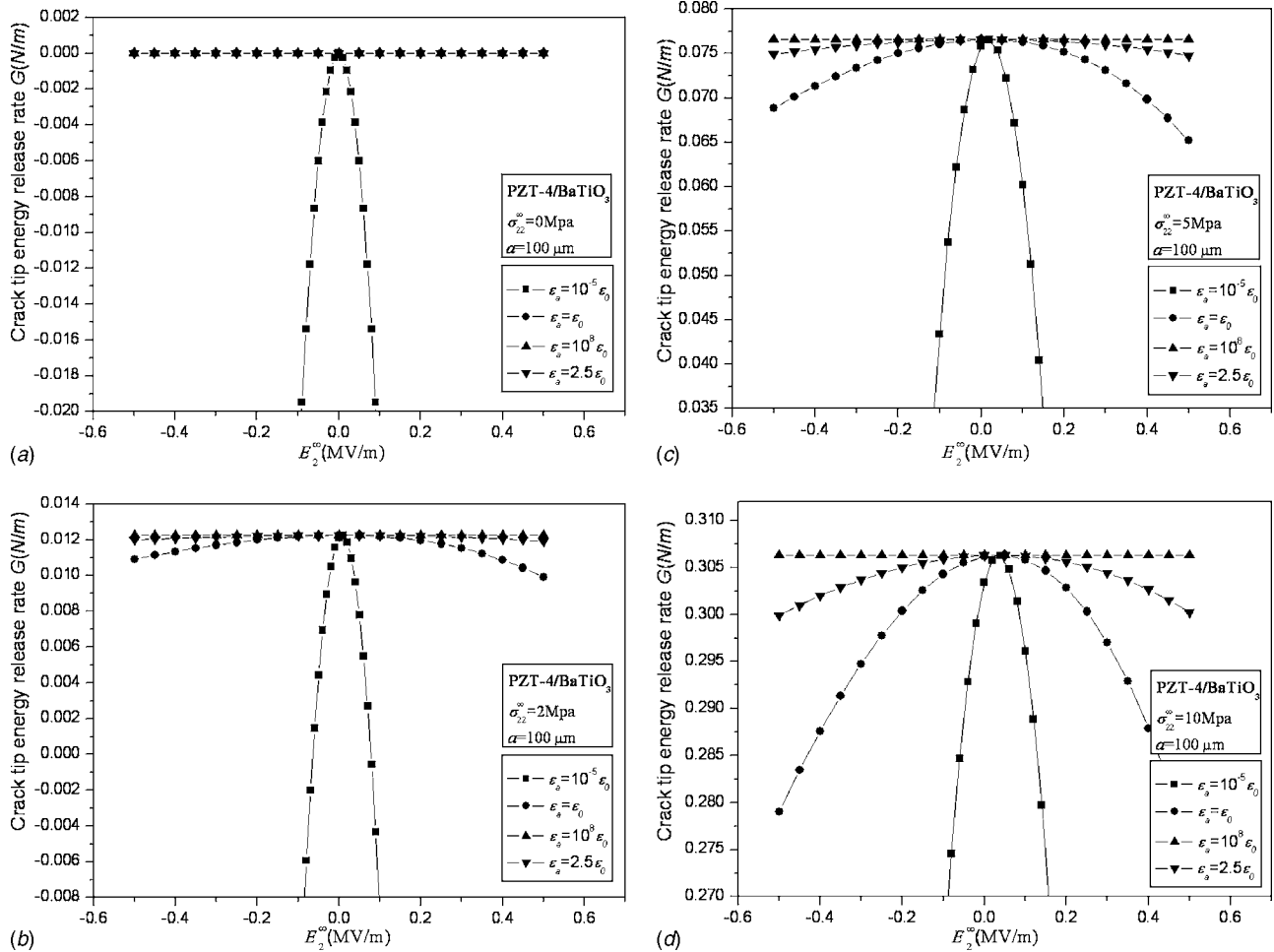


Fig. 2 The crack tip energy release rate for an interface crack in PZT-4/BaTiO₃ bimaterial subjected to the remote tensile stress and electric field. Plots (a), (b), (c), and (d) are given for four different levels of applied tensile stress: 0 MPa, 2 MPa, 5 MPa, and 10 MPa, respectively. Here, curves with the symbol ■ refer to the impermeable crack, curves with the symbol ● refer to the semi-permeable crack filled with air or vacuum, curves with the symbol ▼ refer to the semi-permeable crack filled with silicon oil, and curves with the symbol ▲ refer to the permeable crack.

4 Conclusions

The semi-permeable crack model accounting for the permittivity of the medium inside the crack gap is used for solving the interface crack problem between dissimilar piezoelectric materials. This model shows more physically reasonable features than either the impermeable interface crack model or the permeable interface crack model. Numerical results of the singular parameters ε and κ for the 21 kinds of dissimilar piezoelectric materials show that the singularity for a certain combination of dissimilar piezoelectric materials can either be oscillatory or nonoscillatory when the poling axes of both piezoelectric materials are perpendicular to the interface crack. That is, the crack tip singularity may either be governed by $-1/2 \pm i\varepsilon$ or by $-1/2 \pm \kappa$, depending on the combination of the two dissimilar piezoelectric materials, although the permittivity has no influence on the numerical values of ε or κ obtained previously by Ou and Wu [34] under the impermeable crack model. Among the 21 kinds of dissimilar piezoelectric materials, the following 14 kinds: PZT-4/PZT-5H, PZT-4/PZT-6B, PZT-4/PZT-7A, PZT-4/P-7, PZT-4/BaTiO₃, PZT-5H/PZT-7, PZT-5H/PZT-PIC 151, PZT-6B/BaTiO₃, PZT-7A/P-7, PZT-7A/BaTiO₃, PZT-7A/PZT-PIC 151, P-7/BaTiO₃, P-7/PZT-PIC 151, and BaTiO₃/PZT-PIC 151 belong to the nonoscillatory singularity, i.e., the κ -class bimaterial, whereas the following seven kinds: PZT-4/PZT-PIC 151, PZT-5H/PZT-6B,

PZT-5H/BaTiO₃, PZT-6B/PZT-7A, PZT-6B/P-7, and PZT-6B/PZT-PIC 151 belong to the oscillatory singularity, i.e., the ε -class bimaterial.

There are two new relations Eqs. (37) and (39) among the eigenvector matrices, which are not covered by those given previously by Ting [2] and Suo et al. [6]. These relations provide a useful check to examine numerical results. That is, when some people study interface crack problems in dissimilar piezoelectric materials, all of their numerical results for each kind of dissimilar piezoelectric material should satisfy either these two relations or the orthogonal relations of Ting [2], otherwise they might obtain some incorrect results.

Detailed comparisons and discussions for the crack tip ERR in a typical κ -class bimaterial PZT-4/BaTiO₃ and in a typical ε -class bimaterial PZT-5H/BaTiO₃ are performed, respectively. It is concluded that the discrepancies between the semi-permeable crack and the impermeable crack are remarkable, which shows the unphysical features of the charge-free conditions, whether the mechanical loading is preferred or not [23,24], whereas much smaller discrepancies for the ERR can be seen between the semi-permeable crack with silicon oil and the fully permeable crack even under larger mechanical and electric loadings (e.g., 10 MPa and 0.5 MV/m), which are within a tolerant range of 3% in the present study. However, this conclusion may not be correct if the

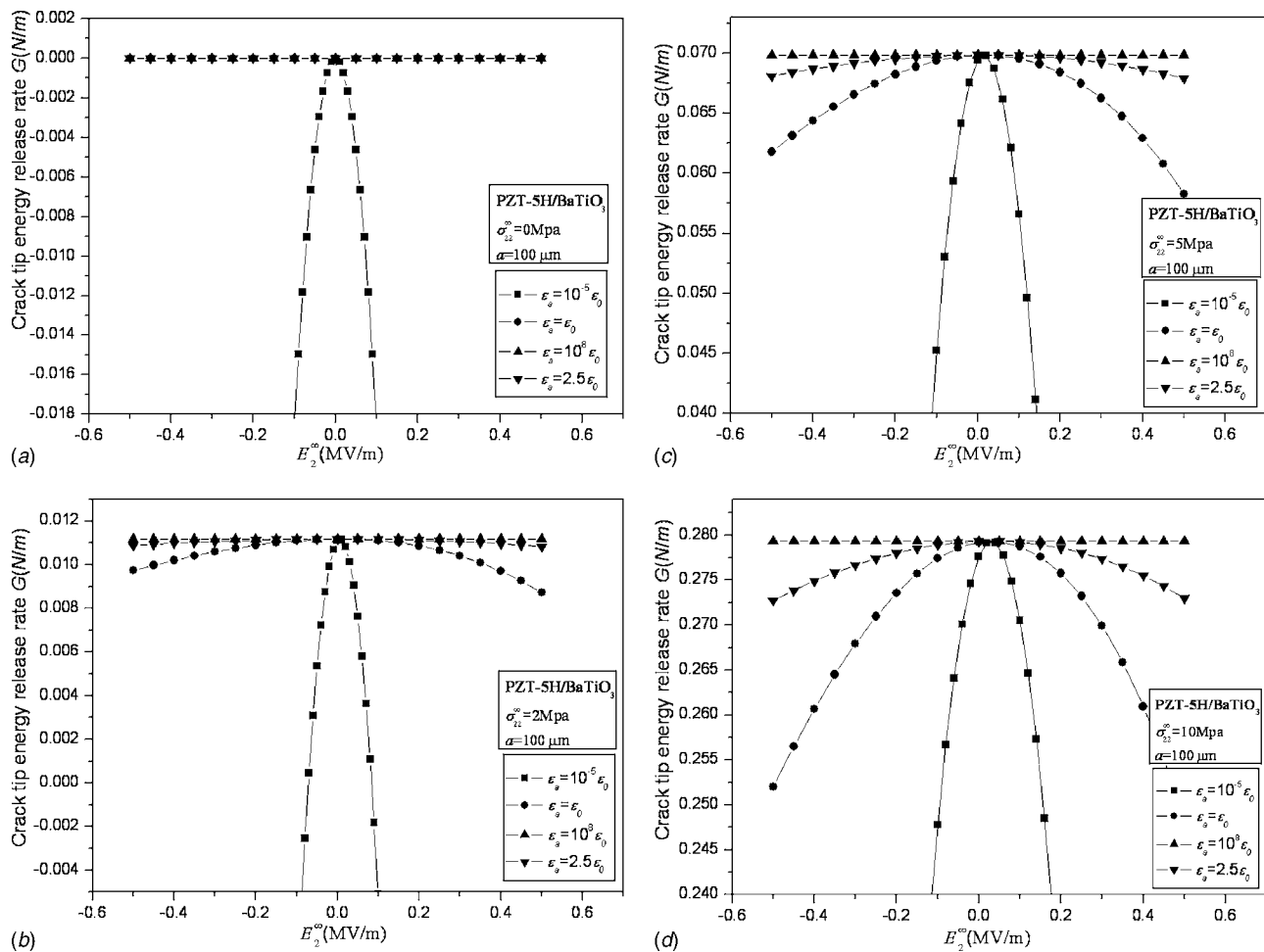


Fig. 3 The crack tip energy release rate for an interface crack in PZT-5H/BaTiO₃ bimaterial subjected to the remote tensile stress and electric field. Plots (a), (b), (c), and (d) are given for four different levels of applied tensile stress 0 MPa, 2 MPa, 5 MPa, and 10 MPa, respectively. Here, curves with the symbol ■ refer to the impermeable crack, curves with the symbol ● refer to the semi-permeable crack filled with air or vacuum, curves with the symbol ▼ refer to the semi-permeable crack filled with silicon oil, and curves with the symbol ▲ refer to the permeable crack.

semi-permeable crack is filled with air or vacuum since the permittivity of air or vacuum is 2.5 times smaller than silicon oil and then the discrepancies will be sometimes over 10% as shown in Figs. 2(b)–2(d) and 3(b)–3(d).

It is concluded that the previous investigations for interface cracks in dissimilar piezoelectric materials under an impermeable crack model might be incorrect or misleading since the model greatly overestimates the effect of the electric field on the ERR, that is, it yields significant discrepancies of the ERR from those obtained from the semi-permeable crack model, whether the mechanical loading is preferred or not, whereas the previous investigations under a permeable crack model may be accepted in a tolerant way since the model yields very small discrepancies from those obtained from the semi-permeable crack model even though the mechanical–electric loading is large. It is emphasized that air/vacuum and silicon oil play quite different roles in a semi-permeable crack. The former always yields smaller values of the ERR and larger discrepancies than those induced from the latter.

It should be emphasized that the fracture criterion for an interface crack in dissimilar piezoelectric materials is still unclear at present and remains to be further investigated. This is either because of the lack of experimental observations or because of the confusion of fracture criteria in homogenous piezoelectric materials (see, Zhang et al. [37]; Chen and Lu [38]; Zhang and Gao [39]; and Chen and Hasebe [27]). To the present authors' knowledge, there is no experimental work for an interface crack in dissimilar

piezoelectric materials in the literature. It is suggested that the four-point bending specimen may be helpful in experiments as Heyer et al. [31] did in homogeneous piezoelectric fracture but the upper material should be with a center-cut gap. It has been proven that Eq. (1) is not only valid in linear analyses in homogeneous piezoelectric crack problems but also in nonlinear analyses with near-tip domain switching [40]. The analysis presented in this paper is a linear piezoelectric analysis, in which the near-tip non-linearity has been entirely neglected. Obviously, the fracture criterion in dissimilar piezoelectric materials should be concerned with some material microstructure such as the near-tip domain switching or microcracks and the mesoscopic failure mechanism. Subsequent investigations into these topics are absolutely needed. Only after doing so, could a more precise fracture criterion be obtained.

Acknowledgment

This work was supported by the National Science Foundation of China (Grant No. 10572110) and the Doctor Foundation of Chinese Education Ministry.

Appendix A: Fundamental Equations of Piezoelectricity

The complete set of basic equations for a linear piezoelectric solid in the absence of body forces and free charges can be written as [6]

constitutive equations

$$\begin{aligned}\sigma_{ij} &= C_{ijkl}u_{k,l} + e_{kij}\phi_{,k} \\ D_i &= e_{ikl}u_{k,l} - \chi_{ik}\phi_{,k}\end{aligned}\quad (A1a)$$

equilibrium equations

$$\begin{aligned}\sigma_{ij,j} &= 0 \\ D_{i,i} &= 0\end{aligned}\quad (A1b)$$

boundary conditions

$$\begin{aligned}\sigma_{ij}n_j &= T_i \\ D_in_i &= -q_s\end{aligned}\quad (A1c)$$

where σ_{ij} , u_i , D_i , and ϕ are separately the stress components, mechanical displacement components, electric displacement components, and electric potential; and C_{ijkl} , e_{kij} , and χ_{ik} are the elastic, piezoelectric, and dielectric constants of the material, respectively.

By using Stroh's theory [1], the general solutions for piezoelectrics under generalized plane strain can be expressed as [6]

$$\mathbf{u} = [u_1, u_2, u_3, \phi]^T = 2 \operatorname{Re}[\mathbf{A}\mathbf{f}(z)] \quad (A2)$$

$$\mathbf{\Sigma} = [\sigma_{i2}] = [\sigma_{21}, \sigma_{22}, \sigma_{23}, D_2]^T = 2 \operatorname{Re}[\mathbf{B}\mathbf{f}'(z)] \quad (A3)$$

where $\mathbf{f}(z)$ is a column of four complex potential functions

$$\mathbf{f}(z) = [f_1(z_1), f_2(z_2), f_3(z_3), f_4(z_4)]^T z_\alpha = x_1 + p_\alpha x_2, \quad (\alpha = 1, 2, 3, 4) \quad (A4)$$

in which the superscript T denotes transpose of a vector; Re denotes the real part of a complex number; the over prime ($'$) denotes the differentiation with respect to the corresponding argument; and p_α are the eigenvalues with positive imaginary parts of the following material's characteristic equation obtained by substituting Eq. (A2) into Eqs. (A1a) and (A1b)

$$[C_{i1k1} + p(C_{i1k2} + C_{i2k1}) + p^2 C_{i2k2}]a_k + [e_{i1i} + p(e_{i12} + e_{2i1}) + p^2 e_{2i2}]a_i = 0$$

$$[e_{i1k} + p(e_{i1k2} + e_{2k1}) + p^2 e_{2k2}]a_k - [\chi_{i11} + p(\chi_{i12} + \chi_{2i1}) + p^2 \chi_{2i2}]a_i = 0 \quad (A5)$$

Equations (A5) can be written in the following compact form

$$[\mathbf{Q} + p(\mathbf{R} + \mathbf{R}^T) + p^2 \mathbf{T}]\mathbf{a} = \mathbf{0} \quad (A6)$$

where $\mathbf{Q}$, $\mathbf{R}$, and $\mathbf{T}$ are 4×4 material matrices defined as follows

$$\mathbf{Q} = \begin{bmatrix} C_{i1k1} & e_{i1i} \\ e_{i1k} & -\chi_{i11} \end{bmatrix}, \quad \mathbf{R} = \begin{bmatrix} C_{i1k2} & e_{2i1} \\ e_{i1k2} & -\chi_{i12} \end{bmatrix}, \quad \mathbf{T} = \begin{bmatrix} C_{i2k2} & e_{2i2} \\ e_{2k2} & -\chi_{2i2} \end{bmatrix} \quad (A7)$$

An auxiliary vector $\mathbf{b}_\alpha$ is introduced by Suo et al. [6], which is defined as

$$\mathbf{b}_\alpha = (\mathbf{R}^T + p_\alpha \mathbf{T})\mathbf{a}_\alpha = -(1/p_\alpha)(\mathbf{Q} + p_\alpha \mathbf{R})\mathbf{a}_\alpha \quad (A8)$$

Thus, we can obtain two 4×4 nonsingular material characteristic matrices $\mathbf{A}$ and $\mathbf{B}$ as follows

$$\mathbf{A} = (\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4), \quad \mathbf{B} = (\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \mathbf{b}_4) \quad (A9)$$

For a bimaterial, we introduce the following material matrices, which will be used subsequently in this paper, defined as

$$\mathbf{H} = \mathbf{Y}_1 + \bar{\mathbf{Y}}_2, \quad \mathbf{Y}_1 = i\mathbf{A}_1\mathbf{B}_1^{-1}, \quad \mathbf{Y}_2 = i\mathbf{A}_2\mathbf{B}_2^{-1} \quad (A10)$$

$$\mathbf{D} = \operatorname{Re}(\mathbf{H}), \quad \mathbf{W} = \operatorname{Im}(\mathbf{H}) \quad (A11)$$

where the subscript 1 and 2 separately indicate that the quantities are defined in materials 1 and 2, respectively; the overbar ($\bar{}$) indicates complex conjugation; and Im denotes the imaginary part

of a complex number. It can be proved that $\mathbf{H}$ is a Hermitian matrix and satisfies

$$\mathbf{H}^T = \bar{\mathbf{H}} \quad (A12)$$

When the piezoelectric bimaterial continuum has certain symmetry or degenerates to be homogeneous, the matrix $\mathbf{H}$ can be real. However, in general, except for the above special case, $\mathbf{H}$ is complex for almost all the piezoelectric bimaterials. It should be emphasized that in Sec. 2, two different cases for piezoelectric bimaterial with real and complex $\mathbf{H}$ have been discussed, respectively.

Appendix B: Singularity Parameters ε and κ and Eigenvectors $\mathbf{w}_\alpha$ ($\alpha=1,2,3,4$) for a Semi-Permeable Interface Crack in Each of Kinds of Piezoelectric Bimaterials

(1) PZT-4/PZT-5H

$$\kappa_1 = 0.0442, \quad \kappa_2 = -0.0442, \quad \varepsilon_1 = \varepsilon_2 = 0$$

$$\mathbf{w} = \begin{bmatrix} 0.8535 & 0.8535 & 0 & 0 \\ -0.3459 & 0.3459 & 0.9995 & 0 \\ 0 & 0 & 0 & 1 \\ 0.3896 & -0.3896 & 0.0312 & 0 \end{bmatrix} \quad (B1)$$

(2) PZT-4/PZT-6B

$$\kappa_1 = 0.0168, \quad \kappa_2 = -0.0168, \quad \varepsilon_1 = \varepsilon_2 = 0$$

$$\mathbf{w} = \begin{bmatrix} 0.7066 & 0.7066 & 0 & 0 \\ 0.5993 & -0.5993 & 0.9455 & 0 \\ 0 & 0 & 0 & 1 \\ 0.3763 & -0.3763 & 0.3256 & 0 \end{bmatrix} \quad (B2)$$

(3) PZT-4/PZT-7A

$$\kappa_1 = 0.0247, \quad \kappa_2 = -0.0247, \quad \varepsilon_1 = \varepsilon_2 = 0$$

$$\mathbf{w} = \begin{bmatrix} 0.8940 & 0.8940 & 0 & 0 \\ 0.1296 & -0.1296 & 0.9646 & 0 \\ 0 & 0 & 0 & 1 \\ 0.4288 & -0.4288 & 0.2636 & 0 \end{bmatrix} \quad (B3)$$

(4) PZT-4/P-7

$$\kappa_1 = 0.0367, \quad \kappa_2 = -0.0367, \quad \varepsilon_1 = \varepsilon_2 = 0$$

$$\mathbf{w} = \begin{bmatrix} 0.8748 & 0.8748 & 0 & 0 \\ -0.1721 & 0.1721 & 0.9906 & 0 \\ 0 & 0 & 0 & 1 \\ 0.4528 & -0.4528 & 0.1364 & 0 \end{bmatrix} \quad (B4)$$

(5) PZT-4/BaTiO₃

$$\kappa_1 = 0.0508, \quad \kappa_2 = -0.0508, \quad \varepsilon_1 = \varepsilon_2 = 0$$

$$\mathbf{w} = \begin{bmatrix} 0.9138 & 0.9138 & 0 & 0 \\ -0.0531 & 0.0531 & 0.9837 & 0 \\ 0 & 0 & 0 & 1 \\ 0.4027 & -0.4027 & 0.1798 & 0 \end{bmatrix} \quad (B5)$$

(6) PZT-5H/PZT-7

$$\kappa_1 = 0.0035, \quad \kappa_2 = -0.0035, \quad \varepsilon_1 = \varepsilon_2 = 0$$

$$\mathbf{w} = \begin{bmatrix} -0.3031 & 0.3031 & 0 & 0 \\ 0.8604 & 0.8604 & 0.9280 & 0 \\ 0 & 0 & 0 & 1 \\ -0.4096 & -0.4096 & -0.3725 & 0 \end{bmatrix} \quad (\text{B6})$$

(7) PZT-5H/PZT-PIC 151

$$\kappa_1 = 0.0447, \quad \kappa_2 = -0.0447, \quad \varepsilon_1 = \varepsilon_2 = 0$$

$$\mathbf{w} = \begin{bmatrix} 0.8601 & 0.8601 & 0 & 0 \\ 0.1680 & -0.1680 & 0.9873 & 0 \\ 0 & 0 & 0 & 1 \\ -0.4817 & 0.4817 & 0.1586 & 0 \end{bmatrix} \quad (\text{B7})$$

(8) PZT-6B/BaTiO₃

$$\kappa_1 = 0.0095, \quad \kappa_2 = -0.0095, \quad \varepsilon_1 = \varepsilon_2 = 0$$

$$\mathbf{w} = \begin{bmatrix} 0.7279 & 0.7279 & 0 & 0 \\ 0.6385 & -0.6385 & 0.9928 & 0 \\ 0 & 0 & 0 & 1 \\ -0.2500 & 0.2500 & -0.1196 & 0 \end{bmatrix} \quad (\text{B8})$$

(9) PZT-7A/P-7

$$\kappa_1 = 0.0023, \quad \kappa_2 = -0.0023, \quad \varepsilon_1 = \varepsilon_2 = 0$$

$$\mathbf{w} = \begin{bmatrix} -0.3020 & 0.3020 & 0 & 0 \\ 0.8810 & 0.8810 & 0.9426 & 0 \\ 0 & 0 & 0 & 1 \\ -0.3632 & -0.3632 & -0.3340 & 0 \end{bmatrix} \quad (\text{B9})$$

(10) PZT-7A/BaTiO₃

$$\kappa_1 = 0.0206, \quad \kappa_2 = -0.0206, \quad \varepsilon_1 = \varepsilon_2 = 0$$

$$\mathbf{w} = \begin{bmatrix} 0.9058 & 0.9058 & 0 & 0 \\ 0.1726 & -0.1726 & 0.9978 & 0 \\ 0 & 0 & 0 & 1 \\ -0.3869 & 0.3869 & 0.0665 & 0 \end{bmatrix} \quad (\text{B10})$$

(11) PZT-7A/PZT-PIC 151

$$\kappa_1 = 0.0175, \quad \kappa_2 = -0.0175, \quad \varepsilon_1 = \varepsilon_2 = 0$$

$$\mathbf{w} = \begin{bmatrix} 0.7006 & 0.7006 & 0 & 0 \\ -0.4683 & 0.4683 & 0.8812 & 0 \\ 0 & 0 & 0 & 1 \\ -0.5383 & 0.5383 & 0.4727 & 0 \end{bmatrix} \quad (\text{B11})$$

(12) P-7/BaTiO₃

$$\kappa_1 = 0.0162, \quad \kappa_2 = -0.0162, \quad \varepsilon_1 = \varepsilon_2 = 0$$

$$\mathbf{w} = \begin{bmatrix} 0.8504 & 0.8504 & 0 & 0 \\ -0.1070 & 0.1070 & 0.9587 & 0 \\ 0 & 0 & 0 & 1 \\ -0.5152 & 0.5152 & 0.2844 & 0 \end{bmatrix} \quad (\text{B12})$$

(13) P-7/PZT-PIC 151

$$\kappa_1 = 0.0351, \quad \kappa_2 = -0.0351, \quad \varepsilon_1 = \varepsilon_2 = 0$$

$$\mathbf{w} = \begin{bmatrix} 0.8354 & 0.8354 & 0 & 0 \\ -0.0033 & 0.0033 & 0.9530 & 0 \\ 0 & 0 & 0 & 1 \\ -0.5496 & 0.5496 & 0.3029 & 0 \end{bmatrix} \quad (\text{B13})$$

(14) BaTiO₃/PZT-PIC 151

$$\kappa_1 = 0.0472, \quad \kappa_2 = -0.0472, \quad \varepsilon_1 = \varepsilon_2 = 0$$

$$\mathbf{w} = \begin{bmatrix} 0.8685 & 0.8685 & 0 & 0 \\ -0.1058 & 0.1058 & 0.9548 & 0 \\ 0 & 0 & 0 & 1 \\ -0.4843 & 0.4843 & 0.2972 & 0 \end{bmatrix} \quad (\text{B14})$$

(15) PZT-4/PZT-PIC 151

$$\varepsilon_1 = 0.0095, \quad \varepsilon_2 = -0.0095, \quad \kappa_1 = \kappa_2 = 0$$

$$\mathbf{w} = \begin{bmatrix} 0.7224 & 0.7224 & 0 & 0 \\ 0.6908i & -0.6908i & -0.6459 & 0 \\ 0 & 0 & 0 & 1 \\ 0.0307i & -0.0307i & 0.7635 & 0 \end{bmatrix} \quad (\text{B15})$$

(16) PZT-5H/PZT-6B

$$\varepsilon_1 = 0.0219, \quad \varepsilon_2 = -0.0219, \quad \kappa_1 = \kappa_2 = 0$$

$$\mathbf{w} = \begin{bmatrix} 0.6832i & -0.6832i & 0 & 0 \\ 0.7300 & 0.7300 & 0.7743 & 0 \\ 0 & 0 & 0 & 1 \\ -0.0203 & -0.0203 & -0.6328 & 0 \end{bmatrix} \quad (\text{B16})$$

(17) PZT-5H/PZT-7A

$$\varepsilon_1 = 0.0069, \quad \varepsilon_2 = -0.0069, \quad \kappa_1 = \kappa_2 = 0$$

$$\mathbf{w} = \begin{bmatrix} 0.3569i & -0.3569i & 0 & 0 \\ 0.8956 & 0.8956 & 0.9350 & 0 \\ 0 & 0 & 0 & 1 \\ -0.2654 & -0.2654 & -0.3548 & 0 \end{bmatrix} \quad (\text{B17})$$

(18) PZT-5H/BaTiO₃

$$\varepsilon_1 = 0.0130, \quad \varepsilon_2 = -0.0130, \quad \kappa_1 = \kappa_2 = 0$$

$$\mathbf{w} = \begin{bmatrix} -0.6480i & 0.6480i & 0 & 0 \\ 0.6944 & 0.6944 & 0.5683 & 0 \\ 0 & 0 & 0 & 1 \\ 0.3128 & 0.3128 & 0.8228 & 0 \end{bmatrix} \quad (\text{B18})$$

(19) PZT-6B/PZT-7A

$$\varepsilon_1 = 0.0055, \quad \varepsilon_2 = -0.0055, \quad \kappa_1 = \kappa_2 = 0$$

$$\mathbf{w} = \begin{bmatrix} -0.6561i & 0.6561i & 0 & 0 \\ 0.7316 & 0.7316 & 0.7861 & 0 \\ 0 & 0 & 0 & 1 \\ 0.1853 & 0.1853 & 0.6181 & 0 \end{bmatrix} \quad (\text{B19})$$

(20) PZT-6B/P-7

$$\varepsilon_1 = 0.0121, \quad \varepsilon_2 = -0.0121, \quad \kappa_1 = \kappa_2 = 0$$

$$\mathbf{w} = \begin{bmatrix} 0.7116 & 0.7116 & 0 & 0 \\ 0.7021i & -0.7021i & -0.4550 & 0 \\ 0 & 0 & 0 & 1 \\ 0.0236i & -0.0236i & 0.8905 & 0 \end{bmatrix} \quad (\text{B20})$$

(21) PZT-6B/PZT-PIC 151

$$\varepsilon_1 = 0.0134, \quad \varepsilon_2 = -0.0134, \quad \kappa_1 = \kappa_2 = 0$$

$$\mathbf{w} = \begin{bmatrix} -0.4841i & 0.4841i & 0 & 0 \\ 0.8048 & 0.8048 & 0.8664 & 0 \\ 0 & 0 & 0 & 1 \\ 0.3435 & 0.3435 & 0.4994 & 0 \end{bmatrix} \quad (\text{B21})$$

References

- [1] Stroh, A. N., 1958, "Dislocations and Cracks in Anisotropic Elasticity," *Philos. Mag.*, **3**, pp. 625–646.
- [2] Ting, T. C. T., 1986, "Explicit Solution and Invariance of the Singularities at an Interface Crack in Anisotropic Composites," *Int. J. Solids Struct.*, **22**, pp. 965–983.
- [3] Herbert, J., 1982, *Ferroelectrics Transducers and Sensors*, Gordon and Breach, New York.
- [4] Harrison, W. B., McHenry, K. D., and Koepke, B. G., 1986, "Monolithic Multilayer Piezoelectric Ceramic Transducer," *Proceedings IEEE 6th International Symposium Applied Ferroelectrics*, Minneapolis, MN, pp. 265–272.
- [5] Kuo, C. M., and Barnett, D. M., 1991, "Stress Singularities of Interface Cracks in Bonded Piezoelectric Half-Spaces," *Modern Theory of Anisotropic Elasticity and Applications*, J. J. Wu, T. C. Ting, and D. M. Barnett, eds., *SIAM Proceeding Series*, Philadelphia, PA.
- [6] Suo, Z., Kuo, C. M., Barnett, D. M., and Willis, J. R., 1992, "Fracture Mechanics of Piezoelectric Ceramics," *J. Mech. Phys. Solids*, **40**, pp. 739–765.
- [7] Zhong, Z., and Meguid, S. A., 1997, "Interfacial Debonding of a Circular In-Homogeneity in Piezoelectric Materials," *Int. J. Solids Struct.*, **34**, pp. 1965–1984.
- [8] Beom, H. G., and Atluri, S. N., 1996, "Near-Tip Fields and Intensity Factors for Interfacial Cracks in Dissimilar Anisotropic Piezoelectric Media," *Int. J. Fract.*, **75**, pp. 163–183.
- [9] Qin, Q. H., and Yu, S. W., 1997, "An Arbitrarily-Oriented Plane Crack Terminating at Interface Between Dissimilar Piezoelectric Materials," *Int. J. Solids Struct.*, **34**, pp. 581–590.
- [10] Ru, C. Q., Mao, X., and Epstein, M., 1998, "Electric-Field Induced Interfacial Cracking in Multiplayer Electrostrictive Actuators," *J. Mech. Phys. Solids*, **46**, pp. 1301–1318.
- [11] Deng, W., and Meguid, S. A., 1998, "Analysis of Conducting Rigid Inclusion at the Interface of Two Dissimilar Piezoelectric Materials," *ASME J. Appl. Mech.*, **65**, pp. 76–84.
- [12] Ding, H. J., Chi, Y., and Guo, F., 1999, "Solutions for Transversely Isotropic Piezoelectric Infinite Body, Semi-Infinite Body and Bimaterial Infinite Body Subjected to Uniform Ring Loading and Charges," *Int. J. Solids Struct.*, **36**, pp. 2613–2631.
- [13] Tian, W. Y., and Chen, Y. H., 2000, "Interaction Between an Interface Crack and Sub Interface Microcracks in Metal/Piezoelectric Bimaterials," *Int. J. Solids Struct.*, **37**, pp. 7743–7757.
- [14] Ma, L. F., and Chen, Y. H., 2001, "Weight Function for Interface Cracks in Dissimilar Anisotropic Piezoelectric Materials," *Int. J. Fract.*, **110**, pp. 263–279.
- [15] Narita, K., and Shindo, Y., 1999, "Anti-Plane Shear Crack in a Piezoelectric Layer Bonded to Dissimilar Half Spaces," *J. Mech. Phys. Solids*, **42**, pp. 66–72.
- [16] Herrmann, K. P., Loboda, V. V., and Govorukha, V. B., 2001, "On Contact Zone Models for an Electrically Impermeable Interface Crack in a Piezoelectric Bimaterial," *Int. J. Fract.*, **111**, pp. 203–227.
- [17] Parton, V. Z., and Kudryavtsev, B. A., 1988, *Electromagnetoelasticity*, Gordon and Breach, New York.
- [18] Hao, T. H., and Shen, Z. Y., 1994, "A New Electric Boundary Condition of Electric Fracture Mechanics and Its Application," *Eng. Fract. Mech.*, **47**, pp. 793–802.
- [19] Dunn, M., 1994, "The Effects of Crack Face Boundary Conditions on the Fracture Mechanics of Piezoelectric Solids," *Eng. Fract. Mech.*, **48**, pp. 25–39.
- [20] Sosa, H., and Khutoryansky, N., 1996, "New Development Concerning Piezoelectric Materials With Defect," *Int. J. Solids Struct.*, **33**, pp. 3399–3414.
- [21] Xu, X. L., and Rajapakse, R. K. N. D., 2001, "On a Plane Crack in Piezoelectric Solids," *Int. J. Solids Struct.*, **38**, pp. 7643–7658.
- [22] McMeeking, R. M., 1999, "Crack Tip Energy Release Rate for a Piezoelectric Compact Tension Specimen," *Eng. Fract. Mech.*, **64**, pp. 217–244.
- [23] McMeeking, R. M., 2001, "Towards a Fracture Mechanics for Brittle Piezoelectric and Dielectric Materials," *Int. J. Fract.*, **108**, pp. 25–41.
- [24] McMeeking, R. M., 2004, "The Energy Release Rate for a Griffith Crack in a Piezoelectric Material," *Eng. Fract. Mech.*, **71**, pp. 1149–1163.
- [25] Liu, Y., and Chen, Y. H., 2005, "An Analytical Solution for a Cracked Piezoelectric Plate According to the PKHS Electric Boundary Condition," *Acta Mech.*, **180**, pp. 233–244.
- [26] Ou, Z. C., and Chen, Y. H., 2005, "On Approach of Crack Tip Energy Release Rate for a Semi-Permeable Crack When Electromechanical Loadings Become Very Large," *Int. J. Fract.*, **133**, pp. 89–105.
- [27] Chen, Y. H., and Hasebe, N., 2005, "Current Understanding on Fracture Behaviors of Ferroelectric/Piezoelectric Materials," *J. Intell. Mater. Syst. Struct.*, **16**, pp. 553–688.
- [28] Ou, Z. C., and Chen, Y. H., 2003, "Discussion of the Crack Face Electric Boundary Condition in Piezoelectric Fracture Mechanics," *Int. J. Fract.*, **123**, pp. L151–L155.
- [29] Park, S. B., and Sun, C. T., 1995, "Fracture Criteria for Piezoelectric Ceramics," *J. Am. Ceram. Soc.*, **78**, pp. 1475–1480.
- [30] Park, S. B., and Sun, C. T., 1995, "Effect of Electric Fields on Fracture of Piezoelectric Ceramics," *Int. J. Fract.*, **70**, pp. 203–216.
- [31] Heyer, V., Schneider, G. A., Balke, H., Drescher, J., and Bahr, H. A., 1998, "A Fracture Criterion for Conducting Cracks in Homogeneously Poled Piezoelectric PZT-PIC 151 Ceramics," *Acta Mater.*, **46**, pp. 6615–6622.
- [32] Muskhelishvili, N. I., 1953, *Some Basic Problems of the Mathematical Theory of Elasticity*, Noordhoff, Leyden, The Netherlands.
- [33] England, A. H., 1965, "A Crack Between Dissimilar Media," *ASME J. Appl. Mech.*, **32**, pp. 400–402.
- [34] Ou, Z. C., and Wu, X. J., 2003, "On the Crack Tip Stress Singularity of Interfacial Cracks in Transversely Isotropic Piezoelectric Bimaterials," *Int. J. Solids Struct.*, **40**, pp. 7499–7511.
- [35] Wu, K. C., 1990, "Stress Intensity Factors and Energy Release Rate for Interfacial Cracks Between Dissimilar Anisotropic Materials," *ASME J. Appl. Mech.*, **57**, pp. 882–886.
- [36] Qu, J. M., and Li, Q. Q., 1991, "Interfacial Dislocation and Its Application to Interface Cracks in Anisotropic Bimaterials," *J. Elast.*, **26**, pp. 169–195.
- [37] Zhang, T. Y., Zhao, M. H., and Tong, P., 2002, "Fracture of Piezoelectric Ceramics," *Adv. Appl. Mech.*, **38**, pp. 147–289.
- [38] Chen, Y. H., and Lu, T. J., 2003, "Cracks and Fracture in Piezoelectric Materials," *Adv. Appl. Mech.*, **39**, pp. 127–221.
- [39] Zhang, T. Y., and Gao, C. F., 2004, "Fracture Behaviors of Piezoelectric Materials," *Theor. Appl. Fract. Mech.*, **41**, pp. 339–379.
- [40] Fulton, C. C., and Gao, H., 2001, "Effect of Local Polarization Switching on Piezoelectric Fracture," *J. Mech. Phys. Solids*, **49**, pp. 927–952.

**Seyed Ali Modarres
Najafabadi**
e-mail: smodar@cim.mcgill.ca

József Kövecses
e-mail: jozsef.kovecses@mcgill.ca

Jorge Angeles
e-mail: angeles@cim.mcgill.ca

Department of Mechanical Engineering,
McGill University,
817 Sherbrooke Street West,
Montreal, Quebec H3A 2K6, Canada

Energy Analysis and Decoupling in Three-Dimensional Impacts of Multibody Systems

This paper discusses an exact decomposition of the kinetic energy to determine the energy content that influences the dynamics of unilateral contacts in multibody systems. This decomposition essentially divides the kinetic energy of the whole multibody system into two completely decoupled parts associated with the constrained and admissible directions of unilateral contacts. This will provide a picture of how the energy absorption/dissipation during impacts is related to the variation of the generalized velocities and the configuration of multibody systems. Potential applications of such a decoupling are highlighted. [DOI: 10.1115/1.2712226]

1 Introduction

Impact is a complex multiscale phenomenon that takes place when at least two separate solid bodies enter in collision. The forces arising during impact are normally large enough to cause not only elastic, but also plastic deformation in the local contact area. This is known to be one of the primary causes of energy loss during impact [1]. The presence of fast-scale physical effects, such as wave propagation, is another important source of energy dissipation. During impact, part of the kinetic energy absorbed in the compression phase is recovered during the restitution phase and the rest remains in the colliding bodies in various forms (e.g., plastic deformations, impact-induced vibrations). This results in energy dissipation, which can be related to the dynamics in the normal or tangential directions of contact. On the other hand, friction also acts as another energy dissipating mechanism; it is associated with the tangential dynamics of the contacting bodies.

In the literature, the problem of impact between two single bodies has been widely discussed. However, the interaction of multibody systems and the related energy considerations have been given less attention. If two single bodies are involved, then energy dissipation due to the normal contact forces results in a direct decrease of the normal relative velocity of the colliding bodies. The energy loss due to friction influences the tangential relative velocity. However, for collisions of multibody systems, the variation of the normal or tangential relative velocities of the contacting bodies does not stand as a direct measure of absorption or restitution of kinetic energy for the whole system. In fact, it is useful to know how energy dissipation/absorption during impact is related to changes in the generalized velocities and the configuration of the system.

In this paper, we discuss an effective and physically meaningful decomposition of the kinetic energy for unilaterally constrained multibody systems. This decomposition is useful in the energy analysis of general three-dimensional contact scenarios in multibody systems. The decoupling of the kinetic energy is achieved through a decomposition of the generalized velocity vector, based on the normal and tangential directions of unilateral contacts. This will also make it possible to determine what parts of the total pre-impact kinetic energy contribute to the dynamics and energy dissipation along the normal and tangential directions. These directions can also be represented in an abstract sense in the tangent space of the configuration manifold of the multibody system.

2 Decoupling of the Generalized Velocity Vector and the Kinetic Energy

Let us consider that n generalized coordinates describe the configuration of the system. These are represented in an n -dimensional vector $\mathbf{q}$. In this paper, we study holonomic, scleronomic systems subject to a set of unilateral constraints. Therefore, the kinetic energy of the system can be expressed as a quadratic function of the generalized velocities. We address the case of simultaneous multiple-point contact scenarios with m contact points. For the instant of impact, the constraints can be written at the velocity level as [2,3]

$$\mathbf{A}(\mathbf{q})\dot{\mathbf{q}} \geq 0 \quad (1)$$

where $\mathbf{A}$ is the $m \times n$ constraint Jacobian matrix and $\dot{\mathbf{q}}$ is the $n \times 1$ vector of generalized velocities. The number of the constraints in Eq. (1) gives the number of contact points. Each constraint parameterizes the normal direction for one contact point, in terms of the generalized coordinates and velocities of the multibody system at hand. For a single-point contact, there is only one constraint, thus $\mathbf{A}$ reduces to an $n \times 1$ row vector. Figure 1 shows a typical multibody impact situation with a single point contact with one normal and one tangential direction.

Let us consider now the tangent space of the configuration manifold of the multibody system for the pre-impact instant. This tangent space is an n -dimensional Euclidean vector space [4]. The constraints in Eq. (1) give m vectors (directions) in this vector space. Those m vectors represent the relative normal velocity for each of the m contact pairs in the tangent space, and define a subspace we shall call the *space of constrained motion*. The dimension of this subspace is equal to the number of independent constraints in Eq. (1) (it is m if all the constraints are linearly independent). The orthogonal complement of the space of constrained motion is another subspace of the tangent space. It shall be termed the *space of admissible motion*. The union of these two subspaces gives back the tangent space of the configuration manifold. The space of admissible motion includes vectors with directions along which the motion is not constrained. These directions can be seen as the mapping of the tangential directions of the contact pairs into the tangent space of the configuration manifold.

One of the essential points of our analysis is the decomposition of the generalized velocity vector of the multibody system ($\dot{\mathbf{q}}$) into two components, while keeping the original parametrization of the system ($\mathbf{q}, \dot{\mathbf{q}}$). One of these components belongs to the space of constrained motion, and the other is an element of the space of admissible motion. It is important to perform this decomposition in a physically meaningful way, by accounting for the metric tensor of the tangent space and the potential inhomogeneity in physi-

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received November 7, 2005; final manuscript received December 20, 2006. Review conducted by Marc P. Mignolet.

cal units. The detailed mathematical procedure and description of such a decomposition was introduced in [5].

Based on these considerations, the following projection operators can be established to decompose a vector [5]

$$\mathbf{H}_c = \mathbf{U}^{-1}(\mathbf{AU}^{-1})^\dagger \mathbf{A} \quad (2)$$

$$\mathbf{H}_a = \mathbf{I} - \mathbf{U}^{-1}(\mathbf{AU}^{-1})^\dagger \mathbf{A} \quad (3)$$

where $\mathbf{U}$ is obtained from the Cholesky decomposition of the generalized mass matrix of the system as $\mathbf{M} = \mathbf{U}^T \mathbf{U}$, and $\mathbf{I}$ is $n \times n$ identity matrix. Exponent “ $\dagger$ ” denotes the right Moore-Penrose generalized inverse, which can be obtained from the singular value decomposition for example [6].¹ Operators $\mathbf{H}_c$ and $\mathbf{H}_a$ are associated with the spaces of constrained and admissible motions, respectively. These expressions were derived for the general case where the constraints do not have to be linearly independent, i.e., the constraint Jacobian need not have full row rank [5]. Below, we simplify the foregoing expressions as we primarily address the case of independent constraints in this work. Using the above operators, the vector of generalized velocities of the system can be decoupled as

$$\dot{\mathbf{q}} = \dot{\mathbf{q}}_c + \dot{\mathbf{q}}_a = \mathbf{H}_c \dot{\mathbf{q}} + \mathbf{H}_a \dot{\mathbf{q}} \quad (4)$$

where vector $\dot{\mathbf{q}}_a = \mathbf{H}_a \dot{\mathbf{q}}$ represents the component in the space of admissible motion, and $\dot{\mathbf{q}}_c = \mathbf{H}_c \dot{\mathbf{q}}$ is the component in the space of constrained motion. It is important to note that we are not using local parameterizations for the subspaces; we keep the original global parameterizations of the tangent space. Note that $\dot{\mathbf{q}}_c$ and $\dot{\mathbf{q}}_a$ in the above equations only represent the decomposed parts of $\dot{\mathbf{q}}$ associated with the spaces of constrained and admissible motions, respectively, and need not be, in fact, total time-derivatives. Vector $\dot{\mathbf{q}}_a$ is completely admissible with the constraints in Eq. (1), i.e., it is unconstrained. On the other hand, $\dot{\mathbf{q}}_c$ is associated with the constrained directions of the unilateral constraints. Using Eq. (4), the total kinetic energy of the system can be written as

$$T = \frac{1}{2} \dot{\mathbf{q}}^T \mathbf{M} \dot{\mathbf{q}} = \frac{1}{2} (\mathbf{H}_c \dot{\mathbf{q}})^T \mathbf{M} (\mathbf{H}_c \dot{\mathbf{q}}) + \frac{1}{2} (\mathbf{H}_a \dot{\mathbf{q}})^T \mathbf{M} (\mathbf{H}_a \dot{\mathbf{q}}) + \frac{1}{2} (\mathbf{H}_a \dot{\mathbf{q}})^T \mathbf{M} (\mathbf{H}_c \dot{\mathbf{q}}) + \frac{1}{2} (\mathbf{H}_c \dot{\mathbf{q}})^T \mathbf{M} (\mathbf{H}_a \dot{\mathbf{q}}) \quad (5)$$

Next, we will show that the said operators completely decouple the whole kinetic energy into parts associated with the constrained and admissible motions of the system. To demonstrate this, we need to prove that the projection operators are orthogonal with respect to the total mass matrix of the system, i.e., $\mathbf{H}_c^T \mathbf{M} \mathbf{H}_a = \mathbf{O}$ and $\mathbf{H}_a^T \mathbf{M} \mathbf{H}_c = \mathbf{O}$, with $\mathbf{O}$ denoting the $n \times n$ zero matrix. Given the way the right generalized inverse can be expressed for a full row rank matrix, the formulation of the above projection operators can be rewritten for the case of independent constraints in the forms

$$\mathbf{H}_c = \mathbf{U}^{-1}(\mathbf{AU}^{-1})^\dagger \mathbf{A} = \mathbf{U}^{-1} \mathbf{U}^{-T} \mathbf{A}^T (\mathbf{AU}^{-1} \mathbf{U}^{-T} \mathbf{A}^T)^{-1} \mathbf{A} = \mathbf{M}^{-1} \mathbf{A}^T (\mathbf{AM}^{-1} \mathbf{A}^T)^{-1} \mathbf{A} \quad (6)$$

$$\mathbf{H}_a = \mathbf{I} - \mathbf{U}^{-1}(\mathbf{AU}^{-1})^\dagger \mathbf{A} = \mathbf{I} - \mathbf{M}^{-1} \mathbf{A}^T (\mathbf{AM}^{-1} \mathbf{A}^T)^{-1} \mathbf{A} \quad (7)$$

Hence, and noting that $\mathbf{H}_a = \mathbf{I} - \mathbf{H}_c$, we obtain

$$\mathbf{H}_c^T \mathbf{M} \mathbf{H}_a = \mathbf{H}_c^T \mathbf{M} (\mathbf{I} - \mathbf{H}_c) = \mathbf{H}_c^T \mathbf{M} - \mathbf{H}_c^T \mathbf{M} \mathbf{H}_c \quad (8)$$

Considering that $\mathbf{AM}^{-1} \mathbf{A}^T$ is a symmetric matrix for $\mathbf{M}$ is symmetric, we obtain

$$\mathbf{H}_c^T \mathbf{M} = \mathbf{A}^T (\mathbf{AM}^{-1} \mathbf{A}^T)^{-1} \mathbf{AM}^{-1} \mathbf{A}^T \mathbf{M} = \mathbf{A}^T (\mathbf{AM}^{-1} \mathbf{A}^T)^{-1} \mathbf{A} \quad (9)$$

¹For a full row rank matrix $\mathbf{S}$, the right Moore-Penrose generalized inverse can be expressed as $\mathbf{S}^\dagger = \mathbf{S}^T (\mathbf{SS}^T)^{-1}$.

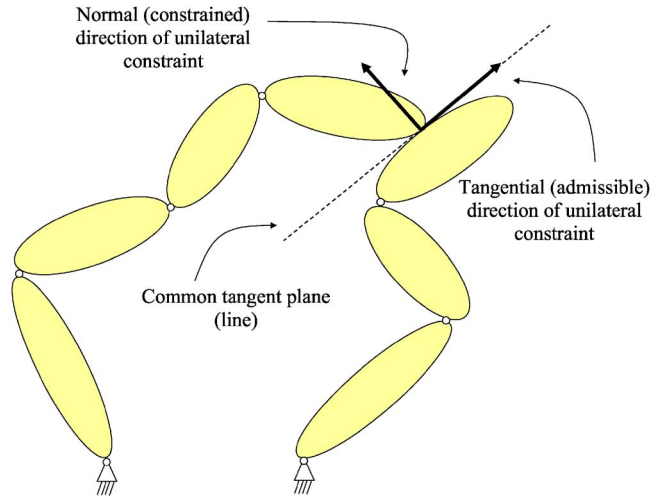


Fig. 1 General unilaterally constrained multibody system

$$\begin{aligned} \mathbf{H}_c^T \mathbf{M} \mathbf{H}_c &= \mathbf{A}^T (\mathbf{AM}^{-1} \mathbf{A}^T)^{-1} (\mathbf{AM}^{-1} \mathbf{A}^T) (\mathbf{AM}^{-1} \mathbf{A}^T)^{-1} \mathbf{A} \\ &= \mathbf{A}^T (\mathbf{AM}^{-1} \mathbf{A}^T)^{-1} \mathbf{A} \end{aligned} \quad (10)$$

Equations (9) and (10) show that $\mathbf{H}_c^T \mathbf{M} = \mathbf{H}_c^T \mathbf{M} \mathbf{H}_c$. Using Eq. (8) and noting that $\mathbf{H}_a^T \mathbf{M} \mathbf{H}_c = (\mathbf{H}_c^T \mathbf{M} \mathbf{H}_a)^T$, we obtain

$$\mathbf{H}_c^T \mathbf{M} \mathbf{H}_a = \mathbf{O} \quad \mathbf{H}_a^T \mathbf{M} \mathbf{H}_c = \mathbf{O} \quad (11)$$

which shows the orthogonality of the projection operators with respect to the total mass matrix of the system.

The direct consequence of Eq. (11) is that the second and third terms in Eq. (5) vanish. These terms represent the coupling between the kinetic energies associated with the spaces of constrained and admissible motions, since they contained both $\dot{\mathbf{q}}_a$ and $\dot{\mathbf{q}}_c$. Therefore, the said decomposition completely decouples the kinetic energy of the whole system, and hence

$$T = T_c + T_a = \frac{1}{2} (\mathbf{H}_c \dot{\mathbf{q}})^T \mathbf{M} (\mathbf{H}_c \dot{\mathbf{q}}) + \frac{1}{2} (\mathbf{H}_a \dot{\mathbf{q}})^T \mathbf{M} (\mathbf{H}_a \dot{\mathbf{q}}) \quad (12)$$

where T_c is the kinetic energy associated with the constrained motion, i.e., the normal directions of the contact pairs; T_a is in turn the kinetic energy associated with the admissible motion. We emphasize again that this decomposition is physically meaningful, i.e., completely consistent with the potentially different physical units of the generalized coordinates. It has to be noted that we present here the proof of decoupling for the case of independent constraints. However, it can be shown that this decoupling also holds for the general case where the constraints are not necessarily independent. In such a case, the simplifications in Eqs. (6) and (7) cannot be performed; the orthogonality of the projection operators can be shown, for example, by directly using a singular-value-decomposition-based interpretation of the generalized inverses in product $\mathbf{H}_c^T \mathbf{M} \mathbf{H}_a$. The formulation provided above are derived for the general case and are valid for the general motion (translation and rotation) of the elements of a multibody system.

We illustrate the above-mentioned complete decoupling, with the aid of Eqs. (4), (5), (11), and (12): For a fixed configuration of the multibody system, any variation of the generalized velocity vector that lies in the space of admissible motion affects only part T_a , while leaving part T_c of the kinetic energy unchanged. On the other hand, if any variation of the vector of generalized velocities lies in the space of constrained motion, it causes part T_c to change, but it does not affect T_a at all. These considerations clearly show that the decomposition discussed completely isolates the kinetic energies of admissible motion and constrained motion. This is an important property, which makes decomposition (12) useful in the energy analysis of unilaterally constrained mechanical systems.

3 Analysis of Energy Dissipation During Impact

The above decomposition can provide new perspectives on both the dynamics of contact and investigation of energy absorption and restitution during impact of multibody systems. If we consider the work done by normal contact forces, this affects only the kinetic energy associated with the space of constrained motion T_c . This can be readily perceived by considering the physical meaning of the decomposition at hand, and that the normal contact forces represent vectors in the dual of the space of constrained motion [7,8]. Therefore, the absorption and restitution of energy due to the effect of energy-dissipating mechanisms acting in the normal directions of contact pairs during impact only changes the value of the kinetic energy of the constrained motion T_c . In other words, the energy dissipation due to the plastic deformation of the contact areas along the normal directions and other effects associated with the normal contact forces causes T_c to decrease, while leaves T_a unchanged. On the other hand, the energy loss due to friction and other energy-dissipating elements acting in the tangential directions of the contact pairs reduces T_a during the impact phase and does not directly affect T_c . This is a direct result of the complete decoupling of the kinetic energy. We have to note that during the impact period, the configuration of the system is assumed to remain unchanged. Therefore, although the generalized inertia matrix is a function of the system configuration specified by the generalized coordinates, it will remain constant during impact. Consequently, the same mutually orthogonal vector spaces can be used to represent the system before and after impact (in the pre- and post-impact instants). Hence, we can directly use this decoupling to study the energy absorption and restitution occurring in the two phases of impact.

Let us consider an impact in a multibody system (Fig. 1), and study the compression and restitution phases of the impact between the two bodies involved. Based on its definition, the compression phase ends when the normal relative velocity of the contact points belonging to the colliding bodies vanishes. This means that, at the end of compression, the projection of the vector of generalized velocities into the space of constrained motion is the zero vector ($\dot{\mathbf{q}}_c = \mathbf{0}$). This consequently implies that the kinetic energy of the constrained motion decreases to zero at the end of compression. Conclusively, the compression phase ends when T_c of the system vanishes. On the other hand, in the restitution phase, the part of the energy that is recovered due to the positive work of the normal contact forces is transformed back into the kinetic energy associated with the normal directions, driving the two bodies apart. This indicates that, during restitution, the elastic strain energy that is recovered due to the effect of normal contact forces, is actually transformed into the kinetic energy associated with the constrained motion (T_c). Thus, the amount of T_c of the interacting multibody systems after impact will be the same as the amount of the energy released during restitution due to the positive work of normal contact forces.

The difference between the energy absorbed during compression and the energy released during restitution, due to the effect of normal contact forces, describes the energy loss that is due to the dynamics characteristics associated with the normal directions of impact [9]. Therefore, we conclude that the difference between the amounts of the kinetic energy of constrained motion before and after impact defines this energy loss. This can result in a new interpretation for the energetic coefficient of restitution. This new interpretation will be introduced particularly for the interaction of complex systems. The energetic coefficient of restitution was originally defined for the impact of two single bodies, considering the normal direction of contact. The definition given in [9] is

$$e_s^2 = -\frac{W_n(p_f) - W_n(p_c)}{W_n(p_c)} \quad (13)$$

where $W_n(p_f)$ is the work done by the normal contact force during the whole period of impact (compression and restitution phases),

$W_n(p_c)$ is the work of the normal contact force in the compression phase, and e_s represents the energetic coefficient of restitution. The new interpretation can be based on the ratio of the post- and pre-impact values of the kinetic energy of the constrained motion (T_c) and can be written as

$$e_s = \sqrt{\frac{T_c^+}{T_c^-}} \quad (14)$$

where T_c^+ and T_c^- are the post- and pre-impact kinetic energies of constrained motion, respectively. Such an interpretation of the energetic coefficient of restitution can be of considerable potential in the dynamics of impact in multibody systems. This coefficient can be used, similar to the other coefficients of restitution, to determine the post-impact state of the system. However, this new definition is originally given for impact scenarios involving multibody systems and can capture the effect of not only local, but also global energy-dissipating features (e.g., losses due to wave propagation). The coefficient of restitution defined in Eq. (14) can address the general case of simultaneous multiple-point impact scenarios, as opposed to the other types of coefficients of restitution. However, for that case, we have to assume that the normal contact impulses developed in all closed contact pairs are compressive. Therefore, they contribute to kinetic energy dissipation during impact. In general, simultaneous multiple-point impact problems, cases might occur where some of the closed contact pairs have zero normal impulses [2,10,11]. These are excluded here. The detailed discussion on such cases, as well as the potential applications and advantages of the proposed definition of the energetic coefficient of restitution are the subjects of our ongoing work and will be addressed in upcoming papers.

Further, the new definition given in Eq. (14) can be used to unify the concepts of the energetic and the Newtonian coefficients of restitution. To elaborate more on this, let us look at the mathematical expression describing our new interpretation for the energetic coefficient of restitution

$$e_s = \sqrt{\frac{T_c^+}{T_c^-}} = \sqrt{\frac{(\frac{1}{2}\dot{\mathbf{q}}_c^T \mathbf{M} \dot{\mathbf{q}}_c)^+}{(\frac{1}{2}\dot{\mathbf{q}}_c^T \mathbf{M} \dot{\mathbf{q}}_c)^-}} = \sqrt{\frac{(\dot{\mathbf{q}}_c^T \mathbf{M} \dot{\mathbf{q}}_c)^+}{(\dot{\mathbf{q}}_c^T \mathbf{M} \dot{\mathbf{q}}_c)^-}} \quad (15)$$

We can consider that the generalized velocities are geometrically interpreted in the tangent space of the configuration manifold, where the metric tensor representation is the mass matrix [4,12]. Hence, the numerator and denominator of the above equation represent the norm of $\dot{\mathbf{q}}_c$ after and before impact, respectively. On the other hand, Newton's coefficient of restitution is defined as the ratio of the normal relative velocities of the contact points after and before impact. Further, note that $\dot{\mathbf{q}}_c$ is the part of the generalized velocity vector which is associated with the normal (constrained) directions of the contact pairs. Thus, one can observe that the Newtonian coefficient of restitution can be seen as a special type of the generalized energetic coefficient of restitution, where the mass matrix (metric tensor representation of the tangent space of the configuration manifold) is in a very simple form and can be eliminated from the expression.

In the tangential directions of contact, the energy dissipated due to friction, and other effects acting in tangential directions, causes the kinetic energy of admissible motion (T_a) to decrease. The difference between the values of T_a before and after impact measures the energy loss due to the energy-dissipating processes associated with the tangential directions of the contact pairs. The decoupling of the kinetic energy provides us with a tool to analyze the energy losses associated with both the normal and tangential (constrained and admissible) directions of the contact pairs.

The formulation discussed here can also have many other potential applications. One can be the optimum design of systems aimed to have high energy-dissipating capabilities (e.g., damping the relative motion) in the normal or tangential directions of impact. Examples of such systems can be found in vehicle suspen-

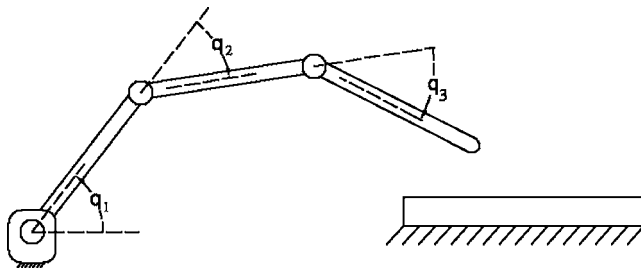


Fig. 2 Unilaterally constrained three-link planar manipulator

sion systems or aircraft landing gears. In such cases, the kinetic energy of constrained motion T_c (or the one associated with admissible motion T_a), expressed in terms of the generalized coordinates and generalized velocities of the system, should be minimized over a desired period of time, which can be considered as a design objective. This can open up possibilities to develop novel design concepts in such applications.

The concept of decoupling of the kinetic energy can also be used to improve the operational conditions in contact onset. For example, we consider the case in which the smooth contact task of a robotic manipulator with a stiff environment is of interest. This means that the normal contact velocity between the end-effector and the environment should decrease to zero in the shortest possible time after the first contact takes place. This objective will be satisfied if the kinetic energy of the constrained motion (T_c) attains its minimum value at the pre-impact instant. Such a decoupling can also be performed for other instants when the unilateral constraints are not active. This can be advantageous in feedback control, when the objective is to minimize the effects of impact on the system. This goal can be met if the decomposed generalized velocities are controlled properly. The decoupling and analysis presented in this work can lead to methods to develop optimum trajectories resulting in a minimum jump in the constrained motion after impact.

4 Numerical Illustration

To illustrate the foregoing ideas, we analyze the single-point contact of the three-link robotic arm shown in Fig. 2. In the initial configuration of the arm, its end-point is a finite distance above a fixed flat plate. The arm is then released from this configuration, and undergoes a series of collisions with the plate. A model-based controller is used to ensure that the end-point of the arm has a desired velocity in the pre-impact instant just before the first collision. In the pre-impact instant, the controller is removed to allow us to focus on the effects of the impacts. Our aim is to give a qualitative type of illustration. Therefore, the magnitude of the various quantities is not really important here. We note that it is assumed that only the end-point of the third link will undergo collisions.

We consider impacts where friction is neglected, in order to highlight the decoupling discussed above. In this case, all three types of the coefficients of restitution (kinematic, kinetic, and energetic) lead to the same results. It can also be shown that for frictionless contact scenarios, our new definition for the energetic coefficient of restitution leads to the same results as the other types of coefficients of restitution, which will be proved in detail in subsequent papers. Here, we developed the simulation based on the energetic coefficient of restitution [9] with a value of 0.85. The parameters used for this example are given in more detail in Appendix A.

Figure 3 shows the displacement and the relative normal velocity histories of the end-point with respect to the plate. Figure 4 shows the decomposition of the generalized velocity vector based on Eq. (4). The results show that the component in the space of constrained motion ($\dot{q}_c$) is heavily influenced by the repeated impacts, while the component in the space of admissible motion ($\dot{q}_a$)

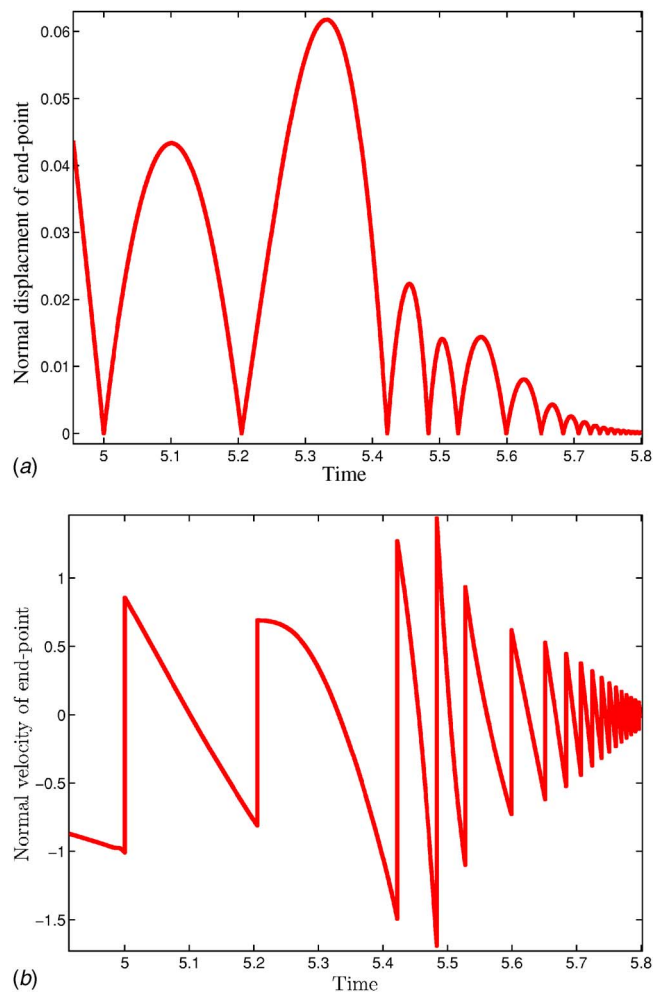


Fig. 3 Displacement and velocity of the end-point in the normal direction of contact

is not affected. The decomposition of the kinetic energy based on Eq. (12) is shown in Fig. 5. It can again be observed that only the part associated with the space of constrained motion (T_c) is involved directly in the impacts. The kinetic energy of admissible motion (T_a) does not contain jump discontinuities; it is not affected directly by the contacts. We also evaluated our new interpretation for the energetic coefficient of restitution. Equation (14) was evaluated for each impact. It gave back exactly the value of the coefficient of restitution that was used to develop the simulation based on the technique described in [9] using Eq. (13). This also shows that the new interpretation is valid. We believe that our interpretation is actually more general. It can also address the case of multiple-point contacts in complex systems where we can define a kind of an "effective" coefficient of restitution that is not necessarily associated with a single-point contact.

Based on the foregoing results, several, nontrivial phenomena can also be observed. We do not discuss these here in detail, but highlight a few important points. Although energy is dissipated in each collision, it can also happen that the kinetic energy of the whole system actually increases after impact. The reason here is the way in which gravitational and kinetic energies are transformed into each other, and how they are distributed between the various elements of the system and the two subspaces of the tangent space. In other words, although kinetic energy is dissipated during the impact period (Fig. 5), it can increase over a finite period of time due to the possible transformation of gravitational potential energy into kinetic energy. This is why, in some collisions, the end-effector bounces higher than in the previous one.

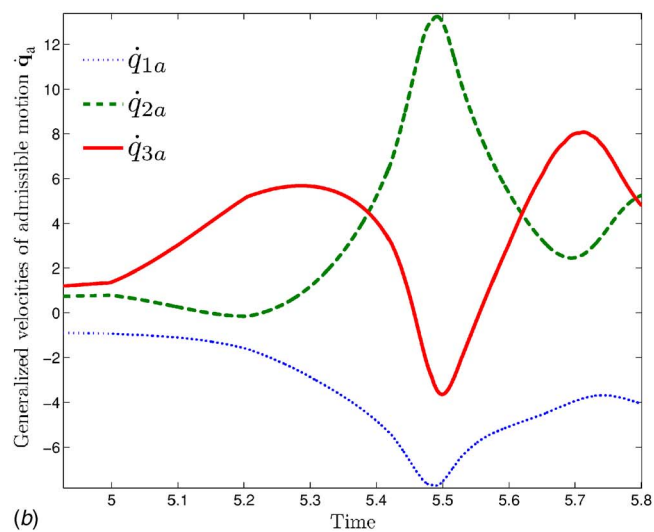
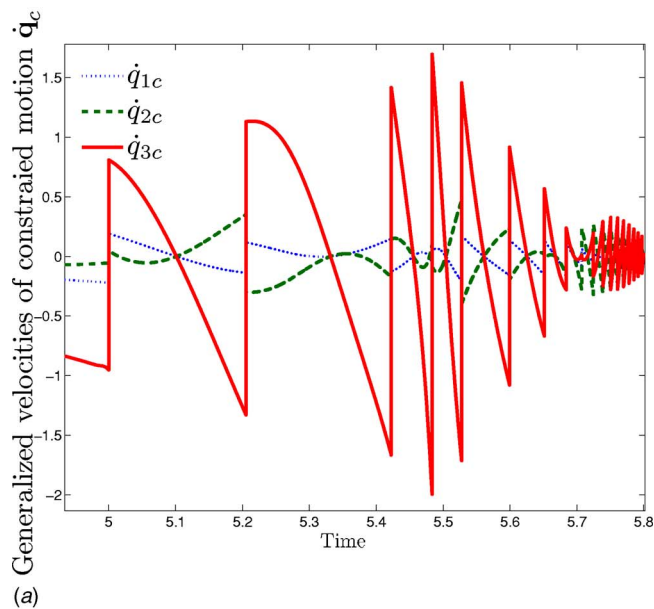


Fig. 4 Generalized velocities associated with admissible and constrained motions

Over a finite period of time, the energy loss can be observed by investigating the total mechanical energy of the system (Fig. 6). This shows again that the energy analysis of impact in complex systems is far from being trivial. The decomposition and decoupling introduced in this work can be a useful tool to further advance our understanding of impacts in unilaterally constrained systems.

In the above study, an example of a frictionless contact problem has been considered. Such an example has been used to put into perspective the useful features of the decomposition approach discussed above. However, the analysis of frictional impact problems involving multibody systems will require a detailed analysis of the tangential dynamics characteristics of the system as well. We believe that the proposed method can also be considerably useful to study dynamic behavior of frictional impact problems, which will be the focus of our future work.

5 Conclusions

In this paper we introduced an analysis intended to characterize impacts in multibody systems. The highlights of this analysis are the decomposition of the unilaterally constrained generalized velocity vector of a multibody system, and the decoupling of the

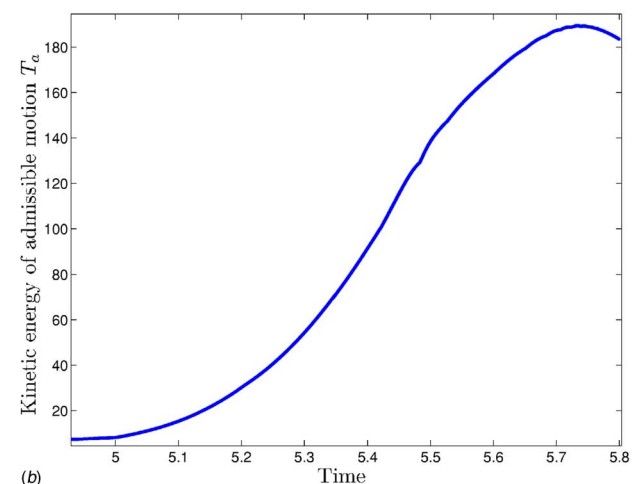
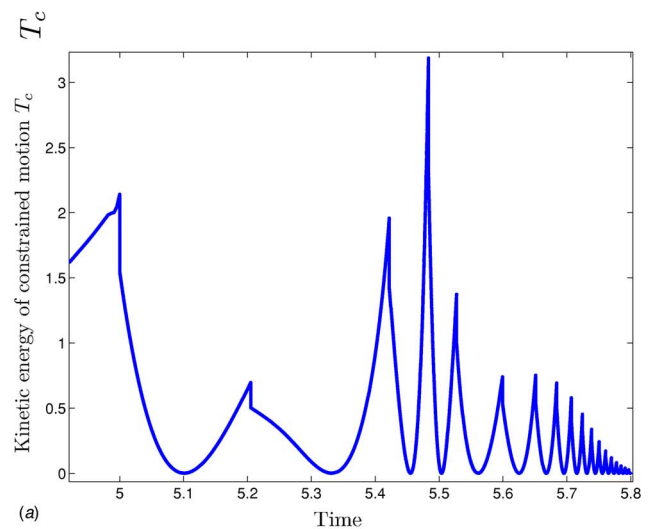


Fig. 5 Kinetic energies of constrained T_c and admissible T_a motions

kinetic energy. The kinetic energy is split into two independent parts: one is associated with the normal directions of contact, the other with the tangential directions of the contact pairs. These normal and tangential directions of the contact pairs are also rep-

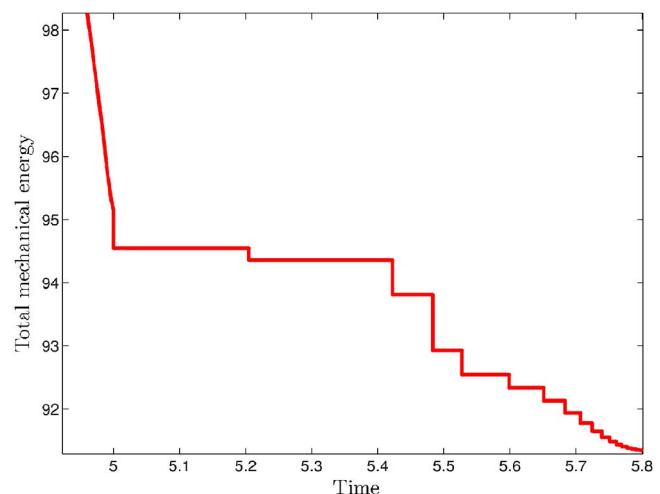


Fig. 6 Total mechanical energy of the manipulator

resented in the tangent space of the configuration manifold, as the space of constrained motion and the space of admissible motion. This representation makes possible the above-mentioned decomposition. The analysis holds for holonomic, scleronomous systems subject to a set of unilateral constraints. We addressed the case primarily when the constraints describing the relative motion of colliding bodies are linearly independent.

This decomposition of the kinetic energy is useful in the analysis of absorption, restitution and dissipation of energy during impacts. The kinetic energy can be decoupled in the pre- and post-impact instants, which makes it possible to gain insight into the nature of the interaction, and determine how the kinetic energy is distributed among the representative directions of the contact pairs. For instance, this can lead to a new interpretation for the energetic coefficient of restitution, defined particularly for collisions involving multibody systems. This new interpretation is also valid for simultaneous multiple-point contact scenarios if all the closed contacts contribute to the energy dissipation. We believe that the material reported here can be applied in the optimum design, dynamic analysis and feedback control of multibody systems undergoing impact.

Acknowledgment

This work was supported by the Natural Sciences and Engineering Research Council of Canada and Le Fonds Québécois de la Recherche sur la Nature et les Technologies, and also by McGill University. The support is gratefully acknowledged.

Appendix A

The example studied in Sec. 4 involves a three-link planar robotic manipulator, where each link is modeled as a slender rigid beam with length and weight of 1 m and 5 kg, respectively. The second moment of inertia about the axis perpendicular to the plane through the center of mass for each link is assumed to be $I_{zz} = 5/12 \text{ kg m}^2$.

A computed torque control algorithm is employed to move the system from its initial configuration to the pre-impact position. The initial and pre-impact configurations of the system are expressed as

$$\begin{aligned} q_1^I &= \pi/4, & q_2^I &= 0, & q_3^I &= 0 \\ q_1^F &= \pi/3, & q_2^F &= -\pi/3, & q_3^F &= -\pi/3 \end{aligned}$$

where the superscripts “I” and “F” denote the initial and pre-impact configurations, respectively (Fig. 2). The controller is designed in a way that in the pre-impact instant, the normal and tangential velocities of the end-point of the third arm would be -1 m/s and 1 m/s , respectively, where the positive normal and tangential directions are assumed upward and to the right, respectively. The third arm is desired to have zero angular velocity in the pre-impact instant. The value of the energetic coefficient of restitution is assumed to be 0.85.

Appendix B

In this appendix, we briefly outline the procedure of derivation of the projection operators given in Eqs. (2) and (3), following the method given in [5]. Let us assume that a general multibody system is subject to a set of bilateral constraints which can be stated at the velocity level in the form

$$\mathbf{A}(\mathbf{q}, t) \dot{\mathbf{q}} + \mathbf{h}(\mathbf{q}, t) = \mathbf{0} \quad (\text{A1})$$

where $\mathbf{A}$ is the constraint Jacobian matrix and $\mathbf{h}$ represents any prescribed rheonomic terms. Therefore, the virtual displacements $\delta \mathbf{q}$ must be compatible (admissible) with such constraints and should satisfy the relation

$$\mathbf{A} \delta \mathbf{q} = \mathbf{0} \quad (\text{A2})$$

Based on the above equation, the vector of virtual displacements ($\delta \mathbf{q}$) can be decomposed into two parts, as follows

$$\delta \mathbf{q} = \delta \mathbf{q}_c + \delta \mathbf{q}_a \quad (\text{A3})$$

where $\delta \mathbf{q}_a$ is the part which is kinematically admissible with the constraints, and $\delta \mathbf{q}_c$ is the part that is specified by the constraints. Thus, using these ideas, it is possible to decompose the tangent space of the configuration manifold into two mutually orthogonal subspaces, namely, the spaces of admissible and constrained motions. Matrix $\mathbf{A}$ can also be interpreted in a more general sense as a representation of a linear transformation operating on vectors in the tangent space of the configuration manifold. Therefore, this decomposition is also valid for any particular configuration where the equality sign is relaxed in Eq. (A1), and we arrive to unilateral constraints. It is of interest to do the decomposition using the original (potentially redundant) set of generalized coordinates, without the introduction of a new independent set. However, due to the potential inhomogeneity of the elements of the virtual displacement vector in physical units, methods of matrix computations cannot directly be used. First, we need to find a transformation taking the virtual displacement vector into a vector homogeneous in physical units. Considering the Cholesky decomposition of the generalized mass matrix of the system, i.e., $\mathbf{M} = \mathbf{U}^T \mathbf{U}$, the following transformation can be established

$$\delta \mathbf{q}_h = \mathbf{U} \delta \mathbf{q} \quad (\text{A4})$$

where the subscript “h” denotes the homogeneity of the corresponding parameter in physical units. Based on the above equation, Eq. (A2) can be rewritten as

$$\mathbf{A} \delta \mathbf{q} = \mathbf{A} \mathbf{U}^{-1} \delta \mathbf{q}_h = \mathbf{A}_h \delta \mathbf{q}_h \quad (\text{A5})$$

where the rows of matrix $\mathbf{A}_h = \mathbf{A} \mathbf{U}^{-1}$ are also homogeneous in physical units. Therefore, the product $\mathbf{A}_h^T \mathbf{A}_h$ makes physical sense and it is possible to perform the singular value analysis of matrix $\mathbf{A}_h$, in a physically meaningful way, to obtain the projection operators associated with the admissible and constrained subspaces of the tangent space, as follows [5,6].

$$\delta \mathbf{q}_h = \delta \mathbf{q}_{ch} + \delta \mathbf{q}_{ah} = \mathbf{A}_h^\dagger \mathbf{A}_h \delta \mathbf{q}_h + (\mathbf{I} - \mathbf{A}_h^\dagger \mathbf{A}_h) \delta \mathbf{q}_h \quad (\text{A6})$$

where $\mathbf{I}$ represents an $n \times n$ identity matrix. This can then be used along with Eq. (A4) to establish the decomposition for the original components as

$$\delta \mathbf{q} = \delta \mathbf{q}_c + \delta \mathbf{q}_a = \mathbf{H}_c \delta \mathbf{q} + \mathbf{H}_a \delta \mathbf{q} \quad (\text{A7})$$

where

$$\begin{aligned} \mathbf{H}_c &= \mathbf{U}^{-1} (\mathbf{A} \mathbf{U}^{-1})^\dagger \mathbf{A} \\ \mathbf{H}_a &= \mathbf{I} - \mathbf{U}^{-1} (\mathbf{A} \mathbf{U}^{-1})^\dagger \mathbf{A} \end{aligned} \quad (\text{A8})$$

where $\mathbf{H}_c$ and $\mathbf{H}_a$ represent the orthogonal projection operators associated with the spaces of constrained and admissible motions, respectively.

References

- [1] Johnson, K., 1985, *Contact Mechanics*, Cambridge University Press, Cambridge, UK.
- [2] Pfeiffer, F., and Glocker, C., 1996, *Multibody Dynamics With Unilateral Contacts*, John Wiley and Sons, New York.
- [3] Brogliato, B., 1999, *Nonsmooth Mechanics: Models, Dynamics and Control*, 2nd ed., Springer, London.
- [4] Blajer, W., 1997, “A Geometric Unification of Constrained System Dynamics,” *Multibody Syst. Dyn.*, 1(7), pp. 3–21.
- [5] Kóvacs, J., and Piedbœuf, J.-C., 2003, “A Novel Approach for the Dynamic Analysis and Simulation of Constrained Mechanical Systems,” *Proc. of the 2003 ASME Design Engineering Technical Conferences*, Chicago, September 2–6, 2003, Vol. 5A, pp. 143–152, Paper No. DETC2003/Vib-48318.
- [6] Golub, G. H., and Van Loan, C. F., 1996, *Matrix Computations*, The Johns Hopkins University Press, Baltimore.
- [7] Halmos, P. R., 1987, *Finite-Dimensional Vector Spaces*, Springer, New York.

- [8] Gel'Fand, I. M., 1989, *Lectures on Linear Algebra*, Dover Publications, New York.
- [9] Stronge, W. J., 2000, *Impact Mechanics*, Cambridge University Press, Cambridge, UK.
- [10] Moreau, J. J., 1988. "Unilateral Contact and Dry Friction in Finite Freedom Dynamics," *Nonsmooth Mechanics and Applications* (Courses and Lectures, Vol. 302), Springer-Verlag, Vienna, pp. 1–82.
- [11] Glocker, C., 2006, "An Introduction to Impacts," *Nonsmooth Mechanics of Solids* (CISM Courses and Lectures, Vol. 485), Springer-Verlag, Vienna, pp. 45–102.
- [12] Synge, J., and Schild, A., 1978, *Tensor Calculus*, Dover Publications, New York.

Nonlinear Behavior and Critical State of a Penny-Shaped Dielectric Crack in a Piezoelectric Solid

Chun-Ron Chiang

Professor

Department of Power Mechanical Engineering,
National Tsing Hua University,
Hsin Chu 30013, Taiwan
e-mail: crchiang@pme.nthu.edu.tw

George J. Weng¹

Professor

Fellow ASME
Department of Mechanical and Aerospace
Engineering,
Rutgers University,
New Brunswick, NJ 08903
e-mail: weng@jove.rutgers.edu

By means of the Hankel transform and dual-integral equations, the nonlinear response of a penny-shaped dielectric crack with a permittivity κ_0 in a transversely isotropic piezoelectric ceramic is solved under the applied tensile stress σ_z^A and electric displacement D_z^A . The solution is given through the universal relation, $D^c/\sigma_z^A = K_D/K_I = M^D/M^\sigma$, regardless of the electric boundary conditions of the crack, where D^c is the effective electric displacement of the crack medium, and K_D and K_I are the electric displacement and the stress intensity factors, respectively. The proportional constant M^D/M^σ has been derived and found to have the characteristics: (i) for an impermeable crack it is equal to D_z^A/σ_z^A ; (ii) for a permeable one it is only a function of the ceramic property; and (iii) for a dielectric crack with a finite κ_0 it depends on the ceramic property, the κ_0 itself, and the applied σ_z^A and D_z^A . The latter dependence makes the response of the dielectric crack nonlinear. This nonlinear response is found to be further controlled by a critical state (σ_c, D_z^A) , through which all the D^c versus σ_z^A curves must pass, regardless of the value of κ_0 . When $\sigma_z^A < \sigma_c$, the response of an impermeable crack serves as an upper bound, whereas that of the permeable one serves as the lower bound, and when $\sigma_z^A > \sigma_c$ the situation is exactly reversed. The response of a dielectric crack with any κ_0 always lies within these bounds. Under a negative D_z^A , our solutions further reveal the existence of a critical κ^* , given by $\kappa^* = -RD_z^A$, and a critical D^* , given by $D^* = -\kappa_0/R$ (R depends only on the ceramic property), such that when $\kappa_0 > \kappa^*$ or when $|D_z^A| < |D^*|$, the effective D^c will still remain positive in spite of the negative D_z^A . [DOI: 10.1115/1.2712227]

1 Introduction

Piezoelectric ceramics are brittle in nature and they often possess sharp cracks, which could inhibit their potential applications. In order to provide better understanding on the fracture behavior of this class of materials, a vast body of literature has been developed over the past 2 decades. A central issue involved in the study is the electrical boundary condition. Earlier studies have often treated it as either a permeable [1] or an impermeable one [2–8]. These approaches have shed some light on the fracture behavior of piezoelectric materials, but as pointed out by McMeeking [9], Pak [3], Suo et al. [7], Zhang and Hack [10], Zhang [11], Dunn [12], Park and Sun [8], Sosa and Khutoryansky [13], and Zhang et al. [14], among others, these two boundary conditions may lead to some unrealistic results under certain conditions. The simplicity offered by these two boundary conditions was also the reason for their adoption in most dynamic analyses [15–18]. Of course this is not to say that analyses based on the permeable or impermeable conditions—especially under Mode II and III loadings—are without merit, as when a crack is idealized as zero thickness and never opens up, one can only set the electric potential as either continuous or discontinuous across the crack surfaces and this would lead to these two types of conditions. For a “closed” crack, the electric permittivity “inside the crack” really plays no role and physically it should never appear in the crack boundary condition. It is only after the crack has opened up that the electric permittivity inside the crack starts to influence the crack-tip behavior.

Under a tensile loading the crack will open up, and the dielectric permittivity of the crack medium, κ_0 , will play a crucial role. In this case it is essential to introduce the dielectric permittivity κ_0 into the boundary condition. With such a perspective, Parton and Kudryavtsev [19] have suggested using the separation of crack surfaces as a measure for the calculation of electric field from the potential drop, while Hao and Shen [20] have suggested the crack opening displacement as the measure for the drop. In addition Shindo et al. [21] and Yang [22] have also suggested direct implementation of the continuity of the normal electric displacement and electric potential from the lower crack surface to the crack medium and then to the upper crack surface. By taking the effect of κ_0 into consideration, various electro-elastic behaviors of a dielectric crack have been reported by Shindo et al. [23], McMeeking [24], Xu and Rajapakse [25], Wang and Jiang [26], Dascalu and Homentcovisch [27], and Wang and Mai [28], Zhang and Gao [29], among others, for a two-dimensional (2D) crack.

This paper is concerned with the response of a 3D penny-shaped dielectric crack. The penny-shaped crack problem can be treated as a limiting case of a spheroidal crack, or directly as a crack with flat surfaces. The spheroidal problem is a piezoelectric analog of Eshelby's [30] elastic problem. This approach has been taken by Wang [31], Kogan et al. [32], Huang [33], and Chiang and Weng [34]. Both Wang's and Huang's works were formulated in Eshelby's framework and performed for an impermeable crack, while that of Kogan et al. was developed using the harmonic potentials. Chiang and Weng made use of the Eshelby-type electromechanical S tensor derived by Dunn and Wienecke [35] in their analysis and gave explicit results for a strongly oblate cavity with a finite dielectric permittivity. While these analyses do offer valuable insights into the general characteristics of a spheroidal void, many important characteristics for a flat crack could not be

¹Corresponding author.

Contributed by the Applied Mechanic Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received May 30, 2006; final manuscript received July 13, 2006. Review conducted by Robert M. McMeeking.

simply extracted from it as a limiting case, for a flat crack with a finite radius, a , cannot be represented by a spheroidal void.

We shall follow the second approach by treating the crack with flat surfaces at the outset. Along this line we have seen several important studies on the penny-shaped crack in recent years. Chen and Shioya [36,37] have used Fabricant's [38] potential theory and the formulation of Ding et al. [39], to find the solution of a pair of concentrated forces and examined some fundamental characteristics of an impermeable crack. Zhao et al. [40] have used the Somigliana identity to solve the boundary integral equations for a penny-shaped crack parallel to the free boundary of a half space. Using the permeable boundary condition Yang and Lee [41,42] have examined the crack characteristics for a piezoelectric strip under both normal and nonaxisymmetric loading. By considering the constitutive equation of the dielectric medium, Lin et al. [43] have included the effect of κ_0 into their formulation. Li and Lee [44] have implemented the Hao-Shen type boundary condition to solve the crack problem in a piezoelectric layer. Their results illustrate the strong dependence of the crack opening displacement intensity factor on κ_0 as a function of the applied electric field. In addition, Chen and Lim [45] have solved the nonaxisymmetric, point force problem that is applied at the crack faces of a permeable crack.

Against the backdrops of these contributions, our focus here is on the development of effective electric displacement D^c of the crack medium due to the presence of κ_0 . More specifically we plan to examine the dependence of D^c on the applied tensile stress σ_z^A , axial electric displacement D_z^A , and dielectric permittivity κ_0 . This issue has not been explored in the past, and it also has direct implications on the determination of the electric displacement intensity factor K_D in terms of the stress intensity factor K_I and the dielectric permittivity κ_0 . Our analysis will reveal that the response of the dielectric crack is nonlinear, but those of the impermeable and permeable ones are linear, and these linear responses will serve as the upper and lower bounds of the nonlinear response. We shall also report that there exist several critical states that define the transition of the nonlinear response of the dielectric crack under both positive and negative D_z^A . The existence of such a critical state is also a new feature that is not found in the literature. Along the way to a full analysis of the dielectric crack, we will also collect some known and some new results for the impermeable and permeable cracks. We shall cast the linear nature of the two idealized cracks and the nonlinear characteristics of the dielectric crack in a universal relation through which their distinct features can all be made transparent.

We note in passing that most piezoelectric ceramics are also ferroelectric and that they could undergo domain switch under certain electromechanical load [46–48]. The constitutive behavior of a ferroelectric ceramic is nonlinear. In this study we will limit our analysis to the linear range of ceramic response.

2 Field Equations of a Transversely Isotropic Solid Under Axisymmetric Load

As pointed out by Ikeda [49], there are four major ways to write the linear constitutive relations of a piezoelectric solid, but for the present problem it is most suitable to choose

$$\begin{aligned}\sigma_{ij} &= C_{ijkl}\varepsilon_{kl} - e_{nij}E_n \\ D_m &= e_{mkl}\varepsilon_{kl} + \kappa_{mn}E_n\end{aligned}\quad (1)$$

where σ , ε , D , and E are tensors of the stress, strain, electric displacement, and electric field, respectively; C and κ are elastic stiffness at constant electric field and dielectric permittivity at constant strain; and e is the piezoelectric moduli tensor. The strain and electric field are derivable from the displacement u_i and electric potential ϕ as

$$\varepsilon_{ij} = \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right); \quad E_i = -\frac{\partial \phi}{\partial x_i} \quad (2)$$

In the absence of body force and space charge, the equation of equilibrium and the charge equation of electrostatics assert

$$\frac{\partial \sigma_{ij}}{\partial x_j} = 0; \quad \frac{\partial D_i}{\partial x_i} = 0 \quad (3)$$

For a transversely isotropic ceramic with the $6mm$ symmetry along the z axis (poling direction), the components of the electro-mechanical moduli in Eq. (1) carry the form

$$\begin{bmatrix} C_{11} & C_{12} & C_{13} & 0 & 0 & 0 & 0 & 0 & -e_{31} \\ C_{12} & C_{11} & C_{13} & 0 & 0 & 0 & 0 & 0 & -e_{31} \\ C_{13} & C_{13} & C_{33} & 0 & 0 & 0 & 0 & 0 & -e_{33} \\ 0 & 0 & 0 & C_{44} & 0 & 0 & 0 & -e_{15} & 0 \\ 0 & 0 & 0 & 0 & C_{44} & 0 & -e_{15} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & C_{66} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & e_{15} & 0 & k_{11} & 0 & 0 \\ 0 & 0 & 0 & e_{15} & 0 & 0 & 0 & k_{11} & 0 \\ e_{31} & e_{31} & e_{33} & 0 & 0 & 0 & 0 & 0 & k_{33} \end{bmatrix} \quad (4)$$

in Nye's [50] contracted notations, where $C_{66} = 1/2(C_{11} - C_{12})$. It follows that, after substituting Eq. (2) into Eq. (1) and rewriting them in the cylindrical coordinates (r, θ, z) , we have the nonvanishing components of stress and electric displacement in terms of u and ϕ , as

$$\begin{aligned}\sigma_r &= C_{11} \frac{\partial u_r}{\partial r} + C_{12} \frac{u_r}{r} + C_{13} \frac{\partial u_z}{\partial z} + e_{31} \frac{\partial \phi}{\partial z} \\ \sigma_\theta &= C_{12} \frac{\partial u_r}{\partial r} + C_{11} \frac{u_r}{r} + C_{13} \frac{\partial u_z}{\partial z} + e_{31} \frac{\partial \phi}{\partial z} \\ \sigma_z &= C_{13} \frac{\partial u_r}{\partial r} + C_{13} \frac{u_r}{r} + C_{33} \frac{\partial u_z}{\partial z} + e_{33} \frac{\partial \phi}{\partial z} \\ \tau_{zr} &= C_{44} \left(\frac{\partial u_r}{\partial z} + \frac{\partial u_z}{\partial r} \right) + e_{15} \frac{\partial \phi}{\partial r} \\ D_r &= e_{15} \left(\frac{\partial u_r}{\partial z} + \frac{\partial u_z}{\partial r} \right) - \kappa_{11} \frac{\partial \phi}{\partial r} \\ D_z &= e_{31} \frac{\partial u_r}{\partial r} + e_{31} \frac{u_r}{r} + e_{33} \frac{\partial u_z}{\partial z} - \kappa_{33} \frac{\partial \phi}{\partial z}\end{aligned}\quad (5)$$

Similarly Eq. (3) can be written as

$$\begin{aligned}\frac{\partial \sigma_r}{\partial r} + \frac{\partial \tau_{zr}}{\partial z} + \frac{\sigma_r - \sigma_\theta}{r} &= 0 \\ \frac{\partial \tau_{zr}}{\partial r} + \frac{\partial \sigma_z}{\partial z} + \frac{\tau_{zr}}{r} &= 0 \\ \frac{\partial D_r}{\partial r} + \frac{\partial D_z}{\partial z} + \frac{D_r}{r} &= 0\end{aligned}\quad (6)$$

This in turn leads to the field equations in terms of displacement and electric potential as

$$\begin{aligned}C_{11} \left[\frac{\partial^2 u_r}{\partial r^2} + \frac{\partial}{\partial r} \left(\frac{u_r}{r} \right) \right] + C_{44} \frac{\partial^2 u_r}{\partial z^2} + (C_{13} + C_{44}) \frac{\partial^2 u_z}{\partial r \partial z} \\ + (e_{31} + e_{15}) \frac{\partial^2 \phi}{\partial r \partial z} = 0\end{aligned}$$

$$\begin{aligned}
& (C_{13} + C_{44}) \left(\frac{\partial^2 u_r}{\partial r \partial z} + \frac{1}{r} \frac{\partial u_r}{\partial z} \right) + C_{33} \frac{\partial^2 u_z}{\partial z^2} + C_{44} \left(\frac{\partial^2 u_z}{\partial r^2} + \frac{1}{r} \frac{\partial u_z}{\partial r} \right) \\
& + e_{15} \left(\frac{\partial^2 \phi}{\partial r^2} + \frac{1}{r} \frac{\partial \phi}{\partial r} \right) + e_{33} \frac{\partial^2 \phi}{\partial z^2} = 0 \\
& (e_{31} + e_{15}) \left(\frac{\partial^2 u_r}{\partial r \partial z} + \frac{1}{r} \frac{\partial u_r}{\partial z} \right) + e_{33} \frac{\partial^2 u_z}{\partial z^2} + e_{15} \left(\frac{\partial^2 u_z}{\partial r^2} + \frac{1}{r} \frac{\partial u_z}{\partial r} \right) \\
& - \kappa_{11} \left(\frac{\partial^2 \phi}{\partial r^2} + \frac{1}{r} \frac{\partial \phi}{\partial r} \right) - \kappa_{33} \frac{\partial^2 \phi}{\partial z^2} = 0 \quad (7)
\end{aligned}$$

We now solve the above partial differential equations with suitable crack conditions prescribed over the crack surfaces by the integral transform method [51].

3 General Solutions and Boundary Conditions

The problem of a penny-shaped crack loaded under an external tensile stress σ_z^A and axial electric displacement D_z^A can be considered as a superposition of a uniform one and another with a suitable boundary condition on the crack surfaces. The first part of the problem is trivial so we will focus on the second part.

Guided by the solution of elasticity, the displacement and electric potential for the second part can be written through the Hankel transform

$$\begin{aligned}
u_r &= \int_0^\infty A_r(s) J_1(sr) \exp(vs z) ds \\
u_z &= \int_0^\infty A_z(s) J_0(sr) \exp(vs z) ds \\
\phi &= \int_0^\infty A_\phi(s) J_0(sr) \exp(vs z) ds \quad (8)
\end{aligned}$$

where J_0 and J_1 are Bessel's functions of order zero and one, respectively. The functions $A_r(s)$, $A_z(s)$, and $A_\phi(s)$ are yet to be determined.

Substituting Eq. (8) into Eq. (7), it turns out that it is sufficient to have these unknown functions in the following form

$$A_r(s) = A_r Q(s), \quad A_z(s) = A_z Q(s), \quad A_\phi(s) = A_\phi Q(s) \quad (9)$$

where $Q(s)$ depends only on the boundary conditions, while the coefficients A_r , A_z , and A_ϕ must satisfy

$$(-C_{11} + C_{44} \nu^2) A_r - (C_{13} + C_{44}) \nu A_z - (e_{31} + e_{15}) \nu A_\phi = 0$$

$$(C_{13} + C_{44}) \nu A_r + (C_{33} \nu^2 - C_{44}) A_z + (e_{33} \nu^2 - e_{15}) A_\phi = 0$$

$$(e_{31} + e_{15}) \nu A_r + (e_{33} \nu^2 - e_{15}) A_z + (\kappa_{11} - \kappa_{33} \nu^2) A_\phi = 0 \quad (10)$$

This is a system of homogeneous equations. To have a nontrivial solution, the value of ν must make the following determinant vanish

$$\begin{vmatrix}
-C_{11} + C_{44} \nu^2 & -(C_{13} + C_{44}) \nu & -(e_{31} + e_{15}) \nu \\
(C_{13} + C_{44}) \nu & C_{33} \nu^2 - C_{44} & e_{33} \nu^2 - e_{15} \\
(e_{31} + e_{15}) \nu & e_{33} \nu^2 - e_{15} & \kappa_{11} - \kappa_{33} \nu^2
\end{vmatrix} = 0 \quad (11)$$

The above condition yields a polynomial function of six orders, which is exactly the characteristic equation introduced in Chiang and Weng [34], a variant of those in Dunn and Wienecke [35]. Ignoring the degenerate conditions, two cases are possible:

$$(i) \quad \nu = \pm p, \quad \pm q, \pm r \quad (12)$$

or

$$(ii) \quad \nu = \pm p, \quad \pm q \pm ir \quad (13)$$

where p , q , and r are real and positive (this r appears here only momentarily and should not be confused with the radial distance). Since it is sufficient to consider only a half-space for the present problem, we shall focus on the region $z < 0$, so only the roots with the positive real part—that is, p , q , r in (i), or p , $q + ir$, $q - ir$ in (ii)—will be taken. Let the three roots be denoted by ν_1 , ν_2 , and ν_3 , and each eigen number corresponds to an eigen “vector,” that is, (A_r, A_z, A_ϕ) . The solution then can be written in the following form

$$\begin{aligned}
u_r &= \sum_{\alpha=1}^3 A_{r\alpha} \int_0^\infty Q(s) J_1(sr) \exp(\nu_\alpha s z) ds \\
u_z &= \sum_{\alpha=1}^3 A_{z\alpha} \int_0^\infty Q(s) J_0(sr) \exp(\nu_\alpha s z) ds \\
\phi &= \sum_{\alpha=1}^3 A_{\phi\alpha} \int_0^\infty Q(s) J_0(sr) \exp(\nu_\alpha s z) ds \quad (14)
\end{aligned}$$

As the components of eigenvectors must satisfy the homogeneous equations Eq. (10), it implies that, for a given ν_α , we have

$$\beta_{\phi\alpha}^z \equiv \frac{A_{z\alpha}}{A_{\phi\alpha}} = \frac{(C_{11} - C_{44} \nu_\alpha^2)(e_{33} \nu_\alpha^2 - e_{15}) - (C_{13} + C_{44})(e_{31} + e_{15}) \nu_\alpha^2}{(C_{13} + C_{44})^2 \nu_\alpha^2 - (C_{33} \nu_\alpha^2 - C_{44})(C_{11} - C_{44} \nu_\alpha^2)} \quad (15)$$

$$\begin{aligned}
\beta_{\phi\alpha}^r &\equiv \frac{A_{r\alpha}}{A_{\phi\alpha}} \\
&= \frac{(C_{33} \nu_\alpha^2 - C_{44})(e_{31} + e_{15}) \nu_\alpha - (C_{13} + C_{44}) \nu_\alpha (e_{33} \nu_\alpha^2 - e_{15})}{(C_{13} + C_{44})^2 \nu_\alpha^2 - (C_{33} \nu_\alpha^2 - C_{44})(C_{11} - C_{44} \nu_\alpha^2)} \quad (16)
\end{aligned}$$

where the β are dependent only on the ceramic property, not on the boundary conditions. In this way we just have three unknowns $A_{\phi 1}$, $A_{\phi 2}$, and $A_{\phi 3}$ remaining to be determined. In reality we may set $A_{\phi 1} = 1$ (i.e., choose the real eigen number as ν_1) without loss of generality, for it is only a normalization factor for the function $Q(s)$, while the other two can be determined from the boundary conditions.

This set of eigenvalues is in fact connected to that of the Stroh–Lekhnitskii formalism [52–54] by $p_\alpha = i \nu_\alpha$, where p_α are their eigenvalues (see, for instance, Eq. (C7) of Suo et al. [7] and Eq. (18) of Sosa [5]). A similar connection was previously established by Chiang [55] in the elastic context.

3.1 Boundary Conditions. It follows that, on the boundary $z=0$, the fields are given by

$$\tau_{zr} = \sum_{\alpha=1}^3 [C_{44}(A_{r\alpha} \nu_\alpha - A_{z\alpha}) - e_{15} A_{\phi\alpha}] \int_0^\infty s Q(s) J_1(sr) ds \quad (17)$$

$$\sigma_z = \sum_{\alpha=1}^3 (C_{13} A_{r\alpha} + C_{33} \nu_\alpha A_{z\alpha} + e_{33} \nu_\alpha A_{\phi\alpha}) \int_0^\infty s Q(s) J_0(sr) ds \quad (18)$$

$$D_z = \sum_{\alpha=1}^3 (e_{31} A_{r\alpha} + e_{33} \nu_\alpha A_{z\alpha} - \kappa_{33} \nu_\alpha A_{\phi\alpha}) \int_0^\infty s Q(s) J_0(sr) ds \quad (19)$$

$$u_z = \sum_{\alpha=1}^3 A_{z\alpha} \int_0^\infty Q(s) J_0(sr) ds \quad (20)$$

$$\phi = \sum_{\alpha=1}^3 A_{\phi\alpha} \int_0^\infty Q(s) J_0(sr) ds \quad (21)$$

Since $\tau_{zr}=0$

$$\sum_{\alpha=1}^3 [C_{44}(\nu_\alpha A_{r\alpha} - A_{z\alpha}) - e_{15} A_{\phi\alpha}] = 0 \quad (22)$$

Furthermore, u_z and ϕ both satisfy

$$\sum_{\alpha=1}^3 A_{z\alpha} \int_0^\infty Q(s) J_0(sr) ds = 0, \quad r > a \quad (23)$$

$$\sum_{\alpha=1}^3 A_{\phi\alpha} \int_0^\infty Q(s) J_0(sr) ds = 0, \quad r > a \quad (24)$$

Now for brevity we set

$$M^\sigma \equiv \sum_{\alpha=1}^3 m_\alpha^\sigma \equiv \sum_{\alpha=1}^3 (C_{13} A_{r\alpha} + C_{33} \nu_\alpha A_{z\alpha} + e_{33} \nu_\alpha A_{\phi\alpha}) \quad (25)$$

$$M^D \equiv \sum_{\alpha=1}^3 m_\alpha^D \equiv \sum_{\alpha=1}^3 (e_{31} A_{r\alpha} + e_{33} \nu_\alpha A_{z\alpha} - \kappa_{33} \nu_\alpha A_{\phi\alpha}) \quad (26)$$

which, in view of Eqs. (15) and (16), can be rewritten as

$$M^\sigma \equiv \sum_{\alpha=1}^3 m_\alpha^\sigma \equiv \sum_{\alpha=1}^3 (C_{13} \beta_{\phi\alpha}^r + C_{33} \nu_\alpha \beta_{\phi\alpha}^z + e_{33} \nu_\alpha) A_{\phi\alpha} \quad (27)$$

$$M^D \equiv \sum_{\alpha=1}^3 m_\alpha^D \equiv \sum_{\alpha=1}^3 (e_{31} \beta_{\phi\alpha}^r + e_{33} \nu_\alpha \beta_{\phi\alpha}^z - \kappa_{33} \nu_\alpha) A_{\phi\alpha} \quad (28)$$

It turns out that the ratio, M^D/M^σ will play a central role in the response of the dielectric crack.

The permittivity of the crack medium obviously has no effect on the mechanical boundary condition, which can be simply set by $-\sigma_z^A$, as

$$M^\sigma \int_0^\infty s Q(s) J_0(sr) ds = -\sigma_z^A, \quad 0 < r < a \quad (29)$$

where we have taken the origin to be at the center of the crack, and a is the crack radius.

The permittivity of the crack medium on the other hand will affect the electric boundary condition. We may consider that its influence is to introduce an extra electric displacement that is caused by the deformation of the crack. This extra electric displacement is given by

$$\Delta D^c = -\kappa_0 \frac{\Delta \phi}{\Delta u_z} \quad (30)$$

So on the crack surfaces instead of imposing $-D_z^A$, we now have to impose the total one, $-D^c$, that is

$$-D^c = -D_z^A + \Delta D^c$$

or

$$D^c = D_z^A - \Delta D^c \quad (31)$$

and the electric boundary condition becomes

$$M^D \int_0^\infty s Q(s) J_0(sr) ds = -D^c, \quad 0 < r < a \quad (32)$$

The total electric displacement D^c will henceforth be called the effective electric displacement of the dielectric crack. Its magnitude defines the electric response of the dielectric crack, and it also provides the connection between the electric displacement intensity factor K_D and the stress intensity factor, K_I .

We now solve these integral equations.

4 Impermeable and Permeable Cracks

Before we proceed to analyze the response of the dielectric crack, it is of interest to explore the consequences of these governing equations for the two limiting cases of $\kappa_0=0$ and $\kappa_0 \rightarrow \infty$ first. These correspond to the conditions of an impermeable and a permeable crack, respectively, and the problem is linear.

4.1 Impermeable Crack: $\kappa_0=0$. It proves advantageous to consider the inhomogeneous boundary conditions Eqs. (29) and (32) separately. We first solve the dual integral equations by assuming $D_z^A=0$. The corresponding eigenvectors must be chosen to satisfy this condition. Since $\sum_{\alpha=1}^3 A_{z\alpha} \neq 0$ in general, standard dual integral equations yield

$$\int_0^\infty s Q^\sigma(s) J_0(sr) ds = -\frac{\sigma_z^A}{M^\sigma}, \quad 0 < r < a \quad (33)$$

$$\int_0^\infty Q^\sigma(s) J_0(sr) ds = 0, \quad r > a \quad (34)$$

The solution to the equations is well known [51,56]

$$Q^\sigma(s) = \frac{-2}{\pi M^\sigma} \int_0^a \sin(st) dt \int_0^t \frac{r \sigma_z^A dr}{(t^2 - r^2)^{1/2}} \quad (35)$$

Likewise by taking $\sigma_z^A=0$, we can write

$$Q^D(s) = \frac{-2}{\pi M^D} \int_0^a \sin(st) dt \int_0^t \frac{r D_z^A dr}{(t^2 - r^2)^{1/2}} \quad (36)$$

We first solve for $Q^\sigma(s)$. With $A_{\phi 1}=1$ and from the boundary conditions Eqs. (29) and (32) (setting $D^c=0$), we obtain

$$A_{\phi 2} = \frac{-n_1 c_3 + n_3 c_1}{n_2 c_3 - n_3 c_2}, \quad A_{\phi 3} = \frac{-n_2 c_1 + n_1 c_2}{n_2 c_3 - n_3 c_2} \quad (37)$$

where

$$n_\alpha = C_{44} \nu_\alpha \beta_{\phi\alpha}^r - C_{44} \beta_{\phi\alpha}^z - e_{15}$$

and

$$c_\alpha = e_{31} \beta_{\phi\alpha}^r + e_{33} \nu_\alpha \beta_{\phi\alpha}^z - \kappa_{33} \nu_\alpha \quad (38)$$

The corresponding results for $Q^D(s)$ are similar to Eq. (37) by simply replacing c with the following definition

$$c_\alpha = C_{13} \beta_{\phi\alpha}^r + C_{33} \nu_\alpha \beta_{\phi\alpha}^z + e_{33} \nu_\alpha \quad (39)$$

This completes the full field solution by superposition of the two.

For the stress and electric displacement intensity factors, we have

$$K_I = \frac{1}{\sqrt{\pi a}} \int_0^a \frac{r \sigma_z^A(r) dr}{\sqrt{a^2 - r^2}} \quad (40)$$

and

Table 1 Material constants of PZT-4 from Park and Sun^a, PZT-5H from Pak^b, and PZT-5 and PZT-7A from Dunn and Taya^c

	C_{11} (GPa)	C_{12} (GPa)	C_{13} (GPa)	C_{33} (GPa)	C_{44} (GPa)	e_{31} (C/m ²)	e_{33} (C/m ²)	e_{15} (C/m ²)	κ_{11} (10 ⁻⁹ C/V m)	κ_{33} (10 ⁻⁹ C/V m)
PZT-4	139	77.8	74.3	113	25.6	-6.98	13.84	13.44	6	5.47
PZT-5	121	75.4	75.2	111	21.1	-5.4	15.8	12.3	8.11	7.35
PZT-5H	126	55	53	117	35.3	-6.5	23.3	17	15.1	13
PZT-7A	148	76.2	74.2	131	25.4	-2.1	9.5	9.2	4.07	2.08

^aSee Ref. 8.^bSee Ref. 4.^cSee Ref. 58.

$$K_D = \frac{1}{\sqrt{\pi a}} \int_0^a \frac{r D_z^A(r) dr}{\sqrt{a^2 - r^2}} \quad (41)$$

as there is no coupling on $z=0$. When the applied stress and electric displacement are uniform, the stress intensity factor K_I and electric displacement intensity factor K_D reduce to

$$K_I = \frac{2}{\pi} \sigma_z^A \sqrt{\pi a}, \quad K_D = \frac{2}{\pi} D_z^A \sqrt{\pi a} \quad (42)$$

We thus recover the well-known results that both intensity factors depend only on their respective loads, and that they are independent of the ceramic property.

4.2 Permeable or Conducting Crack: $\kappa_0 \rightarrow \infty$. On $z=0$, the electric potential ϕ may be set to null. This implies that

$$\sum_{\alpha=1}^3 A_{\phi\alpha} = 0 \quad (43)$$

Again by setting $A_{\phi 1}=1$, the other two eigenvectors follow from Eqs. (22) and (43), or more explicitly using the definitions of β in Eqs. (15) and (16)

$$A_{\phi 2} = \frac{(-\beta_{\phi 1}^r \nu_1 + \beta_{\phi 1}^c) - (-\beta_{\phi 3}^r \nu_3 + \beta_{\phi 3}^c)}{(-\beta_{\phi 3}^r \nu_3 + \beta_{\phi 3}^c) - (-\beta_{\phi 2}^r \nu_2 + \beta_{\phi 2}^c)} \quad (44)$$

$$A_{\phi 3} = \frac{(-\beta_{\phi 2}^r \nu_2 + \beta_{\phi 2}^c) - (-\beta_{\phi 1}^r \nu_1 + \beta_{\phi 1}^c)}{(-\beta_{\phi 3}^r \nu_3 + \beta_{\phi 3}^c) - (-\beta_{\phi 2}^r \nu_2 + \beta_{\phi 2}^c)} \quad (45)$$

Since the mechanical boundary condition remains unchanged, the solution is still given by Eq. (35). As a consequence the stress intensity factor K_I is exactly the same as in the impermeable case and depends only on the applied stress, σ_z^A . The electric displacement intensity factor K_D , on the other hand, is related to K_I through $K_D/K_I = M^D/M^\sigma$, and the latter ratio is given in terms of β_ϕ and A_ϕ which depend only on the ceramic property. As a consequence K_D is linearly proportional to K_I and is independent of the applied electric field D_z^A . Apparently for a fully permeable crack, the electric displacement intensity factor is not caused by the electric condition imposed on the boundary, but is induced by the piezoelectric coupling of the ceramic.

After computing the M^D/M^σ ratio for the four types of PZT listed in Table 1, we find

PZT-4

$$K_D = 0.25330 \times 10^{-9} (\text{C/N}) K_I$$

PZT-5

$$K_D = 0.30352 \times 10^{-9} (\text{C/N}) K_I$$

PZT-5H

$$K_D = 0.27906 \times 10^{-9} (\text{C/N}) K_I \quad (46)$$

PZT-7A

$$K_D = 0.12819 \times 10^{-9} (\text{C/N}) K_I$$

This linear connection between K_D and K_I is believed to be valid not only for the penny-shaped crack but also for a 2D crack. We will illustrate this point in the following example.

McMeeking [57] has conducted a theoretical study on the energy release rate of a crack in piezoelectric materials and provided numerical results for a compact tension specimen having the material properties of PZT-4. In Eq. (A16) of his paper, the energy release rate of an impermeable crack for that particular configuration was

$$G = 0.9884 \frac{K_I^2}{C_{33}} + 0.5505 \frac{K_I K_D}{\sqrt{C_{33} \kappa_{33}}} - 0.2391 \frac{K_D^2}{\kappa_{33}} \quad (47)$$

To find the result for a conducting crack corresponding to the same configuration, we may simply replace K_D by $0.25330 \times 10^{-9} (\text{C/N}) K_I$ shown in Eq. (46). After substituting the material constants of PZT-4 into Eq. (47), we find

$$G = 1.3052 \frac{K_I^2}{C_{33}} \quad (48)$$

An almost identical result was given in his Eq. (A8). The difference in the last digit is believed to be due to the rounding error.

5 Dielectric Crack With a Finite Permittivity κ_0

Now we examine the general problem of a crack with finite dielectric permittivity. Due to the coupling of the traction and electric displacement on the crack surfaces, this problem is nonlinear. As a consequence we cannot decompose $Q(s)$ into $Q^\sigma(s)$ and $Q^D(s)$, and solve the mechanical and electric boundary-value problems separately. This problem needs to be solved under the simultaneous electromechanical boundary conditions.

First we note that the extra electric displacement on the crack surface is given by

$$\Delta D^c = -\kappa_0 \frac{\Delta \phi}{\Delta u_z} = -\kappa_0 \frac{\sum_{\alpha=1}^3 A_{\phi\alpha}}{\sum_{\alpha=1}^3 A_{z\alpha}} \quad (49)$$

Furthermore, from Eqs. (29) and (32), we have

$$\frac{\sigma_z^A}{M^\sigma} = \frac{D_z^A - \Delta D^c}{M^D} \quad (50)$$

On the plane $z=0$, $\tau_{zr}=0$ and so Eq. (22) still holds. Thus for a given σ_z^A and D_z^A , and using Eqs. (22) and (50), we find that, after some lengthy algebra, $A_{\phi 3}$ is the root of the following quadratic equation

$$b_2 A_{\phi 3}^2 + b_1 A_{\phi 3} + b_0 = 0 \quad (51)$$

with

$$b_2 = -\kappa_0 (n_2 - n_3) (n_2 m_3^\sigma - n_3 m_2^\sigma) - D_z^A (n_2 \beta_{\phi 3}^c - n_3 \beta_{\phi 2}^c) \\ \times (n_2 m_3^\sigma - n_3 m_2^\sigma) + \sigma_z^A (n_2 \beta_{\phi 3}^c - n_3 \beta_{\phi 2}^c) (n_2 m_3^D - n_3 m_2^D)$$

$$\begin{aligned}
b_1 = & -\kappa_0[(n_2 - n_3)(n_2 m_1^\sigma - n_1 m_2^\sigma) + (n_2 - n_1)(n_2 m_3^\sigma - n_3 m_2^\sigma)] \\
& - D_z^A[(n_2 \beta_{\phi 3}^\varepsilon - n_3 \beta_{\phi 2}^\varepsilon)(n_2 m_1^\sigma - n_1 m_2^\sigma) + (n_2 \beta_{\phi 1}^\varepsilon - n_1 \beta_{\phi 2}^\varepsilon) \\
& \times (n_2 m_3^\sigma - n_3 m_2^\sigma)] + \sigma_z^A[(n_2 \beta_{\phi 3}^\varepsilon - n_3 \beta_{\phi 2}^\varepsilon)(n_2 m_1^D - n_1 m_2^D) \\
& + (n_2 \beta_{\phi 1}^\varepsilon - n_1 \beta_{\phi 2}^\varepsilon)(n_2 m_3^D - n_3 m_2^D)] \\
b_0 = & -\kappa_0(n_2 - n_1)(n_2 m_1^\sigma - n_1 m_2^\sigma) - D_z^A(n_2 \beta_{\phi 1}^\varepsilon - n_1 \beta_{\phi 2}^\varepsilon) \\
& \times (n_2 m_1^\sigma - n_1 m_2^\sigma) + \sigma_z^A(n_2 \beta_{\phi 1}^\varepsilon - n_1 \beta_{\phi 2}^\varepsilon)(n_2 m_1^D - n_1 m_2^D)
\end{aligned} \quad (52)$$

and

$$A_{\phi 2} = -\frac{n_1}{n_2} - \frac{n_3}{n_2} A_{\phi 3} \quad (53)$$

These constants are seen to depend on the applied stress σ_z^A , applied electric displacement D_z^A , and the dielectric constant of the crack medium κ_0 . Since $A_{\phi 1}=1$ and $A_{z\alpha}=\beta_{\phi\alpha}^\varepsilon A_{\phi\alpha}$, all pertinent quantities, such as M^σ , M^D , and D^c ($D^c=D_z^A-\Delta D^c$), can be determined.

The crack opening displacement and electric potential on $z=0$ then follow from

$$\begin{aligned}
u_z = & -\sum_{\alpha=1}^3 A_{z\alpha} \frac{\sigma_z^A}{2\pi M^\sigma} \sqrt{a^2 - r^2}, \quad r < a \\
\phi = & -\sum_{\alpha=1}^3 A_{\phi\alpha} \frac{D^c}{2\pi M^D} \sqrt{a^2 - r^2}, \quad r < a
\end{aligned} \quad (54)$$

As the electric displacement intensity factor K_D is given by

$$K_D = \frac{2}{\pi} D^c \sqrt{\pi a} \quad (55)$$

it is evident that the relation

$$\frac{D^c}{\sigma_z^A} = \frac{K_D}{K_I} = \frac{M^D}{M^\sigma} \quad (56)$$

holds for the dielectric crack. Since the ratio M^D/M^σ depends on σ_z^A , D_z^A , κ_0 , and the ceramic property, both the effective electric displacement D^c and the electric intensity factor K_D are nonlinear functions of these quantities.

Figure 1 illustrates the variation of the effective electric displacement D^c as a function of the applied stress σ_z^A for PZT-4, by taking the permittivity κ_0 as that of vacuum (i.e., $\kappa_0=8.85 \times 10^{-12}$ C/V m). It is seen that the effective D^c increases nonlinearly with σ_z^A , and that the level of increase is further enhanced by the applied D_z^A . Our calculations for the other three types of PZT are found to exhibit similar characteristics. In light of Eq. (56), the nonlinear D^c versus σ_z^A relations displayed here also reflect the nonlinear connection between K_D and K_I , and the ratio M^D/M^σ also serves as the “secant” modulus in the D^c versus σ_z^A plot.

Figure 2 is a schematic illustration on the influence of κ_0 to the effective electric displacement D^c as the applied σ_z^A increases at a given D_z^A . Here three different kinds of dielectric permittivity inside the crack: $\kappa_0=0$, $\rightarrow\infty$, and finite, corresponding to the cases of an impermeable crack, a conducting one, and one with a finite permittivity, are chosen. It is seen that when $\kappa_0=0$, the traction boundary condition has no influence on D^c , and $D^c=D_z^A$. When $\kappa_0\rightarrow\infty$, D^c is linearly proportional to σ_z^A and is in fact independent of the applied D_z^A . The proportional constant depends on the material properties of the ceramic. The one with a finite permittivity exhibits the nonlinear response as shown in Fig. 1. A key feature of this plot is that, regardless of the value of κ_0 , all the curves must pass through a critical point that is marked by (σ_c, D_z^A) . This critical stress, σ_c , is the applied σ_z^A that will ensure

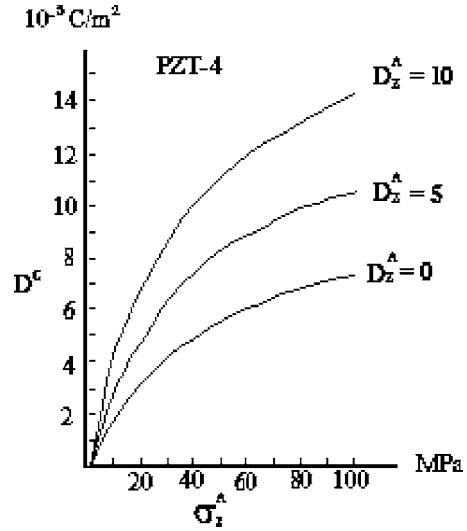


Fig. 1 The effective electric displacement, D^c , of the crack medium versus the applied stress, σ_z^A , for PZT-4. The permittivity κ_0 inside the crack was 8.85×10^{-12} C/V m.

the condition of $\Delta D^c=0$ for a given κ_0 . The existence of this common critical state for a penny-shaped crack is attributed to the relation Eq. (50). That is, with this combination of (σ_c, D_z^A) , any value of κ_0 would not change the condition of $\Delta D^c=0$, and the stress and electric displacement of the crack surfaces will be exactly equal to the applied ones. This critical stress can be easily calculated from the intersection of the two straight lines and can be expressed as

$$\sigma_c = \left. \frac{M^\sigma}{M^D} \right|_{\text{permeable}} \cdot D_z^A \quad (57)$$

For the four types of PZT given in Table 1, they are

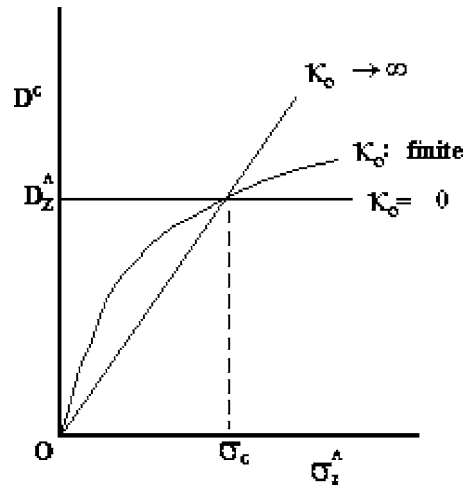


Fig. 2 A schematic plot on the influence of permittivity, κ_0 , of the crack medium to the effective electric displacement, D^c , versus the applied stress, σ_z^A relation. Regardless of the value of κ_0 all the curves must pass through the critical state (σ_c, D_z^A) . When $\sigma_z^A < \sigma_c$, the responses of the impermeable and permeable cracks will serve as the upper and lower bounds, respectively, and when $\sigma_z^A > \sigma_c$ the situation is reversed. The response of a dielectric crack with any κ_0 always lies within these bounds.

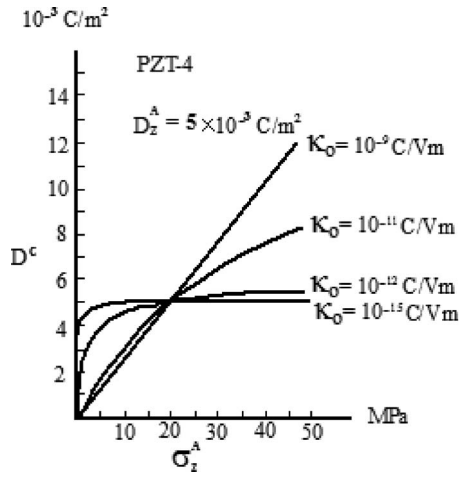


Fig. 3 A quantitative assessment for PZT-4 on the influence of permittivity κ_0 of the crack medium to the effective electric displacement, D^c , versus the applied stress, σ_z^A relation. The result with vacuum ($\kappa_0=8.85 \times 10^{-12}$ C/V m), lying between those of $\kappa_0=10^{-12}$ C/V m and 10^{-11} C/V m, is seen to be far away from those of the permeable and impermeable cracks.

PZT-4

$$\sigma_c = 3.9479 \times 10^9 (\text{N/C}) D_z^A$$

PZT-5

$$\sigma_c = 3.2947 \times 10^9 (\text{N/C}) D_z^A$$

PZT-5H

$$\sigma_c = 3.5834 \times 10^9 (\text{N/C}) D_z^A \quad (58)$$

PZT-7A

$$\sigma_c = 7.8012 \times 10^9 (\text{N/C}) D_z^A$$

These proportional constants are exactly the inverse of those in Eq. (46).

Since the electric displacement intensity factor is calculated from Eq. (55), we may further conclude that, when the applied stress is below σ_c , K_D of an impermeable crack will serve as the upper bound and that of a permeable one will serve as the lower bound of the dielectric crack. In fact, the lower the value of κ_0 , the higher the intensity K_D . On the other hand when the applied stress is greater than σ_c , the behavior is exactly reversed. The K_D of a crack with a finite permittivity always lies within these bounds.

This critical state (σ_c, D_z^A), provides a defining point at which the relative characteristics of the dielectric crack with two different dielectric permittivities are reversed. Its existence appears not to have been reported in the literature.

More detailed quantitative results are illustrated in Fig. 3 for PZT-4 under the condition of $D_z^A = 5 \times 10^{-3}$ C/m². The critical stress σ_c here is 19.74 MPa. Four selected κ_0 values—ranging from 10^{-15} C/V m to 10^{-9} C/V m—are used to compute the D^c versus σ_z^A relations. It is seen that the result with $\kappa_0 = 10^{-9}$ C/V m is quite linear and thus such a crack may effectively be treated as a fully permeable one, but that with $\kappa_0 = 10^{-15}$ C/V m still exhibits certain nonlinearity in the beginning and thus is not exactly an impermeable case. The line with the dielectric constant of vacuum, $\kappa_0 = 8.85 \times 10^{-12}$ C/V m, lies between those of the two middle curves, and it is evident that, even with vacuum, a crack cannot be simply treated as an impermeable (or a permeable) crack.

Finally we present the nonlinear response of the dielectric crack when a negative electric displacement, D_z^A , is applied. The nonlinear D^c versus σ_z^A relations of the dielectric crack with different κ_0

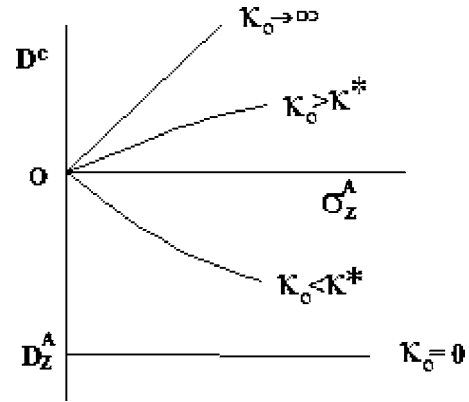


Fig. 4 A schematic plot on the influence of permittivity, κ_0 , of the crack medium to the effective electric displacement, D^c , versus the applied stress, σ_z^A , relation under a negative D_z^A . The critical (σ_c, D_z^A) state in this case only exists at the origin. There exists a critical κ^* for κ_0 , beyond which D^c will remain positive in spite of the negative D_z^A .

values are schematically shown in Fig. 4. In this case the critical state (σ_c, D_z^A) does not depend on the properties of the ceramic, and it can only exist at the origin. All the D^c versus σ_z^A curves are bounded from above by the permeable crack and below by the impermeable one. Furthermore, there exists a transition permittivity κ^* inside the crack that separates, the nonlinear relations from an increase to a decrease of the D^c field as the applied stress σ_z^A increases. At this transition κ^* , the effective electric displacement D^c is always zero, independent of σ_z^A . Its value is a function of the material constants of the ceramic, and it depends linearly on the applied D_z^A , as

$$\kappa^* = -RD_z^A \quad (59)$$

where the proportional constant R for the four ceramics listed in Table 1 are found to be

PZT-4

$$R = 0.78996 \times 10^{-9} \text{ m/V}$$

PZT-5

$$R = 0.98107 \times 10^{-9} \text{ m/V}$$

PZT-5H

$$R = 1.2575 \times 10^{-9} \text{ m/V}$$

PZT-7A

$$R = 0.65419 \times 10^{-9} \text{ m/V} \quad (60)$$

For a PZT-4 under the negative electric displacement $D_z^A = -5 \times 10^{-3}$ C/m², the transition permittivity is $\kappa^* = 3.95 \times 10^{-12}$ C/V m. The corresponding nonlinear relations with four selected κ_0 values are shown in Fig. 5. The horizontal axis then serves as the demarcation line for the response of these two groups of dielectric crack whose permittivity is greater or smaller than this critical value.

Relation (59) in retrospect can be used to find the level of “transition” electric load D^* for a given dielectric permittivity κ_0 , as

$$D^* = -\kappa_0/R \quad (61)$$

When $|D_z^A| < |D^*|$, the induced D^c will still remain positive because the applied stress still dominates the field; only when $|D_z^A| > |D^*|$, will D^c become negative like D_z^A . For a dielectric crack with the permittivity of vacuum ($\kappa_0 = 8.85 \times 10^{-12}$ C/V m) in a

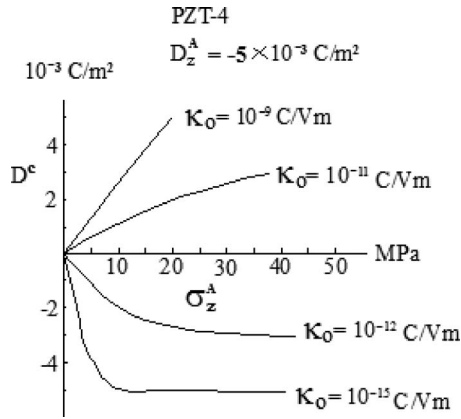


Fig. 5 A quantitative assessment for PZT-4 on the influence of permittivity κ_0 of the crack medium to the effective electric displacement, D^c , versus the applied stress, σ_z^A , relation under a negative D_z^A .

PZT-4, we have $D^* = -11.2 \times 10^{-3} \text{ C/m}^2$. Such a phenomenon can be seen in Fig. 6.

6 Concluding Remarks

By means of the Hankel transform and dual-integral equations, the problem of a penny-shaped dielectric crack in a transversely isotropic piezoelectric medium is solved explicitly. In addition to some well-known results for the intensity factors of impermeable and permeable cracks, our solutions have vividly displayed the nonlinear characteristics of a dielectric crack under a tensile stress and axial electric field. The general results can be cast in the universal relation, as

$$\frac{D^c}{\sigma_z^A} = \frac{K_D}{K_I} = \frac{M^D}{M^\sigma} \quad (62)$$

regardless of the electric boundary conditions. For an impermeable crack the ratio M^D/M^σ is equal to D_z^A/σ_z^A . For a conductive crack this ratio is a function of the material property of the ceramic only, independent of either D_z^A or σ_z^A . For a general dielectric crack with a finite permittivity κ_0 , this ratio depends on the material property of the ceramic, dielectric constant of the crack medium κ_0 , and the applied D_z^A and σ_z^A . The latter dependence makes the response of the crack medium nonlinear.

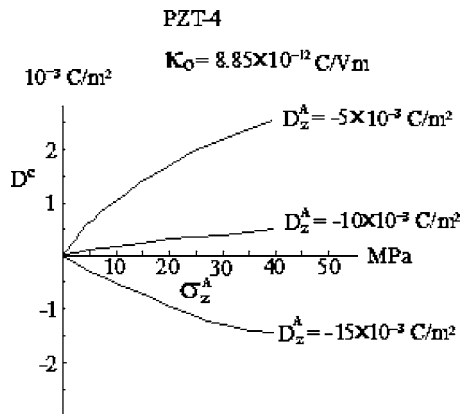


Fig. 6 A quantitative assessment for PZT-4 on the influence of the negative electric load, D_z^A , to the nonlinear relation of D^c versus σ_z^A . There exists a critical D^* , below which the effective D^c will remain positive in spite of the negative D_z^A .

We have demonstrated the nonlinear characteristics of the dielectric crack response by showing the nonlinear growth of D^c as a function of σ_z^A . The growth is monotonically enhanced by a positive D_z^A . The nonlinear dependence of D^c on σ_z^A is found to be closely controlled by a critical state (σ_c, D_z^A) , through which every single curve passes regardless of the dielectric constant κ_0 . When σ_z^A is lower than σ_c , the D^c of the impermeable crack provides the upper bound and that of the permeable crack provides the lower bound. When σ_z^A is greater than σ_c , the situation is exactly reversed. The nonlinear response of the dielectric crack always lies within these bounds, and it undergoes a transition as σ_z^A passes through σ_c . Due to relation (62), the relation between K_D and K_I also exhibits similar transition as the applied stress passes through the critical state. This critical stress depends only on the material constants of the ceramic and is independent of the dielectric permittivity κ_0 of the crack medium.

When the applied electric displacement D_z^A is negative, the critical state (σ_c, D_z^A) can only exist at the origin. In this case there exists a critical dielectric constant κ^* that is linearly proportional to D_z^A in the form of $\kappa^* = -RD_z^A$, where the proportional constant R depends only on the property of the ceramic. For a dielectric crack whose κ_0 is greater than κ^* , its effective electric displacement D^c will continue to be positive due to the greater influence of the tensile stress, but for a crack medium whose κ_0 is lower than κ^* , its D^c will be subjected to a greater influence by D_z^A and turns into the negative range. The values of all these curves, however, always lie between those of the permeable and impermeable ones that also serve as the upper and lower bounds, respectively.

Under a negative D_z^A , the response of a dielectric crack with a given κ_0 also depends on a critical electric displacement D^* . This critical D^* is given by $D^* = -\kappa_0/R$. When the negative applied load D_z^A has a lower magnitude than that of D^* , the effective electric displacement D^c will still be positive but, when it is greater, the D^c of the crack medium will also turn negative.

We have thus demonstrated the nonlinear nature of the crack response, and identified several key critical states that have significant implications on the nonlinear response and the electric intensity factor of the dielectric crack.

Acknowledgment

The study of CRC was supported partly by National Science Council of Taiwan, and that of GJW by NSF through Grant No. CMS-0510409.

References

- [1] Parton, V. Z., 1976, "Fracture Mechanics of Piezoelectric Materials," *Acta Astronaut.*, **3**, pp. 671–683.
- [2] Deeg, W. F., 1980, "The Analysis of Dislocation, Crack and Inclusion Problems in Piezoelectric Solids," Ph.D. dissertation, Stanford University, Stanford, CA.
- [3] Pak, Y. E., 1990, "Crack Extension Force in a Piezoelectric Material," *J. Appl. Mech.*, **57**, pp. 647–653.
- [4] Pak, Y. E., 1992, "Linear Electro-elastic Fracture Mechanics of Piezoelectric Materials," *Int. J. Fract.*, **54**, pp. 79–100.
- [5] Sosa, H., 1991, "Plane Problems in Piezoelectric Media with Defects," *Int. J. Solids Struct.*, **28**, pp. 491–505.
- [6] Sosa, H., 1992, "On the Fracture Mechanics of Piezoelectric Solids," *Int. J. Solids Struct.*, **29**, pp. 2613–2622.
- [7] Suo, Z., Kuo, C.-M., Barnett, D. M., and Willis, J. R., 1992, "Fracture Mechanics for Piezoelectric Ceramics," *J. Mech. Phys. Solids*, **40**, pp. 739–765.
- [8] Park, S. B., and Sun, C. T., 1995, "Effect of Electric Field on Fracture of Piezoelectric Ceramics," *Int. J. Fract.*, **70**, pp. 203–216.
- [9] McMeeking, R. M., 1989, "Electrostrictive Stress near Crack-like Flaws," *J. Appl. Math.*, **40**, pp. 615–627.
- [10] Zhang, T.-Y., and Hack, J. E., 1992, "Mode III Cracks in Piezoelectric Materials," *J. Appl. Phys.*, **71**, pp. 5865–5870.
- [11] Zhang, T.-Y., 1994, "Effect of Sample Width on the Energy Release Rate and Electric Boundary Conditions along Crack Surfaces in Piezoelectric Materials," *Int. J. Fract.*, **66**, pp. 33–38.
- [12] Dunn, M. L., 1994, "The Effect of Crack Face Boundary Conditions on the Fracture Mechanics of Piezoelectric Solid," *Eng. Fract. Mech.*, **48**, pp. 25–39.
- [13] Sosa, H., and Khutoryansky, N., 1996, "New Developments Concerning Pi-

- ezoelectric Materials with Defects," *Int. J. Solids Struct.*, **33**, pp. 3399–3414.
- [14] Zhang, T.-Y., Zhao, M.-H., and Tong, P., 2001, "Fracture of Piezoelectric Ceramics," *Adv. Appl. Mech.*, **38**, pp. 147–289.
 - [15] Li, S. F., and Mataga, P. A., 1996, "Dynamic Crack Propagation in Piezoelectric Materials. Part I: Electrode Solution," *J. Mech. Phys. Solids*, **44**, pp. 1799–1830.
 - [16] Li, S. F., and Mataga, P. A., 1996, "Dynamic Crack Propagation in Piezoelectric Materials. Part II: Vacuum," *J. Mech. Phys. Solids*, **44**, pp. 1831–1866.
 - [17] Wang, X. D., and Meguid, S. A., 2000, "Effect of Electromechanical Coupling on the Dynamic Interaction of Cracks in Piezoelectric Materials," *Acta Mech.*, **143**, pp. 1–15.
 - [18] Li, C. Y., and Weng, G. J., 2002, "Yoffe-type Moving Crack in a Functionally Graded Piezoelectric Material," *Proc. R. Soc. London, Ser. A*, **458**, pp. 381–399.
 - [19] Parton, V. Z., and Kudryavtsev, B. A., 1988, *Electromagnetoelasticity*, Gordon and Breach, New York.
 - [20] Hao, T. H., and Shen, Z. Y., 1994, "A New Electric Boundary Condition of Electric Fracture Mechanics and Its Applications," *Eng. Fract. Mech.*, **47**, pp. 793–802.
 - [21] Shindo, Y., Tanaka, K., and Narita, F., 1997, "Singular Stress and Electric Fields of a Piezoelectric Ceramic Strip with a Finite Crack under Longitudinal Shear," *Acta Mech.*, **120**, pp. 31–45.
 - [22] Yang, F., 2001, "Fracture Mechanics for a Mode I crack in Piezoelectric Materials," *Int. J. Solids Struct.*, **38**, pp. 3813–3830.
 - [23] Shindo, Y., Watanabe, K., and Narita, F., 2000, "Electroelastic Analysis of a Piezoelectric Ceramic Strip with a Central Crack," *Int. J. Eng. Sci.*, **38**, pp. 1–19.
 - [24] McMeeking, R. M., 2001, "Towards a Fracture Mechanics for Brittle Piezoelectric and Dielectric Materials," *Int. J. Fract.*, **108**, pp. 25–41.
 - [25] Xu, X.-L., and Rajapakse, R. K. N. D., 2001, "On a Plane Crack in Piezoelectric Solids," *Int. J. Solids Struct.*, **38**, pp. 7643–7658.
 - [26] Wang, X. D., and Jiang, L. Y., 2001, "Fracture Behavior of Cracks in Piezoelectric Media with Electromechanically Coupled Boundary Conditions," *Proc. R. Soc. London, Ser. A*, **458**, pp. 2545–2560.
 - [27] Dascalu, C., and Homentcovschi, D., 2002, "An Intermediate Crack Model for Flaws in Piezoelectric Solids," *Acta Mech.*, **154**, pp. 85–100.
 - [28] Wang, B. L., and Mai, Y.-W., 2003, "On the Electrical Boundary Conditions on the Crack Surfaces in Piezoelectric Ceramics," *Int. J. Eng. Sci.*, **41**, pp. 633–652.
 - [29] Zhang, T. Y., and Cao, C. F., 2004, "Fracture Behaviors of Piezoelectric Materials," *Theor. Appl. Fract. Mech.*, **41**, pp. 339–379.
 - [30] Eshelby, J. D., 1957, "The Determination of the Elastic Field of an Ellipsoidal Inclusion, and Related Problems," *Proc. R. Soc. London, Ser. A*, **241**, pp. 376–396.
 - [31] Wang, B., 1992, "Three-Dimensional Analysis of a Flat Elliptical Crack in a Piezoelectric Material," *Int. J. Eng. Sci.*, **30**, pp. 781–791.
 - [32] Kogan, L., Hui, C.-Y., and Molokov, V., 1996, "Stress and Induced Field of a Spheroidal Inclusion or a Penny-Shaped Crack in a Transversely Isotropic Piezoelectric Material," *Int. J. Solids Struct.*, **33**, pp. 2719–2737.
 - [33] Huang, J. H., 1997, "A Fracture Criterion of a Penny-shaped Crack in Transversely Isotropic Piezoelectric Media," *Int. J. Solids Struct.*, **34**, pp. 2631–2644.
 - [34] Chiang, C. R., and Weng, G. J., 2005, "The Nature of Stress and Electric-displacement Concentration around a Strongly Oblate Cavity in a Transversely Isotropic Piezoelectric Material," *Int. J. Fract.*, **134**, pp. 319–337.
 - [35] Dunn, M. L., and Wienecke, H. A., 1997, "Inclusions and Inhomogeneity in Transversely Isotropic Piezoelectric Solids," *Int. J. Solids Struct.*, **34**, pp. 3571–3582.
 - [36] Chen, W. Q., and Shioya, T., 1999, "Fundamental Solution of a Penny-shaped Crack in a Piezoelectric Medium," *J. Mech. Phys. Solids*, **47**, pp. 1459–1475.
 - [37] Chen, W. Q., and Shioya, T., 1999, "Complete and Exact Solutions of a Penny-Shaped Crack in a Piezoelectric Solid: Antisymmetric Shear Loadings," *Int. J. Solids Struct.*, **37**, pp. 2603–2619.
 - [38] Fabricant, V. I., 1989, *Application of Potential Theory in Mechanics: A Selection of New Results*, Kluwer Academic, Dordrecht, The Netherlands.
 - [39] Ding, H. J., Chen, B., and Liang, J., 1997, "On the Green's Function for Two-phase Transversely Isotropic Piezoelectric Media," *Int. J. Solids Struct.*, **34**, pp. 3041–3057.
 - [40] Zhao, M. H., Shen, Y. P., Liu, G. N., and Liu, Y. J., 1999, "Crack Analysis in Semi-infinite Transversely Isotropic Piezoelectric Solid II. Penny-shaped Crack near the Surface," *Theor. Appl. Fract. Mech.*, **32**, pp. 233–240.
 - [41] Yang, J. H., and Lee, K. Y., 2001, "Penny-Shaped Crack in a Three-Dimensional Piezoelectric Strip under In-plane Normal Loadings," *Acta Mech.*, **148**, pp. 187–197.
 - [42] Yang, J. H., and Lee, K. Y., 2002, "Three-Dimensional Non-Axisymmetric Behavior of a Penny-Shaped Crack in a Piezoelectric Strip Subjected to In-Plane Loads," *Eur. J. Mech. A/Solids*, **21**, pp. 223–237.
 - [43] Lin, S., Narita, F., and Shindo, Y., 2003, "Electroelastic Analysis of a Penny-shaped Crack in a Piezoelectric Ceramic under Mode I Loading," *Mech. Res. Commun.*, **30**, pp. 371–386.
 - [44] Li, X. F., and Lee, K. Y., 2004, "Effects of Electric Field on Crack Growth for a Penny-shaped Dielectric Crack in a Piezoelectric Layer," *J. Mech. Phys. Solids*, **52**, pp. 2079–2100.
 - [45] Chen, W. Q., and Lim, C. W., 2005, "3D Point Force Solution for a Permeable Penny-Shaped Crack Embedded in an Infinite Transversely Isotropic Piezoelectric Medium," *Int. J. Fract.*, **131**, pp. 231–246.
 - [46] Hwang, S. C., Lynch, C. S., and McMeeking, R. M., 1995, "Ferroelectric/Ferroelastic Interactions and a Polarization Switching Model," *Acta Metall. Mater.*, **43**, pp. 2073–2084.
 - [47] Li, J., and Weng, G. J., 1999, "A Theory of Domain Switch for the Nonlinear Behavior of Ferroelectrics," *Proc. R. Soc. London, Ser. A*, **455**, pp. 3493–3511.
 - [48] Li, W. F., and Weng, G. J., 2002, "A Theory of Ferroelectric Hysteresis with a Superimposed Stress," *J. Appl. Phys.*, **91**, pp. 3806–3815.
 - [49] Ikeda, T., 1996, *Fundamentals of Piezoelectricity*, Oxford University Press, Oxford, UK.
 - [50] Nye, J. F., 1979, *Physical Properties of Crystals*, Oxford University Press, Oxford, UK.
 - [51] Sneddon, I. N., 1951, *Fourier Transforms*, McGraw-Hill, New York.
 - [52] Stroh, A. N., 1958, "Dislocations and Cracks in Anisotropic Elasticity," *Philos. Mag.*, **3**, pp. 625–646.
 - [53] Stroh, A. N., 1962, "Steady State Problems in Anisotropic Elasticity," *J. Math. Phys.*, **41**, pp. 77–103.
 - [54] Lekhnitskii, S. G., 1963, *Theory of Elasticity of an Anisotropic Elastic Body*, Holden-Day, San Francisco, CA.
 - [55] Chiang, C. R., 2004, "Some Crack Problems in Transversely Isotropic Solids," *Acta Mech.*, **170**, pp. 1–9.
 - [56] Kassir, M. K., and Sih, G. C., 1975, *Three-Dimensional Crack Problems*, Noordhoff, Leyden, The Netherlands.
 - [57] McMeeking, R. M., 1999, "Crack Tip Energy Release Rate for a Piezoelectric Compact Tension Specimen," *Eng. Fract. Mech.*, **64**, pp. 217–244.
 - [58] Dunn, M. L., and Taya, M., 1993, "An Analysis of Piezoelectric Composite Materials Containing Ellipsoidal Inhomogeneities," *Proc. R. Soc. London, Ser. A*, **443**, pp. 265–287.

Designing Optimal Volume Fractions For Functionally Graded Materials With Temperature-Dependent Material Properties

Florin Bobaru

Department of Engineering Mechanics,
University of Nebraska-Lincoln,
Lincoln, Nebraska 68588-0526
e-mail: fbaru2@unl.edu

We present a numerical approach for material optimization of metal-ceramic functionally graded materials (FGMs) with temperature-dependent material properties. We solve the non-linear heterogeneous thermoelasticity equations in 2D under plane strain conditions and consider examples in which the material composition varies along the radial direction of a hollow cylinder under thermomechanical loading. A space of shape-preserving splines is used to search for the optimal volume fraction function which minimizes stresses or minimizes mass under stress constraints. The control points (design variables) that define the volume fraction spline function are independent of the grid used in the numerical solution of the thermoelastic problem. We introduce new temperature-dependent objective functions and constraints. The rule of mixture and the modified Mori-Tanaka with the fuzzy inference scheme are used to compute effective properties for the material mixtures. The different micromechanics models lead to optimal solutions that are similar qualitatively. To compute the temperature-dependent critical stresses for the mixture, we use, for lack of experimental data, the rule-of-mixture. When a scalar stress measure is minimized, we obtain optimal volume fraction functions that feature multiple graded regions alternating with non-graded layers, or even non-monotonic profiles. The dominant factor for the existence of such local minimizers is the non-linear dependence of the critical stresses of the ceramic component on temperature. These results show that, in certain cases, using power-law type functions to represent the material gradation in FGMs is too restrictive. [DOI: 10.1115/1.2712231]

1 Introduction

1.1 Literature Review. Inspired by some naturally occurring material systems such as bamboo, bone, sea shells, teeth, functionally graded materials (FGMs) are heterogeneous materials featuring a smooth material composition gradation for achieving a specific function. Most commonly, metal-ceramic FGMs are used to make the transition from a metal part, which is strong but cannot operate at high temperatures, and a ceramic part, which is efficient in shielding against high temperatures but has low tensile strength. The smooth gradation in space of the volume fraction of one constituent reduces the stress concentrations or jumps, minimizing, therefore, the likelihood of material failure. Some current and potential applications include thermal protective shields of high reliability, thin films and coatings for cutting tools, and dental and prosthetic implants, to name just a few. Note that here we consider materials in which the microstructure scale is such that homogenization can be applied. The limitations and validity of micromechanics-based homogenization schemes for FGMs under thermal cycling conditions are analyzed in Pindera et al. [1], for example.

In most studies found in the literature, the defining parameter of a FGM is the function giving the volume fraction spatial variation for one of the two constituents. Other types of parameters can also be selected for describing the FGM, such as Young's modulus E and thermal expansion coefficient α as is done in, for example, [2,3], or [4]. In such cases, however, one has to solve an additional inverse problem to determine the corresponding volume fraction function, which is more relevant from a manufacturing

point of view. Such an inverse problem may be even more complicated by temperature-dependent material properties, and in fact, its solution might not even exist. Moreover, as noted in [5], it may not always be possible to impose the effective thermal expansion coefficient and Young's modulus independent from each other.

Among the publications, theoretical and numerical, which analyze the thermomechanical behavior of FGMs, a common feature is the selection of a power-law type function ($x^p, p \in \mathbb{R}$) for the volume fraction (or other material parameters). Here, x is a coordinate along which the material gradation takes places from one side (at $x=0$) of the material to the other (at $x=1$). The monotonic power law is selected only as a matter of convenience since analytical results are difficult, if not impossible, to obtain otherwise. On the other hand, such a gradation offers a large selection of possible gradations for FGMs (see Fig. 1). We stress, however, that the power-law function is unable to represent, for example, non-monotonic volume fraction variations.

One material optimization problem that can be defined for FGMs is to determine the optimal function for the material gradation in order to minimize, for example, thermal stresses in the FGM. Previous material optimization studies of FGMs under thermomechanical loadings [6–11] limit their design space to that of power-law type or monotonic functions. In Tanaka et al. [10], for example, the radial distribution volume fraction function is sought in a space of monotonic functions of a certain type: their first derivative is a positive “wedge” function. The design variables in this case are the locations of the start, tip, and end of the piecewise linear “wedge”-like derivative. The objective is to minimize the maximum departure of the circumferential stress (which, for the examples we consider here, is significantly larger than the other stress components in the hollow cylinder structure) from a given reference stress distribution under transient heat-conduction conditions. Effective material properties are obtained with the

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received March 18, 2006; final manuscript received October 10, 2006. Review conducted by Subrata Mukherjee.

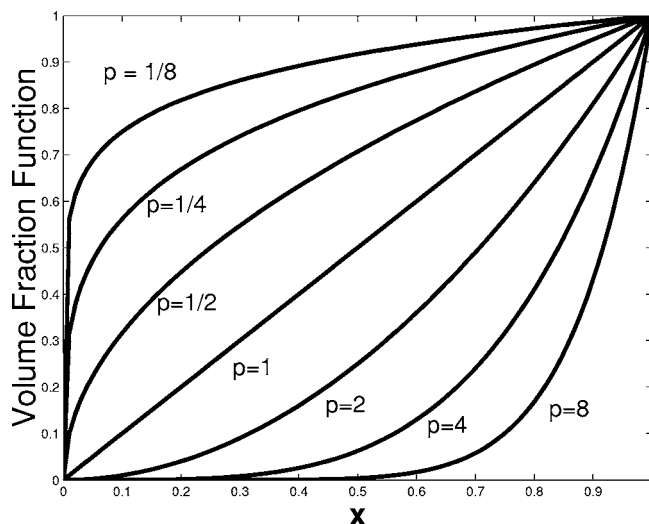


Fig. 1 Various realizations of material gradations as given by the power law x^p for different values of the parameter p . The types of volume fraction variations is rather limited; for instance, non-monotonic variations are excluded.

Mori-Tanaka (M-T) theory in regions where one of the constituents is dilute (metal-rich or ceramic-rich), and these approximations are linked between them by a “fuzzy inference” method. The temperature dependence of the material properties is not considered. Limiting bounds are used for the design variables to preserve toughness and stress constraints are also considered for keeping the circumferential stresses inside a region bounded by the temperature-independent critical stresses. No indication is given as to the meaning and origin of the critical stresses.

A different critical stress criterion from that used in [10] is proposed by Cho and Choi [12], where the objective function is to minimize the ratio of the peak effective stress and the yield stress. Both these quantities are functions of position in [12] but the temperature dependence is not considered in the problem of finding the optimum material gradation for maximizing the yield strength against residual stresses. The 2D analysis in [12] looks at the material starting at uniform temperature and being cooled down to a lower, uniform, temperature. This situation would not correspond to expected working conditions of the FGM structure, but rather to natural cooling encountered in the manufacturing process.

Genetic algorithms are used in [13] to optimize the volume fraction distribution in 2D FGMs. Temperature-dependent properties are not used and thus extra temperature constraints are enforced in order to avoid reaching designs in which the metal component yields. Convergence of the genetic algorithms is obtained after a few hundred generations.

Lipton [14] considers the optimal design of functionally graded shafts subject to purely mechanical loading (torsion). The homogenization theory for periodic microstructures is used and microstresses are included in the formulation. The coupled microstress is shown to account for less than 10% of the total stress. The analysis is limited, however, to FGMs in which one component is dominant (in terms of volume fraction), so the “fiber” and the “matrix” cannot switch roles. The validity range of homogenization methods for FGMs in the context of transient thermal loadings is analyzed in [1] with the use of a high-order method.

An interesting optimal volume fraction variation is found by Boussaa [15], who gives an example of a material optimization for a hollow cylinder FGM (under thermomechanical loadings) where the solution is monotonic but *discontinuous*. In [15] the thicknesses of the homogeneous metal and ceramic coatings and that of the FGM interlayer in the hollow cylinder are fixed, non-

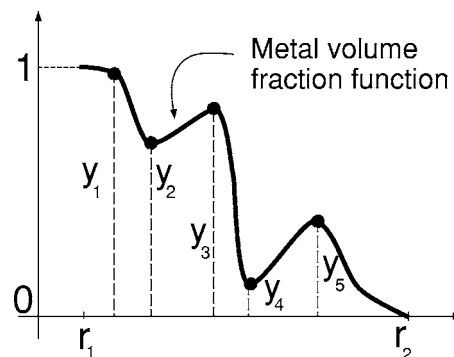


Fig. 2 Schematic representation for a possible continuous, non-monotonic volume fraction function along the radial direction of the hollow cylinder in Fig. 3, and sample design variables ($y_1, \dots, y_5$) that control its profile. The end points may also serve as design variables.

designable parameters. The temperature dependence of the material parameters is neglected and the cost function is selected to be the maximum circumferential stress across the ceramic coating and the graded interlayer. No constraints are used and the thermomechanical effective properties are computed via self-consistent schemes. The result in [15] shows the important role played by the cost function selection on the optimal volume fraction profile.

The minimum structural compliance problem for FGMs under thermomechanical loading is addressed in [16]. The strong influence of the choice of the material gradation on the response of the FGM under impact loading is noticed [17]. The effect of considering temperature-dependent material properties on the stress intensity factors and the energy release rate for an edge crack in an FGM is investigated in [3]. Important developments regarding isothermal fracture in FGMs are found in, for example, Kim and Paulino [18] and references therein.

Generally, the numerical methods used for the analysis of the mechanical and thermomechanical behavior of FGMs are the finite element method [8–10,19–21], the boundary element method [22,23] and meshfree methods [24–26].

A review article on thermal stresses in FGMs is presented in [27]. The focus in [27] is on transient thermoelasticity and the reduction of thermal stresses and the stress intensity factors in FGMs by changing the continuously graded composition of the material. The importance of considering temperature-dependent properties in thermal fracture of ceramics has been discussed in Jin and Batra [28], while [29] show the influence of using temperature-dependent material properties when modeling thermal post-buckling of FGM plates.

Two main issues do not seem to have been addressed in the literature on *optimal material design of FGMs*:

- allowing unrestricted volume fraction variations in the design space;
- considering temperature-dependent material properties and temperature-dependent objective function and constraints.

1.2 Our Contributions. The focus of the present paper is to analyze, using numerical models, the existence of optimal non-monotonic volume fraction functions in metal-ceramic FGMs that cannot be approximated by power-law type profiles. In our analysis we do not impose any predefined shape for the volume fraction function of one of the FGM’s components. For the continuous volume fraction representation we use a design space of shape-preserving cubic splines (see Sec. 2.4), which interpolate a set of control points (the design variables shown in Fig. 2). The design variables determine the “shape” of the optimal volume fraction function, and their number (density) controls the spline’s ability to approximate any continuous profile in the given interval. The op-

timization, however, is one of material, not of shape. While searching in a space of smooth functions, discontinuous solutions may be indicated by sudden changes in the shape-preserving spline.

We consider the weakly coupled thermoelasticity equations for heterogeneous materials (for the type of problems considered here, only the temperature influence on the mechanical response is needed). We use temperature-dependent material properties, and this leads to non-linear heat-transfer equations. In the specific examples we treat, the temperature-dependence of the conductivity parameter is, however, only weakly non-linear. To our knowledge, the present study seems to be the first one focusing on *optimal material design of FGMs with temperature-dependent material parameters*.

We formulate and solve the heterogeneous thermoelastic problem using the element-free Galerkin method (EFG) originally proposed by Belytschko et al. [30] for elasticity. We impose the Dirichlet boundary conditions with the “transformation matrix” method [31] for the thermal problem, and with the penalty method for the thermoelastic equations. The choice of a meshfree method for the analysis problem is here more a matter of convenience. The advantage compared to, for example, the finite element method (FEM), consists of a better resolution of the volume fraction function with the same number of discretization nodes. This is possible since the volume fraction variation is captured at the integration points of the background cells, which, in the EFG method, are independent from the discretization nodes. We use 5×5 quad integration cells that have nodes at their ends (see [32] for motivation of this selection). If one would use the FEM and tried to achieve the same resolution of the integration points, one would need many more nodes than used in the EFG solution. Consequently, the problem size would increase significantly. Our future plans are to combine material and shape optimization, and we have shown elsewhere [33] that meshfree methods have great advantages in shape optimization problems compared to the FEM, for example.

The EFG formulation for homogeneous linear thermoelasticity is given in Bobaru and Mukherjee [34]. In the present formulation, the material parameters depend on position (heterogeneous material) and temperature, and the heat-transfer part of the thermoelastic equations is non-linear. We compute the effective properties of the heterogeneous material with the simple rule of mixture (ROM) and with a combined Mori-Tanaka/fuzzy inference scheme [10]. Other micromechanics-based models can be used as well with our formulation, such as those developed by Reiter and Dvorak [35] that combine the Mori-Tanaka scheme with the self-consistent method.

We introduce new objective function and constraints based on temperature-dependent “effective” critical stresses for the FGM, computed from the yield stress for the metal phase, the critical strength for the ceramic component, and the rule of mixture. The strong dependence of the critical stresses on temperature dominates the optimal volume fraction solution, especially since the yield stress of the metal component drops very fast at high temperatures and the variation of the critical tensile strength of the ceramic is highly non-linear.

The paper is organized as follows: in the following section we describe the material optimization problem and give the EFG formulation and discretization for the thermoelastic problem for 2D heterogeneous solids with temperature-dependent material properties. To determine a sufficiently dense grid for an accurate solution, in Sec. 3 we compare the 2D EFG solution for the axisymmetric heat-transfer problem in a hollow FGM cylinder with the analytical solution for the case of a linearly varying (with the radial coordinate) temperature-independent conductivity. In the rest of the paper, temperature-dependent material properties are used. In Sec. 4 we define two material optimization problems for a functionally graded hollow cylinder subject to thermomechanical loads in plane strain conditions: a mass minimization under

stress constraints and a stress minimization problem. We present the numerical results for the two optimization problems in Sec. 4.2. In these examples, a radially dependent volume fraction function is designed since we try to see if imposed forms on the volume fraction functions (like the power-law profiles) used in many previous studies are too restrictive or not. A two-dimensional, (r, θ) surface type variation of the volume fraction can be easily used if desired without any modifications to the solution method proposed here. We present conclusions in Sec. 5.

2 Problem Description

2.1 Material Optimization Problem Setup. In two-phase composite FGMs, the material composition is defined by the volume fraction function taking values between 0 and 1. We define the optimal volume fraction function as the solution to the functional minimization problem

$$\text{minimize } \mathcal{F}[y(\mathbf{x})]$$

$$\text{subject to } \mathcal{G}[y(\mathbf{x})] \geq 0$$

$$0 \leq y(\mathbf{x}) \leq 1 \quad (1)$$

where $y=y(\mathbf{x})$ is the volume fraction function, while the objective function $\mathcal{F}$ and the constraint $\mathcal{G}$ are, for example, measures of mass, scalar measures of stresses, etc. and thus can be computed by solving the heterogeneous thermoelastic problem described in the next sections. To be able to assess the validity of imposing specific shapes for the volume fraction function, we will assume that this function depends on one spatial coordinate: the radial direction for the optimal material design of a functionally graded hollow cylinder. Methods of non-linear programming [36] can be used for approximating the solution of the optimization problem above if we reduce it to an optimization problem in a finite-dimensional space. We do this by using interpolating shape-preserving cubic splines (Akima splines) whose shapes are defined by a set of points (material design variables) like those shown in Fig. 2. By increasing the number of design variables, we approach the solution of the infinite-dimensional optimization problem (1).

2.2 Temperature-Dependent Thermoelasticity. We consider a domain Ω that is occupied by the FGM under thermomechanical loading. The non-linear heat-transfer equations and the equations of the thermoelastic equilibrium can be written as in Eqs. (2) and (3) below

$$\begin{cases} \nabla \cdot [\kappa(\mathbf{x}, T) \nabla T] + Q(\mathbf{x}, T) = 0 & \text{in } \Omega \\ T = T_0 & \text{on } \Gamma_0^T \\ \kappa(\mathbf{x}, T) \nabla T \cdot \mathbf{n} = \bar{q} & \text{on } \Gamma_1^T \\ \kappa(\mathbf{x}, T) \nabla T \cdot \mathbf{n} + h(\mathbf{x}, T)(T - T_\infty) = 0 & \text{on } \Gamma_2^T \end{cases} \quad (2)$$

where T is the unknown temperature field over the cross-sectional domain Ω , T_0 is the given temperature on the Γ_0^T , $\bar{q}$ is the imposed heat-flux density on Γ_1^T , while T_∞ is the given ambient temperature over the convective boundary Γ_2^T . In addition, $\kappa(\mathbf{x}, T)$ is the heat conduction coefficient, $h(\mathbf{x}, T)$ the convective heat transfer coefficient, and Q the internal heat source. We assume here that the material is thermally isotropic. The exterior normal to a boundary is denoted by $\mathbf{n}$. The dot $(\cdot)$ in the above equations denotes the scalar (or dot) product between tensors of rank 1. In the examples shown in Sec. 4, only temperature and flux boundary conditions are used and the internal heat source is zero. Once the temperature field is solved for from Eq. (2), the thermoelasticity equations for heterogeneous materials give the displacements, stresses, and strains as follows

$$\begin{cases} \nabla \cdot \sigma(\mathbf{x}, T) + \mathbf{b}(\mathbf{x}, T) = \mathbf{0} & \text{in } \Omega \\ \sigma(\mathbf{x}, T)\mathbf{n} = \bar{\mathbf{t}}(\mathbf{x}, T) & \text{on } \Gamma_t \\ \mathbf{u}(\mathbf{x}, T) = \bar{\mathbf{u}}(\mathbf{x}, T) & \text{on } \Gamma_u \end{cases} \quad (3)$$

where

$$\sigma = \mathbf{C}(\mathbf{x}, T) : [\varepsilon(\mathbf{x}, T) - \alpha(\mathbf{x}, T)T(\mathbf{x})\mathbf{I}] \quad (4)$$

with σ , ε , and $\mathbf{u}$ representing the stress tensor, strain tensor, and displacement field, respectively; $\mathbf{C}$ is the heterogeneous elasticity fourth-order tensor, $\mathbf{I}$ is the second-order identity tensor, and α is the coefficient of thermal expansion that depends on the position and temperature. Body forces are indicated by $\mathbf{b}$, while given tractions $\bar{\mathbf{t}}$ are applied over a portion Γ_t of the boundary of Ω . The imposed displacements over Γ_u are $\bar{\mathbf{u}}$.

The heat-transfer problem in Eq. (2) is non-linear since the conductivity depends on the temperature. For the examples we consider, the conductivity dependence on temperature is weakly non-linear (see data in Tables 4 and 5 in Appendix A). We linearize the problem with a “fixed-point” type iterative scheme. Assume that at iteration k we have solved for the temperature field T_k over the domain Ω . The solution for the temperature field T_{k+1} at the next iteration is then given by the solution to the following linear boundary value problem

$$\begin{cases} \nabla \cdot (\kappa(\mathbf{x}, T_k) \nabla T_{k+1}) + Q(\mathbf{x}, T_k) = 0 & \text{in } \Omega \\ T_{k+1} = T_0 & \text{on } \Gamma_0^T \\ \kappa(\mathbf{x}, T_k) \nabla T_{k+1} \cdot \mathbf{n} = \bar{q} & \text{on } \Gamma_1^T \\ \kappa(\mathbf{x}, T_k) \nabla T_{k+1} \cdot \mathbf{n} = h(\mathbf{x}, T_k)(T_{k+1} - T_\infty) & \text{on } \Gamma_2^T \end{cases} \quad (5)$$

The iterations continue until the relative error between the solution vector (which contains the approximate solution at nodes along the radial direction) at the current and previous iteration $\|\mathbf{T}_{k+1} - \mathbf{T}_k\| / \|\mathbf{T}_k\| \leq \delta$, with δ a preset tolerance. We use norm-2 here ($\|\mathbf{v}\| = \sqrt{v_1^2 + \dots + v_n^2}$). The solution converges in only a few iterations for the example of the hollow cylinder shown in Sec. 4 and for a tolerance of $\delta = 10^{-6}$. The starting guess for the iterations is defined as a linear interpolation, along the radial direction of the hollow cylinder, of the imposed boundary temperatures at the inner and outer surfaces of the cylinder.

2.3 Meshfree Solution of the Thermoelastic Problem for the Temperature-Dependent Heterogeneous Materials. The meshfree solution for the non-linear thermal problem amounts to solving a set of linear systems that are formed by discretizing the weak forms corresponding to the equations in (5)

$$\begin{aligned} & \int_{\Omega} \kappa(\mathbf{x}, T_k) \nabla T_{k+1} \cdot \nabla \eta \, d\Omega - \int_{\Gamma_1^T} \bar{q} \eta \, d\Gamma + \int_{\Gamma_2^T} h(\mathbf{x}, T_k) T_{k+1} \eta \, d\Gamma \\ & - \int_{\Gamma_2^T} h(\mathbf{x}, T_k) T_\infty \eta \, d\Gamma = 0 \quad \text{for any } \eta \in V \end{aligned} \quad (6)$$

where $\eta \in V = \{\eta \in H^1(\Omega), \eta = 0 \text{ on } \Gamma_0^T \text{—in the sense of trace}\}$ are the test functions.

The element-free Galerkin (EFG) discretization of Eq. (6) leads to the following linear system of equations

$$\mathbf{M}(T_k) \mathbf{T}_{k+1} = \mathbf{f}(T_k) \quad (7)$$

where the global matrix $\mathbf{M}(T_k)$ and the right-hand side vector $\mathbf{f}(T_k)$ are defined by

$$\begin{aligned} \mathbf{M}(T_k) &= \int_{\Omega} \kappa(\mathbf{x}, T_k) \Lambda^{-1} \mathbf{B}^T \mathbf{B} \Lambda^{-T} d\Omega + \int_{\Gamma_2^T} h(\mathbf{x}, T_k) (\Lambda^{-1} \phi) \\ &\quad \otimes (\Lambda^{-1} \phi) d\Gamma \end{aligned} \quad (8)$$

$$\mathbf{f}(T_k) = \int_{\Omega} Q \Lambda^{-1} \phi \, d\Omega + \int_{\Gamma_1^T} \bar{q} \Lambda^{-1} \phi \, d\Gamma + \int_{\Gamma_2^T} h(\mathbf{x}, T_k) T_\infty \Lambda^{-1} \phi \, d\Gamma \quad (9)$$

The symbol $\otimes$ above stands for the exterior (tensor) product. The matrix $\mathbf{B}$ above is used to approximate the gradient in the N -dimensional discretization space as $\nabla \theta^{(N)} = \mathbf{B} \hat{\boldsymbol{\theta}}$ and is given by

$$\mathbf{B} = \begin{bmatrix} \frac{\partial \phi_1}{\partial \mathbf{x}_1} & \dots & \frac{\partial \phi_N}{\partial \mathbf{x}_1} \\ \frac{\partial \phi_1}{\partial \mathbf{x}_2} & \dots & \frac{\partial \phi_N}{\partial \mathbf{x}_2} \end{bmatrix} \quad \text{with } \phi = \begin{bmatrix} \phi_1 \\ \vdots \\ \phi_N \end{bmatrix}, \quad \boldsymbol{\theta} = \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_N \end{bmatrix} \quad (10)$$

where ϕ_i are the moving-least-squares approximation functions. The matrix Λ in Eq. (8) is accounting for imposing the geometric (or essential) boundary conditions via a transformation method [31,37]. The transformation of the shape functions is necessary in the EFG method since the shape functions do not satisfy the Kronecker delta property. Other methods for imposing the geometric boundary conditions are available and a review of these methods is presented in [38]. The temperature values the nodes, i.e., $\theta_j = \theta(x_j)$, $j=1, \dots, N$, are expressed in terms of some “fictitious” temperatures as

$$\theta_j = \theta(x_j) = \sum_i \phi_i(x_j) \hat{\theta}_i = \sum_i \Lambda_{ij} \hat{\theta}_i = (\Lambda^T \hat{\boldsymbol{\theta}})_j \quad (11)$$

$$\hat{\boldsymbol{\theta}} = \Lambda^{-T} \boldsymbol{\theta}, \quad \hat{\boldsymbol{\eta}} = \Lambda^{-T} \boldsymbol{\eta} \quad (12)$$

where $\hat{\theta}_i$ are the “fictitious” nodal temperature values, and $\Lambda_{ij} = \phi_i(x_j)$ is the transformation matrix.

The weak form for the heterogeneous thermoelastic Eqs. (3) and (4) is

$$\begin{aligned} & \int_{\Omega} \mathbf{C}(\mathbf{x}, T) : \varepsilon(\mathbf{u}) \cdot \varepsilon(\boldsymbol{\xi}) \, d\Omega + \delta \int_{\Gamma_u} \mathbf{u} \cdot \boldsymbol{\xi} \, d\Gamma = \int_{\Omega} \mathbf{b}(\mathbf{x}, T) \cdot \boldsymbol{\xi} \, d\Omega \\ & + \int_{\Gamma_t} \bar{\mathbf{t}} \cdot \boldsymbol{\xi} \, d\Gamma - \delta \int_{\Gamma_u} \bar{\mathbf{u}} \cdot \boldsymbol{\xi} \, d\Gamma + \int_{\Omega} \beta(\mathbf{x}, T) T \operatorname{div} \boldsymbol{\xi} \, d\Omega \\ & - \int_{\Gamma_u} \beta(\mathbf{x}, T) T \mathbf{n} \cdot \boldsymbol{\xi} \, d\Gamma \quad \text{for any } \boldsymbol{\xi} \in [H^1(\Omega)]^2, \end{aligned} \quad (13)$$

where $\varepsilon(\mathbf{u})$ is the symmetric part of $\nabla \mathbf{u}$. In the examples with a hollow cylinder that follow, we will use plane strain conditions for which $\beta(\mathbf{x}, T) = \alpha(\mathbf{x}, T)E(\mathbf{x}, T)/[1 - 2\nu(\mathbf{x}, T)]$ and

$$\mathbf{C}(\mathbf{x}, T) = \frac{E(\mathbf{x}, T)}{(1 + \nu)(1 - 2\nu)} \begin{bmatrix} 1 - \nu & \nu & 0 \\ \nu & 1 - \nu & 0 \\ 0 & 0 & \frac{1 - 2\nu}{2} \end{bmatrix} \quad (\text{plane strain}).$$

Notice that in the computer implementation, we use the customary engineering vector notation for the strain tensor, namely

$$\boldsymbol{\varepsilon} = [\varepsilon_{11}, \varepsilon_{22}, 2\varepsilon_{12} = \gamma_{12}]$$

which leads to the above matrix form of $\mathbf{C}(\mathbf{x}, T)$.

We use a penalty parameter δ between $10^7 \times E$ to $10^9 \times E$ to impose the displacement boundary conditions, and this avoids ill-conditioning. For the examples considered here, the transformation method with inversion of the full matrix described above for the thermal problem is about 50% slower than the penalty method in the case of the thermoelastic weak form. One could use the transformation method with the inversion of the reduced matrix as explained in [31], but the simplicity of the implementation is then lost. This weak form is discretized using the EFG approximation functions. In [34], this is done for the homogeneous material case.

Here we have to include both the spatial and the temperature dependence of material parameters. Computing the material properties for the heterogeneous case is described in the next section.

2.4 Computing Effective Material Properties for the Composite. For the numerical approximation of domain and boundary integrals in (6) and (13) we use Gaussian integration. For domain integrals we employ background quad cells with 5×5 Gauss points. For simplicity, the integration cells are defined by the discretization nodes. All material parameters for the metal-ceramic composite (ZrO_2 and Ti-6Al-4V) have to be evaluated at the integration points and the data in tables in Appendix A is used as follows:

- At a radial location r of the hollow cylinder cross section, where the volume fraction of metal is $y(r)$, we evaluate the “composite” material parameters via the ROM scheme or the combined Mori-Tanaka/fuzzy inference method. For the latter, the Mori-Tanaka method is used over metal- or ceramic-rich regions, while the fuzzy inference method (see Appendix B) connects these regions over the areas where the inclusions and the matrix cannot be easily separated. Other schemes for determining effective properties [35,39] can easily be used with the present formulation.
- Since the micromechanics models require data at temperatures other than those in Appendix A, we use the Akima shape-preserving spline to interpolate the data and to evaluate material properties at any given temperature.
- The continuous volume fraction function profile that defines the FGM is approximated by an Akima spline controlled by a set of design variables (see Fig. 2). While the oscillations of cubic spline interpolation are reduced in the Akima spline interpolation, some portions of the Akima spline may still exit the physical domain allowed for the volume fraction function: the interval $[0,1]$. In such cases, instead of using the Akima spline $s(r)$ we use the *modified* continuous function $\hat{s}(r)$ defined below

$$\begin{cases} \hat{s}(r) = s(r) \\ \text{where } s(r) < 0, \quad \text{set } \hat{s}(r) = 0 \\ \text{where } s(r) > 1, \quad \text{set } \hat{s}(r) = 1 \end{cases} \quad (14)$$

If the control points (design variables) at the ends of the spline are collinear, then the end-condition quadratic of the Akima spline becomes linear [40] and no parts of the spline fall below 0 or above 1. The design variables are independent of the 2D grid nodes used for the solution of the thermoelastic problem. We could also select independent design variables in the θ direction of the hollow cylinder, thus defining a surface for controlling the material variation in the FGM.

An important advantage of the meshfree solution when compared to a FEM solution is that the Gauss integration points are *independent* from the approximation nodes. As the values of the volume fraction function are evaluated at these integration points, one can accurately capture the variations in this function without requiring a large number of discretization nodes. If one uses the FEM and tries to achieve the same resolution of the integration points as in the 5×5 integration points per integration cell used in this work, one needs a much larger number of nodes than the EFG solution. Consequently, the size of the equivalent FEM problem increases significantly compared to the EFG solution. For a study of the efficiency of various integration schemes in EFG see, for example, [32].

3 A Convergence Test

The examples we analyze in Sec. 4 are for the 2D cross section of a hollow cylinder subject to a high temperature gradient and boundary conditions as shown in Fig. 3. We perform a conver-

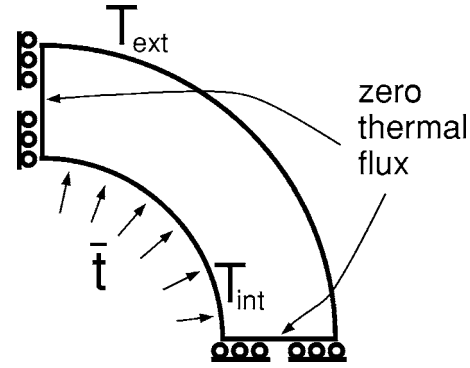


Fig. 3 Boundary conditions for the hollow cylinder under thermomechanical loading: temperature values imposed on the inner (T_{int}) and outer (T_{ext}) surfaces, inner pressure ($\bar{p}$), and symmetry conditions for the thermal flux and displacements

gence test in order to determine a sufficiently dense grid for an accurate solution of the thermoelastic problem. The analytical solution for the linear heat transfer with a linearly varying thermal conductivity can be found. We compare this analytical solution with the numerical temperature field distributions of a functionally graded hollow cylinder for which the material properties vary along the radial direction. We consider 100% ceramic material at the inner surface and pure metal at the outer surface of the hollow cylinder. Assume that the ceramic side is at temperature T_1 (high) and the metal surface at temperature T_2 (low). The steady state equation of the one-dimensional heat transfer, with *temperature-independent* conductivity, can be expressed as the boundary-value problem

$$\frac{1}{r} \frac{d}{dr} \left[\kappa(r) r \frac{dT}{dr} \right] = 0$$

with boundary conditions

$$T(r)|_{r=r_1} = T_1 \quad T(r)|_{r=r_2} = T_2$$

Here, $\kappa(r)$ is the thermal conductivity of the given FGM. The general solution to this equation is given by

$$T(r) = A \int_{r_1}^r \frac{1}{r\kappa(r)} dr + B \quad (15)$$

The coefficients A and B are obtained by applying the boundary conditions, as follows

$$A = (T_1 - T_2) \left[\int_{r_1}^{r_2} \frac{1}{r\kappa(r)} dr \right]^{-1} \quad B = T_2$$

Assuming that the thermal conductivity varies linearly with respect to the position r , $\kappa(r) = a + br$, where a and b are dimensionally corresponding constants, the temperature field solution becomes

$$T(r) = T_2 + (T_1 - T_2) \left(\ln \frac{r}{a + br} - \ln \frac{r_1}{a + br_1} \right) \left(\ln \frac{r_2}{a + br_2} - \ln \frac{r_1}{a + br_1} \right)^{-1} \quad (16)$$

We check this analytical solution against the 2D solution using the EFG method described above. The 2D numerical solution developed here can be used for the case of material variation in the angular direction as well, and in which the sought-after volume fraction is a surface instead of a curve. We consider, in consistent units, $r_1 = 0.7$, $r_2 = 1$, $T_1 = 800$, $T_2 = 0$, $a = -8.733$, and $b = 15.333$. The values for a and b are so chosen as to match the conductivity of ceramic ($\kappa = 2$) at $r = r_1$ and that of metal at $r = r_2$ ($\kappa = 6$). For

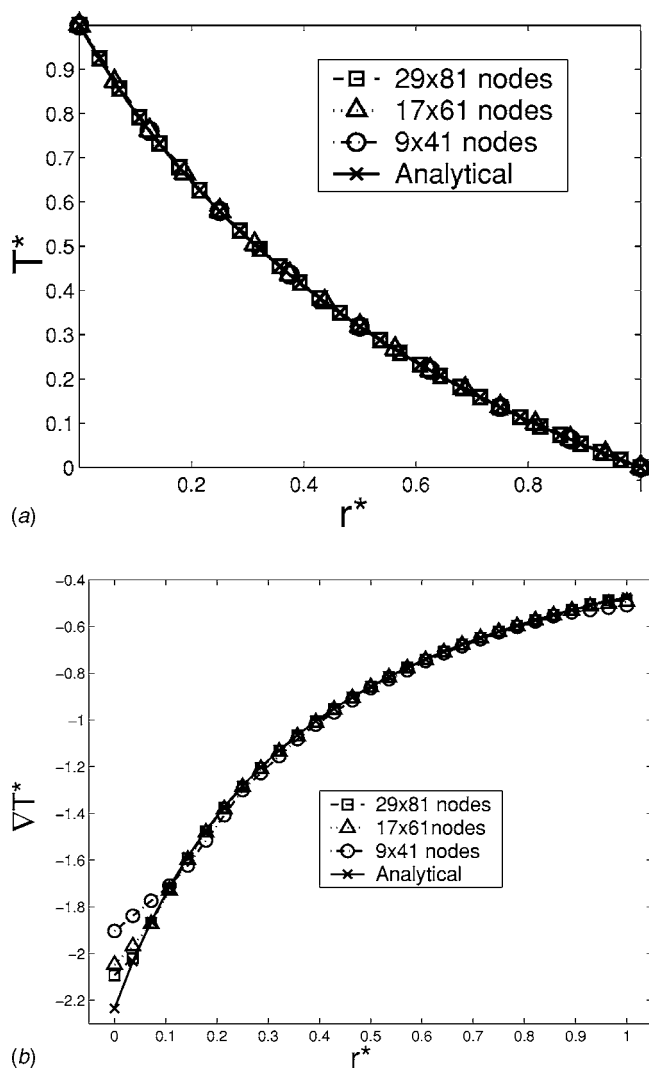


Fig. 4 Analytical and numerical results for the temperature (a) and temperature gradient (b) along the radial direction of a hollow FGM cylinder under temperature-imposed boundary conditions. Dimensionless quantities are as in Eq. (17).

this convergence test we ignore the temperature-dependence of the thermal conductivity. We use grids of 9×41 , 17×61 , and 29×81 nodes in the radial and angular directions, respectively. The quads naturally formed by the nodes are selected, for convenience, as background integration cells in the meshfree solution. We employ 5×5 Gaussian integration in each cell. A linear basis is used for computing the moving-least-squares approximation [41]. The numerical results we plot in Fig. 4 are for points along the 45 deg direction and we use dimensionless quantities as follows:

$$r^* = \frac{r - r_1}{r_2 - r_1}, \quad T^* = \frac{T}{T_1 - T_2}, \quad \nabla T^* = \nabla T \frac{r_2 - r_1}{T_1 - T_2} \quad (17)$$

In Fig. 4 we give the temperature and the temperature gradient for the three grids together with the analytical solutions. The small error present at the end point in the gradient values are due to post-processing: for end points we use the forward or backward finite difference (first-order accuracy) to compute the gradient values, whereas the gradient at the interior is computed using the central difference method (second-order accuracy). The relative error in norm 2 (see Sec. 2.2) for the temperature gradient drops

from 0.0674, to 0.0335, to 0.0239 for the three different grids, respectively. We select the 29×81 discretization for all the subsequent numerical tests.

4 Material Optimization Example Problems

4.1 Computing Effective Critical Stresses. We define critical stress measures to employ as the material optimization objective function and constraints. Since no experimental values are available for various metal-ceramic functionally graded mixtures at various temperatures, we evaluate the composite “critical” tensile and compressive stresses in the FGM using the temperature-dependent yield stress of the metal component and critical strength of the ceramic, together with the simple ROM. If experimental values for critical stresses in FGMs become available for arbitrary volume fractions and the range of temperatures that we cover in this study, the present formulation can easily integrate them. It is worthwhile to note that, in the purely mechanical problem of optimal design of graded composites with predetermined matrix and fiber components [14] that cannot change roles, where a more accurate homogenization technique is used to account for the stresses due to the microstructure, the microstresses account for no more than 10% of the total stress.

Let $\sigma_m(T)$ and $\sigma_c(T)$ be, respectively, the critical tensile (or compressive) yield stress for the homogeneous metal and the tensile (or compressive) strength for the homogeneous ceramic material at a given temperature T (see Tables 4 and 5 in Appendix A, and also note the strong dependence on temperature of the yield and critical stresses). The estimated effective critical tensile (or compressive) stress in the FGM at a location where the volume (area) fraction of the metal is $y(r)$ will then be given by

$$\sigma(r, T) = \sigma_m(T)y(r) + \sigma_c(T)[1 - y(r)] \quad (18)$$

This critical tensile (or compressive) stress is used to define an objective function and/or constraint in the next section.

4.2 Optimization of Volume Fraction Function. We define two non-linear optimization problems: find the best continuous metal volume fraction function in the space of Akima splines controlled by the set of p equally spaced points in the radial direction $(r_1, y_1), \dots, (r_p, y_p)$, which

- (1) minimizes mass under critical stress violation constraints
- (2) minimizes a critical stress violation measure

4.2.1 Problem 1: Mass Minimization Under Stress Violation Constraints. The non-linear optimization problem in this case reads

$$\text{minimize } \mathcal{F}(\mathbf{y}) = m(\mathbf{y})$$

$$\text{subject to } \mathcal{G}(\mathbf{y}) = 1 - \max[v(r, T, \mathbf{y})] \geq 0$$

$$0 \leq y \leq 1 \quad (19)$$

where the design variable (DV) vector $\mathbf{y} = [y_1, \dots, y_p]$ represents the heights of the control points for the volume fraction shape-preserving spline (see Fig. 2) and the objective function is the mass m of the 2D cross section calculated using the variable local density, which is determined by the volume fraction function. The stress violation constraint $\mathcal{G}(\mathbf{y}) \geq 0$ is based on a local stress measure. The function v below gives the violation of the critical effective stress in the composite computed along a radial direction

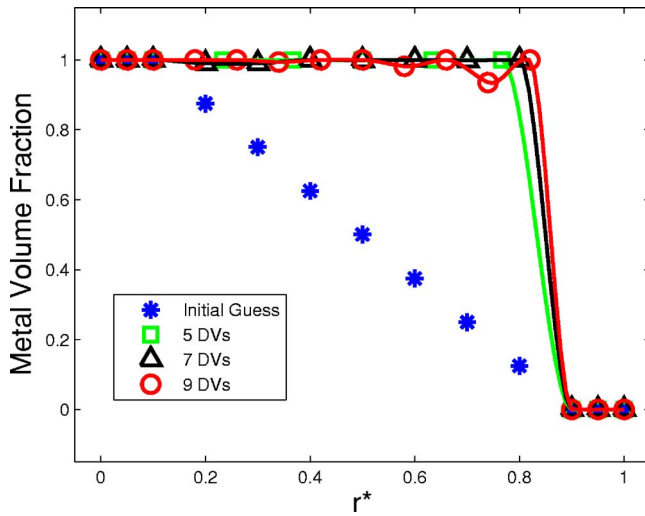


Fig. 5 The initial guess (*) and optimal designs for the metal volume fraction function with five ($\square$), seven ($\triangle$), and nine ($\circ$) design variables selected between the pure metal and ceramic coatings

$$\left\{ \begin{array}{l} \text{if } \sigma_{\theta}(r, T, \mathbf{y}) > 0 \text{ then } v(r, T, \mathbf{y}) = \frac{\sigma_{\theta}(r, T, \mathbf{y})}{\gamma \sigma_{ic}(r, T, \mathbf{y})} \\ \text{otherwise} \\ v(r, T, \mathbf{y}) = \frac{|\sigma_{\theta}(r, T, \mathbf{y})|}{\gamma |\sigma_{cc}(r, T, \mathbf{y})|} \end{array} \right. \quad (20)$$

In the above equation, $\sigma_{ic}(r, T, \mathbf{y})$ and $\sigma_{cc}(r, T, \mathbf{y})$ are the critical effective stresses in tension and compression, respectively, computed via Eq. (18). We emphasize that these critical effective stresses are very strongly dependent on temperatures, as seen from the strongly nonlinear dependence of the metal and ceramic critical stresses given in Tables 4 and 5. The factor γ is a safety factor selected to be 0.8 in what follows. We use the safety factor since the SQP nonlinear optimization algorithm [36] we employ to solve does not require iterates to be feasible. Consequently, “optimal” solutions with small violation of the constraints are possible, but this is undesirable from a design point of view. The safety factor value we use insures that minimizers that violate the constraint by less than 10^{-1} do not result in stresses larger than the critical composite stress at the corresponding temperature and position. In this study we use the dnconf routine from the IMSL library, which finds sensitivities internally by finite difference approximations.

4.2.2 Numerical Results for Problem 1. The boundary conditions are (see Fig. 3): on the boundary Γ_0^T ($r=0.7$ and $r=1$) $T_{\text{ext}} - T_{\text{int}} = 800^\circ\text{C}$, with $T_{\text{int}}=0$, while zero-flux conditions are imposed on the Γ_1^T boundary ($\theta=0$ and $\theta=\pi/2$). An inner pressure $\|\bar{\mathbf{f}}\|=100$ MPa is also imposed.

We select as initial volume fraction as in Fig. 5. A pure metal inner coating for $r \in [0.7, 0.73]$ and a pure ceramic outer coating over $r \in [0.97, 1.0]$. We start with a linear variation of the metal volume fraction between these coatings. This starting guess falls inside the optimization feasible region.

The Akima shape-preserving spline interpolates the design variables $y_1, \dots, y_p$ equally spaced in the radial direction between 0.73 and 0.97, and the fixed points in the imposed coatings. When values outside the interval $[0, 1]$ are produced by the Akima spline, we modify the spline as in (14). Due to this procedure the interpolation function, while in C^0 , may become non-differentiable at some locations in the pure metal or pure ceramic regions. When-

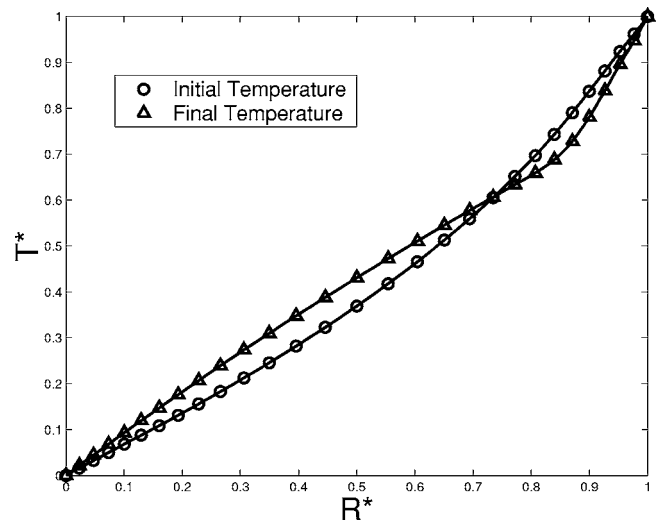


Fig. 6 The temperature along the radial direction of the hollow cylinder at the initial design ($\circ$) and final metal volume fraction ($\triangle$) in the case of seven design variables chosen between the fixed coatings

ever the optimizer stops because of this, we restart computations after a small perturbation of the last computed value of the design variables. Only a few restarts are needed.

With five, seven, and nine design variables placed equally spaced from $r=0.73$ to $r=0.97$, we obtain the optimal variation along the radial direction of the hollow cylinder for the metal volume fraction as in Fig. 5. We obtain these results after a few small perturbations (of relative magnitude less than 1%). We note that the metal component has a smaller density than the ceramic material, so that the metal-rich optimal design is expected.

The temperature profile for the initial and final volume fractions are represented in Fig. 6 for the case of seven design variables. The objective function (mass) history for the case of seven design variables is shown in Fig. 7 and the history of the stress constraint is given in Fig. 8. There is a slight violation of the constraint (see Fig. 8), but the use of the safety factor insures that the result is still inside the elastic design, therefore, the solution is valid. The stress profiles and critical stresses along a radial direction at two intermediate iterations (see Fig. 9) show the strong dependence of

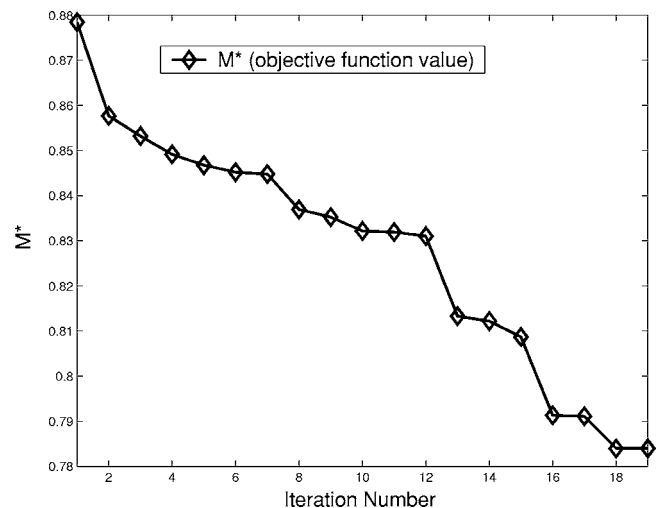


Fig. 7 The history of the objective for the case of seven design variables. We use dimensionless mass $M^* = M/M_c$, where M_c is the mass of a purely ceramic hollow cylinder.

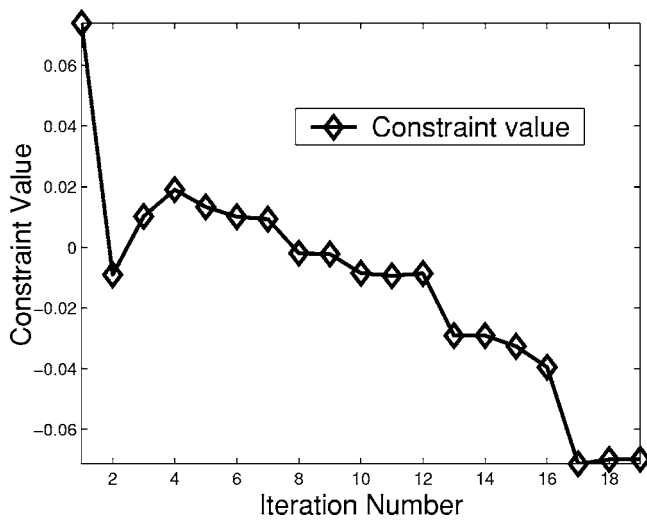


Fig. 8 History of constraint value for the case of seven design variables. The constraint in Eqs. (19) and (20) is slightly violated, but the use of the safety factor gives stresses that are nowhere larger than the critical stresses.

the constraints on the temperature-dependent material properties. The dimensionless radius is defined in Eq. (17) while the dimensionless stress is given by

$$\sigma^* = \frac{\sigma}{E_{c0}\alpha_{c0}\Delta T} \quad (21)$$

where E_{c0} and α_{c0} are Young's modulus and the coefficient of thermal expansion for the ceramic material at room temperature, and $\Delta T = T_{\text{ext}} - T_{\text{int}}$.

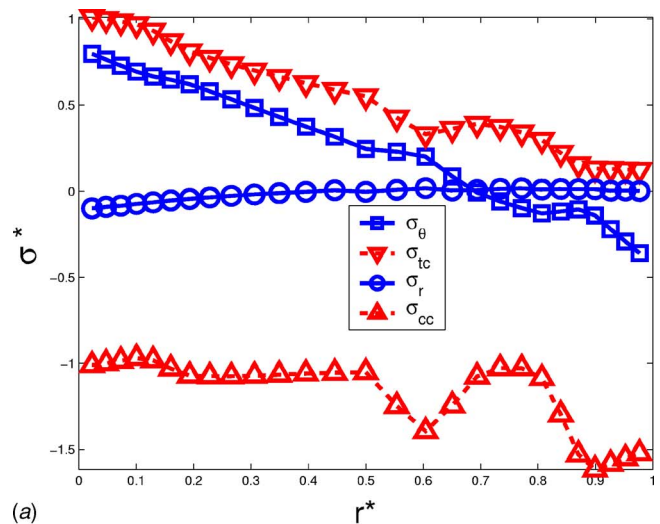
The sharp (thin) FGM layer obtained in front of the imposed ceramic coating may indicate that the optimal solution could change if we removed the fixed metal and ceramic coatings imposed above. We therefore eliminate the inner and outer coatings and assign equally spaced design variables from $r=0.7$ to $r=1$. The initial guess for 6, 9, and 13 design variables is selected as before: two constant values at the ends connected by a linear variation. This time, however, all control points are design variables allowed to vary in the $[0,1]$ interval in search for the optimal configuration. The location of the 13 design variables at the initial guess is shown in Fig. 10. Note that the coordinates of the seven inner design variables for the present case with 13 design variables are the same as those in the case of seven design variables with enforced metal and ceramic coatings. The convergence with the constraint in Eq. (20) becomes more difficult as the design space increases. We propose a different measure for the critical stress violation than the one defined by (19) and (20), equivalent to it but giving larger absolute values for stresses that violate the temperature-dependent critical stresses. The new constraint provides better scaling for the SQP optimizer and convergence is improved. When no coatings are enforced, instead of the constraint defined by (19), we use

$$\mathcal{G}(\mathbf{y}) = \min[\hat{v}(r, T, \mathbf{y})] \geq 0$$

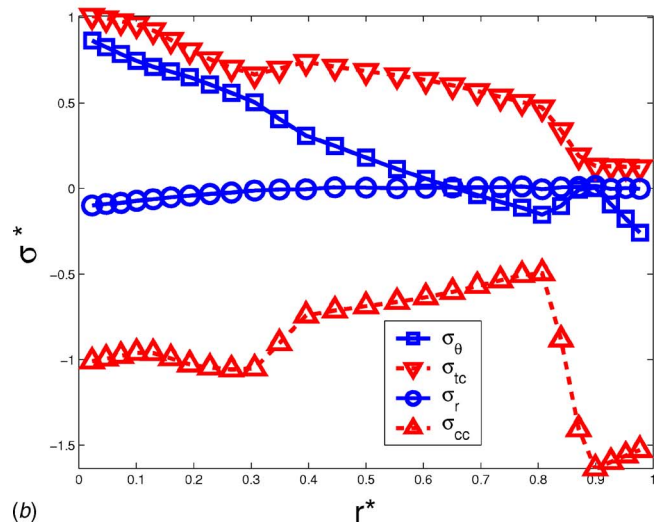
$$0 \leq \mathbf{y} \leq 1 \quad (22)$$

where $\hat{v}$ gives the violation of the critical effective stress in the composite computed along a radial direction as follows

$$\begin{cases} \text{if } \sigma_{\theta}(r, T, \mathbf{y}) > 0 & \text{then } \hat{v}(r, T, \mathbf{y}) = \gamma \sigma_{tc}(r, T, \mathbf{y}) - \sigma_{\theta}(r, T, \mathbf{y}) \\ \text{otherwise} & \hat{v}(r, T, \mathbf{y}) = \sigma_{\theta}(r, T, \mathbf{y}) - \gamma \sigma_{cc}(r, T, \mathbf{y}). \end{cases} \quad (23)$$



(a)



(b)

Fig. 9 The stress profiles along the radial direction of the hollow cylinder for two intermediate iterations for the case of seven design variables in the problem defined by Eqs. (19) and (20). The critical stresses are dependent on the temperature and, implicitly, on the design variables. Shown are the tangential (σ_{θ} , $\square$) and radial (σ_r , $\circ$) stresses, and the critical tensile (σ_{tc} , ∇) and compressive (σ_{cc} , $\triangle$) stresses.

The optimal volume fraction for the mass minimization with stress constraints when no coatings are imposed is shown in Fig. 10 for a value of the safety factor γ equal to 0.9. Notice that we recover the result in the previous analysis, and this demonstrates that the sharp (thin) FGM layer is not a mere artifact of the previously imposed metal and ceramic coatings. The ceramic coating protects the FGM from high temperatures that induce stress states close to the temperature-dependent critical values. The nine design variables case stops at a profile that has a large constraint violation caused by the absence of a sufficiently thick ceramic layer.

4.2.3 Problem 2: Stress Violation Minimization. In this case, the unconstrained non-linear optimization problem we propose is to minimize the largest ratio between the actual circumferential stresses and the critical stresses, given by Eq. (18). The non-linear optimization problem reads

$$\text{minimize } \mathcal{F}(\mathbf{y}) = \max[w(r, T, \mathbf{y})]$$

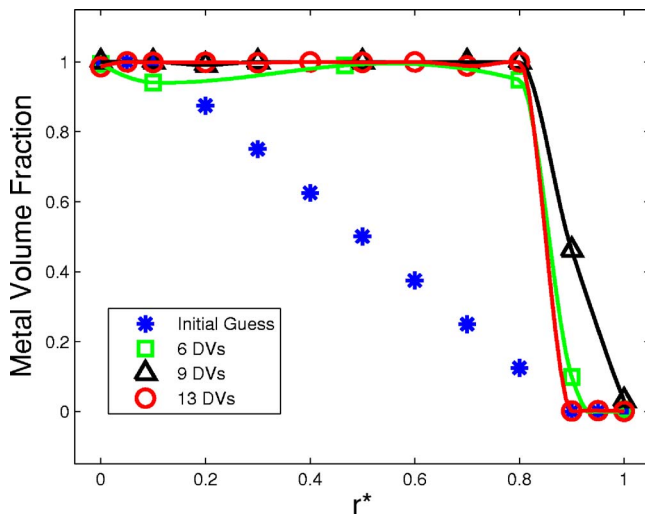


Fig. 10 Volume fraction profile for the mass minimization with no enforced coatings. The initial guess (*) and optimal designs for the metal volume fraction function with six (□), nine (△), and 13 (○) design variables.

$$\text{subject to } 0 \leq y \leq 1 \quad (24)$$

where the design variable vector $\mathbf{y}$ is defined as before, and the objective function is the largest value of the ratio between the local actual and critical stresses. The function w , when larger than 1, gives the violation of the critical effective stress in the composite computed at some point along a radial direction

$$\left\{ \begin{array}{ll} \text{if } \sigma_{\theta}(r, T, \mathbf{y}) > 0 \text{ then } w(r, T, \mathbf{y}) = \frac{\sigma_{\theta}(r, T, \mathbf{y})}{\sigma_{ic}(r, T, \mathbf{y})} \\ \text{otherwise} & w(r, T, \mathbf{y}) = \frac{|\sigma_{\theta}(r, T, \mathbf{y})|}{|\sigma_{cc}(r, T, \mathbf{y})|} \end{array} \right. \quad (25)$$

Any value of w larger than 1 represents a violation of the critical stress, at the corresponding location, by the actual tangential stress. A safe design in this case is one in which the objective function in Eq. (24) reaches sub-unitary values. We emphasize that the critical effective stresses $\sigma_{ic}(r, T, \mathbf{y})$ and $\sigma_{cc}(r, T, \mathbf{y})$, computed with the help of Eq. (4) and the data in Tables 4 and 5, are very strongly dependent on temperature. For this unconstrained optimization problem we use the dbconf routine from the IMSL library which implements a quasi-Newton method and an active set strategy to solve minimization problems subject to simple bounds on the variables. Sensitivities are computed internally by finite difference approximations.

4.2.4 Numerical Results for Problem 2. We perform several tests. The first set of tests is with no inner pressure, while for the second set we use the same inner pressure as in the case of Problem 1 ($\|\bar{\mathbf{f}}\| = 100$ MPa).

We start with an initial design and boundary conditions that produce stresses violating the critical ones by a large margin. We select, therefore, the boundary conditions as follows: inner temperature $T_{\text{int}} = 100^\circ\text{C}$, outer temperature $T_{\text{ext}} = 900^\circ\text{C}$, zero inner pressure. The initial guess for the metal volume fraction variation is constant equal to 0.9 (see Fig. 11). This corresponds to a homogenized metal-rich (90% metal by volume) composite material.

The solution for the metal volume fraction variation is given in Fig. 11 for five (test A), six (test B), and nine (test C) design variables. In Table 1 we give the values of the objective function (24) and (25) at the initial and final designs as well as the number of iterations performed during the optimization process. The temperature profiles at the initial and final volume fraction designs for

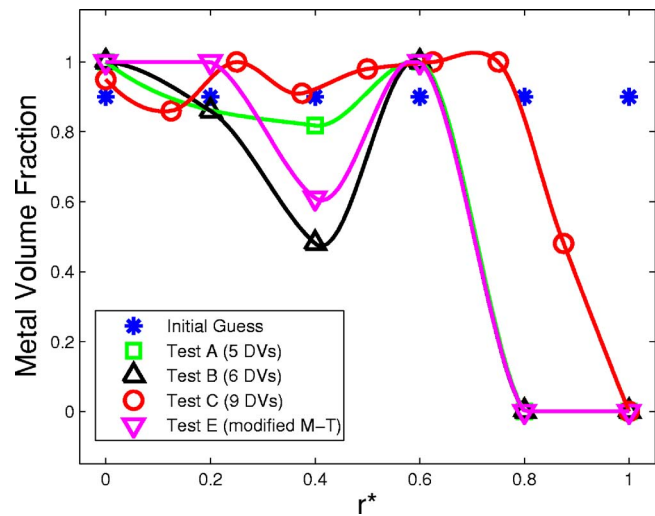


Fig. 11 Metal volume fraction profile for the stress minimization case without inner pressure. The initial guess (*) and optimal metal volume fraction function with five (□), six (△), and nine (○) design variables selected. The solution with the modified Mori-Tanaka scheme with six design variables case is also shown (▽).

the case of 6 design variables (test B) are shown in Fig. 12. The history of the objective function for the same case can be seen in Fig. 13. Notice the large initial value of the objective function since the metal-rich material is exposed to the high temperatures, and the actual tangential stress (which is compressive towards the outer radius) overcomes the critical effective compressive stress significantly. For tests A, B, and C we use one restart from a perturbed state of the design variables with a relative size of less than 0.1%. The details of the perturbation are given in Table 2 for the case of six design variables (test B). Subsequent perturbations of similar magnitude produced no additional gain in the objective function value.

The minimizers obtained (all of the solutions in this study are local minimizers; none can be guaranteed to be a global minimizer) in tests A and B (five and six design variables) show a remarkable feature: multiple graded regions and a volume fraction profile that is a non-monotonic function. We call such a material *non-monotonic functionally graded material* or NM-FGM. Test C with nine design variables stops at a local minimum far from the better results of tests A and B.

The explanation for the existence of such minimizers and of NM-FGMs in this problem can be understood once we analyze the temperature-dependent data in Tables 4 and 5. The critical tensile strength of the ceramic component is highly non-linear in terms of temperature, and it features a peak at around 500°C , while rapidly dropping at higher and lower nearby temperatures. This allows insertion of more than 50% ceramic at $r^* = 0.4$ for the local optimal design of test B since at that location the temperature is around that value. A slightly larger volume fraction of ceramic at that location would increase the objective function value since the critical tensile stress would be reduced while the effective tangential stress would increase. A slightly lower amount of ceramic at that same location would result in an increase in temperature, which in turn reduces the yield stress for the metal. It is interesting to observe that, when starting from an initial design as in test D (see Fig. 14 and Table 1), which represents a material featuring a *monotonic functionally graded region* (which we will call an M-FGM), the final design gives a value of the objective function better by no more than 5% compared to the NM-FGM design. Even this monotonic design is not representable by a power-law type function.

We also performed tests E and F (see Table 1, and Figs. 11 and

Table 1 Stress-minimization problem with no inner pressure. Objective function values at initial and final design and the number of optimization iterations performed in each case.

Test	A Five DVs	B Six DVs	C Nine DVs	D Best ^a Six DVs	E Six DVs M-T	F Best ^b DVs M-T
Obj. function value	4.980	4.980	4.980	0.534	4.959	0.530
$\mathcal{F}(\mathbf{y})$ at initial design						
Obj. function values	1.760	0.805	1.075	No restart	0.775	0.523
$\mathcal{F}(\mathbf{y})$ at restarts						
Obj. function value	0.554 (6) ^c	0.544 (9)	0.997 (6)	0.514 (2)	0.531 (10)	0.503 (6)
$\mathcal{F}(\mathbf{y})$ at final design						

^aSee Fig. 14 for initial guess for the design variables.

^bSee Fig. 14 for initial guess for the design variables.

^cThe number in parentheses indicates the total number of optimization iterations performed until reaching convergence.

14) in which, instead of the ROM scheme for evaluating the effective material properties at a point where the metal is mixed with the ceramic, we use the more accurate modified Mori-Tanaka/fuzzy inference method (see Appendix B for details). The optimal metal volume fraction function when the initial design starts with a constant metal-rich composition (test E) is given in

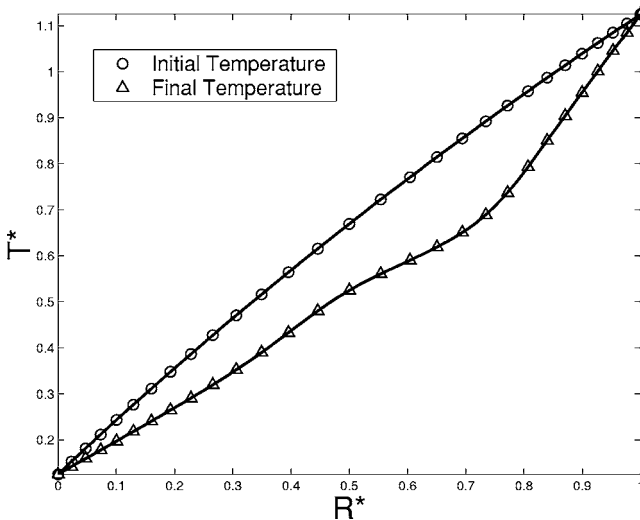


Fig. 12 Temperature profiles along the radial direction for the initial and final material design for the stress minimization problem with six design variables (test B)

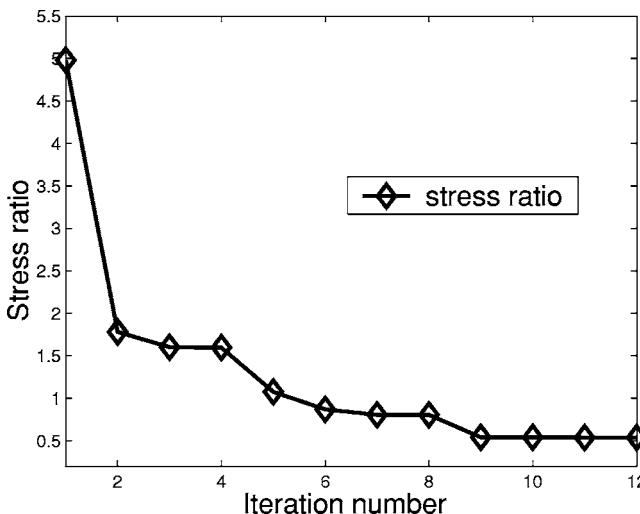


Fig. 13 History of the objective function for the stress minimization problem with six design variables (test B)

Fig. 11. Again, a NM-FGM is obtained and the difference from the results of test B, in which the ROM is used, is only quantitative. Qualitatively the results look the same. Note that we still compute the effective critical stresses using the ROM according to Eq. (18). For the initial guess in Fig. 14 (test F) the M-FGM final design reaches a slightly better value for the objective function than the NM-FGM of test E.

The second set of tests is with inner pressure as in Problem 1. The inner pressure increases the effective tensile tangential stresses. We expect, therefore, a reduction in the ceramic volume fraction around $r^* = 0.4$ that we obtained in test B. We perform several tests (G, H, and K), starting with various initial guesses for six design variables, as shown in Fig. 15. The initial and final values of the objective function for tests G, H, and K are given in Table 3. The result from test G confirms the expectations mentioned above: no more than 25% ceramic volume is allowed in the NM-FGM design, compared to over 52% in the case without pressure (test B). The best result is obtained from test K where the final volume fraction function is monotonic (M-FGM) but still not reproducible by a power-law type function. The objective function values of the NM-FGM design of test G and that of the M-FGM from test K are less than 5% apart. In contrast, the design of test H, which may be closest to be approximated by a power-law function, is 17% worse than the design of test K. The design of test K is quite interesting as it can be roughly approximated by a multi-layered FGM with two graded regions alternated by two homogeneous regions. To study the convergence in terms of the number of design variables of the design in test K, we perform test L, in which we use 13 design variables with the starting guess being a perturbation of the end design of test K. The final design from test L is only a minor improvement (objective function value is reduced by less than 0.4%) over the design in test K.

No restarts have been used in tests G, H, K, and L. The inner pressure regularizes the problem, to some extent, in the sense that some of the local minima that were so prevalent in the case with no inner pressure are now eliminated. This was also the case for Problem 1 (mass minimization with stress constraints), where the higher ceramic density precluded optimal NM-FGM designs. The main source for the NM-FGM designs is the non-linear temperature dependence of the critical strength of the ceramic and the major drop in yield stress of the metal component with increased temperature.

The purely metal and ceramic layers at, respectively, the inner and outer regions of the hollow cylinder, are seen as a common feature of the best designs in the stress minimization case (Problem 2), as was the case in the mass minimization with stress constraints (Problem 1). None of the optimal designs in Problem 2 can be represented by a power-law type function (see results from tests A, B, D, E, F, G, K, L).

We conclude that three factors play a major role in optimal design of material composition for FGMs and in generation of optimal solutions with multiple graded regions and non-monotonic graded profiles:

Table 2 Perturbations used in the stress minimization problem for restarting the iterations in the case of six design variables

Design variables							$\mathcal{F}(y)^b$
Initial guess	0.9 (0.70) ^a	0.9 (0.76)	0.9 (0.82)	0.9 (0.88)	0.9 (0.94)	0.9 (1.0)	4.981
Final values	1.0	0.97192	0.55996	1.0	0.0	0.443291	0.805
Perturbed restart	1.0	0.97193	0.55997	1.0	0.0	0.44329	0.805
Final design	1.0	0.856	0.479	1.0	0.0	0.0	0.544

^aThe numbers in parentheses indicate the radial coordinates of the design variable.

^bThe value of the objective function defined by Eqs. (24) and (25).

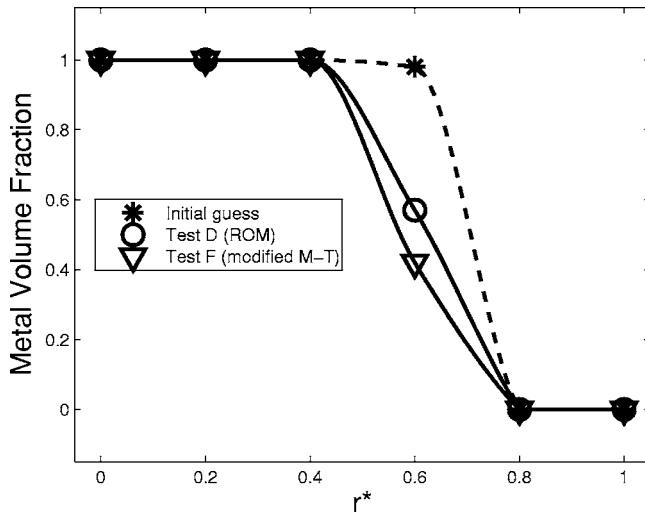


Fig. 14 The best metal volume fraction profile for the stress minimization problem with no inner pressure. Initial guess (*) and the optimal solutions with six design variables with the ROM (○) or the modified Mori-Tanaka method (▽).

- the consideration of temperature-dependent materials properties, especially the dependence of the critical stresses on temperature;
- the definition of the objective function and constraints. Here we used a scalar critical stress measure to optimize against reaching critical stress values anywhere in the FGM.

5 Conclusions

We considered the problem of optimal material design of metal-ceramic FGMs with temperature-dependent materials parameters. We have solved the 2D thermoelastic problem in plane strain for the cross section of a hollow cylinder under thermal and mechanical loadings and defined two nonlinear programming problems in terms of variables that control the continuous metal volume fraction along the radial direction of the cylinder.

A meshfree method, the element-free Galerkin method, was used to solve the heterogeneous thermoelastic problems. The advantage to using the EFG over, for example, the finite element method, may seem minimal here: better resolution of the continuous volume fraction variation with fewer nodes. For combined shape-material optimization problems that involve large shape changes, the use of such a meshfree method is critical.

We introduced new optimization criteria in terms of temperature-dependent critical stresses and solved two problems:

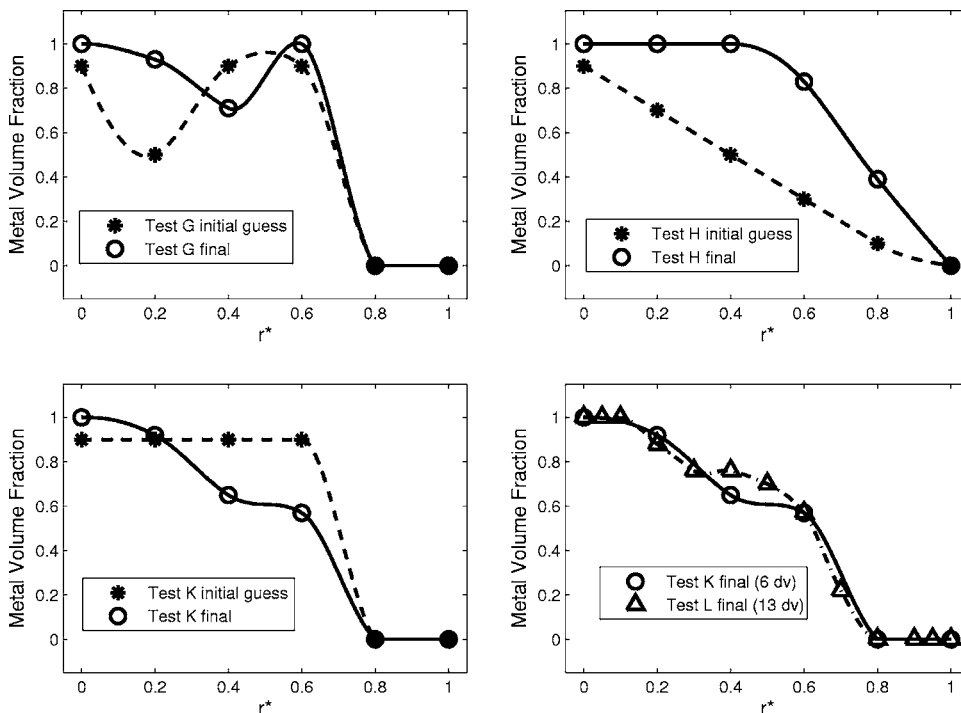


Fig. 15 Initial and final profile of the metal volume fraction for the stress minimization problem with inner pressure. The designs of tests G and K provide the best values of the objective function. Convergence in terms of the number of design variables is shown by comparing test K and L final profiles.

Table 3 Stress-minimization problem with inner pressure. Objective function values at initial and final design and the number of optimization iterations performed in each case. See Fig. 15 for initial guesses for the design variables.

Test	G Six DVs	H Six DVs	K Best Six DVs	L 13 DVs
Obj. function value	1.60	1.026	0.970	0.836
$\mathcal{F}(\mathbf{y})$ at initial design				
Obj. function value	0.840 (7) ^a	0.940 (5)	0.806 (4)	0.803 (6)
$\mathcal{F}(\mathbf{y})$ at final design				

^aThe number in parentheses indicates the total number of optimization iterations performed.

a mass minimization with stress constraints, and a stress minimization problem. Due to the higher density of the ceramic component used here, the mass minimization with stress constraints led to optimal shapes of the volume fraction function that featured a single thin functionally graded region following a narrow purely ceramic thermally protective layer. The optimal solution did not change whether fixed metal and ceramic coatings at the inner and outer regions were imposed or not. Attention has to be paid to the formulation of the constraints, since mathematically equivalent forms do not lead to numerically equivalent solutions due to differences in scaling.

In the stress minimization problem we obtained a variety of interesting minimizers given by *non-monotonic* volume fraction functions or profiles with *multiple graded regions*. Designs featuring non-monotonic volume fraction variations produced values of the objective functions which were close to designs with multiple functionally graded regions alternating with homogeneous layers.

The method for computing critical stresses for a mixture, and the non-linearity of the temperature-dependent critical stress of

the ceramic component combined with the rapid drop with increased temperature of the metal yield stress, were the most important factors that led to the particular shapes of optimal volume fraction function designs. Other critical aspects that control the material design of FGMs were found to be the selection of the objective function and constraints, and the specific thermomechanical loading. The particular effective medium theory used for computing material properties in regions of material mixtures (the rule-of-mixture and the more complex and realistic combination of Mori-Tanaka and fuzzy inference method) also influenced the optimal design but only quantitatively.

Further analysis should be focused on determining measures of critical stresses in FGMs that correlate well with experiments. The methodology developed in the present work can directly use such results if they become available.

The results obtained here advocate for the need of allowing sufficient generality in the selection of the function that represents the volume fraction of the FGM components. We have shown that power-law type variations for the volume fraction function may exclude significantly better solutions given by non-monotonic or multi-layered functionally graded profiles.

Acknowledgment

This work was partially supported by a NASA Epscor Space grant and by a University of Nebraska-Lincoln Layman award. We also thank Han Jiang for the help provided in running some of the numerical tests.

Appendix A: Temperature-Dependent Material Properties

We consider metal-ceramic FGMs made from titanium alloy Ti-6Al-4V and zirconium oxide ZrO₂. In Tables 4 and 5, we list the thermomechanical properties of ZrO₂ and Ti-6Al-4V, respectively. The values are from [41], with a few exceptions for which

Table 4 Temperature-dependent material properties for ZrO₂

T	(°C)	20	1000	1200					
E	(GPa)	175	133	119					
T	(°C)	20	500	1000	1050	1080			
α	(1/°C × 10 ⁻⁶)	6.53	6.53	7.59	7.72	8.0			
T	(°C)	20	100	200	400	800	1200		
κ	(W/m °C)	2	2	2	2	2.2	2.5		
T	(°C)	20	200	400	500	600	800	1000	1100
σ_{tc}	(MPa)	126	117	122	140	123	112	105 ^a	96
T	(°C)	20	500	1000	1200				
σ_{cc}	(MPa)	2050	1600	1200	798				

^aLinearly interpolated value from values at 800°C and 1100°C.

Table 5 Temperature-dependent material properties for Ti-6Al-4V

T	(°C)	20	1000							
E	(GPa)	113.8	113.8							
T	(°C)	20	100	205	315	425	540	650	815	900
α	(1/°C × 10 ⁻⁶)	8.6	8.6	9.0	9.2	9.4	9.5	9.7	10.5 ^a	10.91
T	(°C)	20	500	1000						
κ	(W/m °C)	6.6	6.6	6.6						
T	(°C)	20	315	425	540	815	900			
σ_{tc}	(MPa)	921	655	572	427	(80) ^b	(1)			

^aLinearly extrapolated value.

^bAkima spline extrapolated values.

we linearly extrapolate from the last two values in the tables. In one case, using linear extrapolation would give negative values for the critical stress at high temperatures. In such a case, we employ the Akima shape-preserving spline interpolating the given values to extrapolate to values not given in [42]. We take the critical stresses in compression (σ_{cc}) for the metal to be the same as the critical tensile stresses (σ_{tc}). In calculating material properties for temperature values between the data points in these tables, we use Akima spline interpolation.

Appendix B: Effective Material Properties

The effective material properties at a point in the graded part of the FGM composite can be evaluated using various micromechanics-based formulations, such as the simple ROM, the Mori-Tanaka method, the self-consistent method, etc. Here we use the ROM and also a modified Mori-Tanaka with fuzzy inference briefly described below.

For the rule of mixture, a generic material property P of the FGM is expressed in terms of the corresponding material properties of the matrix (P_m) and the inclusion (P_i) components as follows

$$P(r) = y(r)P_i(r) + [1 - y(r)]P_m(r) \quad (B1)$$

where $y(r)$ is the inclusion's volume fraction at a location r in the FGM.

We use the Mori-Tanaka method to calculate effective properties of FGM at a point where one of the components is dilute: no more than 30% (volume fraction is 0.3), so that the component can still be considered as the inclusion in a matrix. We find the bulk (K) and the shear (G) moduli using the Mori-Tanaka method

$$K = K_m + \frac{y(r)(K_i - K_m)(3K_m + 4G_m)}{3K_m + 4G_m + 3[1 - y(r)](K_i - K_m)}$$

$$G = G_m + \frac{y(r)(G_i - G_m)[5G_m(3K_m + 4G_m)]}{5G_m(3K_m + 4G_m) + 6[1 - y(r)](K_m + 2G_m)(G_i - G_m)}$$

$$\alpha = y(r)\alpha_i + [1 - y(r)]\alpha_m + \frac{y(r)[1 - y(r)](\alpha_i - \alpha_m)(K_i - K_m)}{3K_iK_m/4G_m + [1 - y(r)]K_m}$$

$$\kappa = \kappa_m + \frac{3y(r)\kappa_m(\kappa_i - \kappa_m)}{[1 - y(r)](\kappa_i - \kappa_m) + 3\kappa_m} \quad (B2)$$

Here, $y(r)$ is the volume fraction of the inclusion at point r and subscripts i and m stand for the inclusion and the matrix, respectively. Young's modulus E and Poisson's ratio ν are calculated from K and G . For locations where the inclusions cannot be easily distinguished from the matrix (we take the range $0.3 \leq y(r) \leq 0.7$), the material parameters are evaluated by means of a fuzzy inference [10]. Note that the material properties used here are functions of position as well as temperature.

References

- [1] Pindera, M.-J., Aboudi, J., and Arnold, S., 1998, "Thermomechanical Analysis of Functionally Graded Thermal Barrier Coatings With Different Microstructural Scales," *J. Am. Ceram. Soc.*, **81**(6), pp. 1525–1536.
- [2] Horgan, C., and Chan, A., 1999, "Pressurized Hollow Cylinder or Disk Problem for Functionally Graded Isotropic Linearly Elastic Materials," *J. Elast.*, **55**(1), pp. 43–59.
- [3] Yildirim, B., and Erdogan, F., 2004, "Edge Crack Problems in Homogenous and Functionally Graded Material Thermal Barrier Coatings Under Uniform Thermal Loading," *J. Therm. Stresses*, **27**(4), pp. 311–329.
- [4] Zimmerman, R. W., and Lutz, M. P., 1999, "Thermal Stresses and Thermal Expansion in a Uniformly Heated Functionally Graded Cylinder," *J. Therm. Stresses*, **22**(2), pp. 177–188.
- [5] Nadeau, J., and Meng, X., 2000, "On the Response Sensitivity of an Optimally Designed Functionally Graded Layer," *Composites, Part B*, **31**(4), pp. 285–297.
- [6] Kawamura, R., and Tanigawa, Y., 1998, "Multipurpose Optimization Problem of Material Composition for Thermal Stress Relaxation Type of a Functionally Graded Plate," *JSME Int. J., Ser. A*, **41**(3), pp. 318–325.
- [7] Ootao, Y., Kawamura, R., Tanigawa, Y., and Nakamura, T., 1998, "Neural Network Optimization of Material Composition of a Functionally Graded Material Plate at Arbitrary Temperature Range and Temperature Rise," *Arch. Appl. Mech.*, **68**, pp. 662–676.
- [8] Tanaka, K., Tanaka, Y., Enomoto, K., Poterasu, V., and Sagano, Y., 1993, "Design of Thermoelastic Materials Using Direct Sensitivity and Optimization Methods. Reduction of Thermal Stresses in Functionally Graded Materials," *Comput. Methods Appl. Mech. Eng.*, **106**(1–2), pp. 271–284.
- [9] Tanaka, K., Tanaka, Y., Watanabe, H., Poterasu, V. F., and Sugano, Y., 1993, "Improved Solution to Thermoelastic Material Design in Functionally Graded Materials; Scheme to Reduce Thermal Stresses," *Comput. Methods Appl. Mech. Eng.*, **109**(3–4), pp. 377–389.
- [10] Tanaka, K., Watanabe, H., Sugano, Y., and Poterasu, V., 1996, "Multicriterial Material Tailoring of a Hollow Cylinder in Functionally Graded Materials: Scheme to Global Reduction of Thermoelastic Stresses," *Comput. Methods Appl. Mech. Eng.*, **135**(3–4), pp. 369–380.
- [11] Tanigawa, Y., Oka, N., Akai, T., and Kawamura, R., 1997, "One-Dimensional Transient Thermal Stress Problem for Nonhomogeneous Hollow Circular Cylinder and its Optimization of Material Composition for Thermal Stress Relaxation," *JSME Int. J., Ser. A*, **40**(2), pp. 117–127.
- [12] Cho, J., and Choi, J., 2004, "A Yield-Criteria Tailoring of the Volume Fraction in Metal-Ceramic Functionally Graded Material," *Eur. Phys. J. A*, **23**(2), pp. 271–281.
- [13] Goupee, A., and Vel, S., 2006, "Two-Dimensional Optimization of Material Composition of Functionally Graded Materials Using Meshless Analyses and a Genetic Algorithm," *Comput. Methods Appl. Mech. Eng.*, **195**, pp. 5926–5948.
- [14] Lipton, R., 2002, "Design of Functionally Graded Composite Structures in the Presence of Stress Constraints," *Int. J. Solids Struct.*, **39**, pp. 2575–2586.
- [15] Boussaa, D., 2000, "Optimizing a Compositionally Graded Interlayer to Reduce Thermal Stresses in a Coated Tube," *C. R. Acad. Sci., Ser. IIb Mec.*, **328**(3), pp. 209–215.
- [16] Turteltaub, S., 2002, "Optimal Control and Optimization of Functionally Graded Materials for Thermomechanical Processes," *Int. J. Solids Struct.*, **39**(12), pp. 3175–3197.
- [17] Li, Y., Ramesh, K., and Chin, E., 2001, "Dynamic Characterization of Layered and Graded Structures Under Impulsive Loading," *Int. J. Solids Struct.*, **38**(34–35), pp. 6045–6061.
- [18] Kim, J.-H., and Paulino, G. H., 2005, "Consistent Formulations of the Interaction Integral Method for Fracture of Functionally Graded Materials," *ASME J. Appl. Mech.*, **72**(3), pp. 351–364.
- [19] Cho, J., and Oden, J. T., 2000, "Functionally Graded Material: A Parametric Study on Thermal-Stress Characteristics Using the Crank-Nicolson-Galerkin Scheme," *Comput. Methods Appl. Mech. Eng.*, **188**(1), pp. 17–38.
- [20] Reddy, J., 2000, "Analysis of Functionally Graded Plates," *Int. J. Numer. Methods Eng.*, **47**(1), pp. 663–684.
- [21] Shabana, Y. M., and Noda, N., 2002, "Thermo-Elasto-Plastic Stresses of Functionally Graded Material Plate With a Substrate and a Coating," *J. Therm. Stresses*, **25**(12), pp. 1133–1146.
- [22] Sladek, J., Sladek, V., and Zhang, C., 2004, "A Local Biem for Analysis of Transient Heat Conduction With Nonlinear Source Terms in FGMS," *Eng. Anal. Boundary Elem.*, **28**(1), pp. 1–11.
- [23] Sutradhar, A., and Paulino, G. H., 2004, "The Simple Boundary Element Method for Transient Heat Conduction in Functionally Graded Materials," *Comput. Methods Appl. Mech. Eng.*, **193**(42–44), pp. 4511–4539.
- [24] Bobaru, F., and Jiang, H., 2004, "Optimization of Functionally Graded Materials With Temperature Dependent Properties: A Meshfree Solution," in: *The XXI International Congress of Theoretical and Applied Mechanics*, Warsaw, Poland, ISBN 83-89687-01-1.
- [25] Qian, L., Batra, R., and Chen, L., 2004, "Analysis of Cylindrical Bending Thermoelastic Deformations of Functionally Graded Plates by a Meshless Local Petrov-Galerkin Method," *Comput. Mech.*, **33**(4), pp. 263–273.
- [26] Sladek, J., Sladek, V., and Zhang, C., 2003, "Transient Heat Conduction Analysis in Functionally Graded Materials by the Meshless Local Boundary Integral Equation Method," *Comput. Mater. Sci.*, **28**(3–4), pp. 494–504.
- [27] Noda, N., 1999, "Thermal Stresses in Functionally Graded Materials," *J. Therm. Stresses*, **22**(4–5), pp. 477–512.
- [28] Jin, Z.-H., and Batra, R., 1998, "Thermal Fracture of Ceramics With Temperature-Dependent Properties," *J. Therm. Stresses*, **21**, pp. 157–176.
- [29] Liew, K., Yang, J., and Kitipornchai, S., 2004, "Thermal Post-Buckling of Laminated Plates Comprising Functionally Graded Materials With Temperature-Dependent Properties," *ASME J. Appl. Mech.*, **71**(6), pp. 839–850.
- [30] Belytschko, T., Lu, Y., and Gu, L., 1994, "Element-Free Galerkin Methods," *Int. J. Numer. Methods Eng.*, **37**(2), pp. 229–256.
- [31] Chen, J.-S., and Wang, H.-P., 2000, "New Boundary Condition Treatments in Meshfree Computation of Contact Problems," *Comput. Methods Appl. Mech. Eng.*, **187**(3), pp. 441–468.
- [32] Rabczuk, T., and Belytschko, T., 2005, "Adaptivity for Structured Meshfree Particle Methods in 2D and 3D," *Int. J. Numer. Methods Eng.*, **63**, pp. 1559–1582.
- [33] Bobaru, F., and Rachakonda, S., 2006, "E(FG)²: A New Fixed-Grid Shape Optimization Method Based on the Element-Free Galerkin Meshfree Analysis," *Struct. Multidiscip. Optim.*, **32**(3), pp. 215–228.
- [34] Bobaru, F., and Mukherjee, S., 2002, "Meshless Approach to Shape Optimization of Linear Thermoelastic Solids," *Int. J. Numer. Methods Eng.*, **53**(4), pp. 765–796.

- [35] Reiter, T., and Dvorak, G., 1998, "Micromechanical Models for Graded Composite Materials: II. Thermomechanical Loading," *J. Am. Soc. Mass Spectrom.*, **46**, pp. 1655–1673.
- [36] Nocedal, J., and Wright, S., 1999, *Numerical Optimization*, Springer-Verlag, New York.
- [37] Chen, J.-S., Pan, C., Wu, C.-T., and Liu, W. K., 1996, "Reproducing Kernel Particle Methods for Large Deformation Analysis of Non-Linear Structures," *Comput. Methods Appl. Mech. Eng.*, **139**(1-4), pp. 195–227.
- [38] Li, S., and Liu, W. K., 2002, "Meshfree and Particle Methods and Their Applications," *Appl. Mech. Rev.*, **55**(1), pp. 1–34.
- [39] Suresh, S., and Mortensen, A., 1998, *Fundamentals of Functionally Graded Materials*, Institute of Materials, London.
- [40] Akima, H., 1970, "New Method of Interpolation and Smooth Curve Fitting Based on Local Procedures," *J. Assoc. Comput. Mach.*, **17**(4), pp. 589–602.
- [41] Bobaru, F., and Mukherjee, S., 2001, "Shape Sensitivity Analysis and Shape Optimization in Planar Elasticity Using the Element-Free Galerkin Method," *Comput. Methods Appl. Mech. Eng.*, **190**(32-33), pp. 4319–4337.
- [42] Shackelford, J., and Alexander, W., 2001, *CRC Materials Science and Engineering Handbook*, CRC Press, Boca Raton, FL.

Thermoelastic Dynamic Instability (TEDI) in Frictional Sliding of a Half-Space Against a Rigid Non-Conducting Wall

L. Afferrante

M. Ciavarella

CEMEC-PoliBA-Centre of Excellence in
Computational Mechanics,
Politecnico di Bari,
V.le Japigia 182,
70125 Bari, Italy

In the sliding of half-spaces with constant friction coefficient, two classes of instabilities are well known: thermoelastic instability (TEI), which occurs for sufficiently long wavelengths and Dynamic Instability (DI), which happens at sufficiently high friction coefficient, and whose growth factor increases linearly with wave number. Although the two phenomena look therefore quite distinct, their coupling is discussed here for an elastic and conducting half-space sliding against a rigid and non-conducting wall. The coupling between thermal and dynamic effects is not always negligible. In fact, surprisingly, new areas of instability are found, called thermoelastic dynamic instabilities (TEDI), similar to TEI at high speeds and DI at low speeds. TEDI lowers the critical speed and friction coefficient in many conditions even to zero. At low speeds, TEDI is ill-posed as DI at small wavelengths, and hence a regularized friction law like the Rice-Ruina one would probably be needed to correct the results. [DOI: 10.1115/1.2712232]

Keywords: thermoelastic contact, TEI, thermoelastic instability, squeal, frictional vibrations

1 Introduction

Many models of friction explain the intrinsically “unstable” nature of friction, including the difference between static and dynamic coefficient, the dependence of friction on temperature, pressure, speed, and so on (Dowson [1]). In addition, scatter, oscillations, and even “capriciousness” in friction experiments are so typical that they do not discourage the even unexperienced experimentalist (Duffour and Woodhouse [2]).

Friction instabilities are of interest in the sliding of tectonic plates during earthquakes (Ben-Zion [3], Rice et al. [4]), in the sliding of rubber-like materials (Barquins et al. [5]), in the generation of noise and vibration (Kinkaid et al. [6]), or the appearance of hot spots and other thermally induced vibrations in automotive brakes or clutches (Kennedy and Ling, [7], Anderson and Knapp, [8]). Moreover, stick-slip vibrations are a limiting factor in machine tool slides and other linear positioning devices (Popp and Rudolph [9]). Citing Johnson [10]: “*Dynamic friction, leading to frictionally excited vibrations, is a vast subject, covering extreme ranges of space and time: from seismic faults stretching many kilometers to crystal lattice spacing in the Atomic Friction Microscope; from the years which separate earthquakes to the kilohertz frequencies of squealing brakes and railway wheels.*” A full understanding of the transient response is often impossible, and the definition of frictional instabilities is often very helpful, even assuming the simplest models of Coulomb constant coefficient of friction. Limiting the attention to this case (of constant friction coefficient) there are two basic frictional instability phenomena, those involving Dynamic Instabilities (DI) and those involving thermoelastic instabilities (TEI). They result from the interaction between relatively simple physical processes, notably, the elastic deformation of the contacting bodies, the development of frictional forces at the interface opposing the motion, and the consequent generation of frictional heat. However, they are typically

treated separately, as people interested in DI often neglect thermal aspects, and vice versa, under the assumption that the time scales involved are widely separated. We shall briefly recollect some basic findings about DI and TEI, before proceeding with a new coupled analysis, including both inertia terms, and frictional heating.

1.1 Dynamic Instabilities. The dynamic class of friction instabilities includes unstable propagation of surface waves in continuous bodies. The existence of self-excited oscillations at the sliding interface can be understood if we start from the simplest case of a free surfaces, for which we can only have propagation of Rayleigh waves. These waves have amplitudes which decay exponentially with distance from the free surface. For bonded interfaces, Stoneley found that if the shear wave speeds of the contacting materials do not differ greatly, a different type of waves can exist, named Stoneley waves (Stoneley, [11]). They can travel along the interface of two bonded contacting elastic bodies, i.e., when friction is infinite. At the other extreme, i.e., for frictionless contact, Achenbach and Epstein [12] and Murty [13] then found “smooth contact Stoneley waves,” “slip waves,” or finally “frictionless generalized Rayleigh waves,” which occur for a wider range of material combinations than the Stoneley waves for bonded contact. Recently, a lot of analytical effort has been devoted to the study of similar types of dynamic stability of frictional sliding. In particular, Martins [14] considered a single half-plane sliding against a rigid wall, and found that instability emerges only for friction coefficients larger than 1. However, Adams [15], in the general case for two elastic half-planes, showed that the critical coefficient can be much lower than 1 and can even be zero for two elastic materials of similar material properties, especially in the case of layered geometries (Adams [16]). More precisely, when the so-called “generalized Rayleigh waves” exist already for the frictionless contact, steady frictional sliding along an interface between dissimilar elastic solids with Coulomb friction is unstable for *arbitrarily small* values of friction (Rice et al. [4], Ranjith et al. [17]), Ranjith et al. [18]).

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received April 7, 2006; final manuscript received October 16, 2006. Review conducted by Antoinette Maniatty.

1.2 Thermoelastic Instabilities. Another, so far separate, class of instability, is that due to coupling of thermal and mechanical contact boundary conditions at the interface. This class of instability emerges either with static contact, when the conduction of heat across an interface between two thermoelastic bodies *even in the absence of sliding* and hence frictional heating, or when there is a pressure-dependent thermal contact resistance at the interface that varies with local contact pressure, as is reasonable to assume, because the microscopic contact area varies with pressure (Barber et al. [19], Clausen and Chao [20]). An example application involves the solidification of a metal against a plane mold, where thermoelastic contact between the partially solidified casting and the mold can be unstable, leading to non-uniform pressure distribution and alloy composition (Richmond and Huang [21], Yigit and Barber [22], Yigit [23]).

In the case of frictional sliding, the typical and most important effect is that frictional heating at the interface depends on the contact pressure, and this couples the thermal and mechanical parts of the problem, giving rise to frictionally excited thermoelastic instability (TEI) (Barber [24], Dow and Burton [25]). In this process, any perturbation in contact pressure causes a corresponding perturbation in heating and hence thermal distortion, which exaggerates the initial perturbation. TEI has received more attention in technological applications for which it is of critical importance, such as the design of brakes and clutches (Kennedy and Ling [7], Anderson and Knap [8]). For a given friction coefficient, there will be some sliding speed V_{cr} above which the system will be unstable (whereas DI is typically due to a critical friction coefficient alone).

1.3 Thermoelastic Dynamic Instabilities. Recently (Afferrante et al. [26]), we discovered an interesting coupling between dynamics and thermal effects, which we named TEDI, in a simple 1D model, with a constant coefficient of friction. The model without coupling shows a trivial response; in particular, no dynamic instability, and a frictional heating (TEI) instability.¹ The coupled model leads to instability at all speeds, although at small speeds the non-linear limit cycle obtained is clearly an oscillation closer to the quasi-static approximation. At larger speeds, TEDI gives rise to significant increase of maximum or mean force and frictional heating. Hence, we expect to see that in cases where both instabilities are at play, the coupling would also show some interesting and unexpected features.

In the present paper, we aim at discussing the continuum case of the one half-space sliding against a rigid and non-conductor. This is the extension of Martins' paper (Martins et al. [14]) of DI, with frictional heating added, or the extension of Burton et al. [28] paper, with inertia terms added. In the case without frictional heating (Martins et al. [14]), the dynamic instability only occurs for large friction coefficient ($f_{cr}=1$), whereas it would be much lower if the bodies were elastically more similar. For the TEI case, a critical speed can be defined for each friction coefficient, as a function on the wavelength. We shall find that the combined case shows instability in a much larger range of cases, sometimes having features close to DI and sometimes close to TEI.

2 Problem Statement

We shall consider the case of an elastic, conducting half-space, sliding against a rigid and non-conducting wall (Fig. 1).

This case generalizes the TEI case by Dow and Burton [25], and the Martins et al. [14] case for DI. The choice of a non-conducting rigid wall geometry obviously corresponds to a limit case, which is only approximately met by practical systems such as a metallic material against a ceramic material. However, it can be shown that the general case of two elastic non-conducting ma-

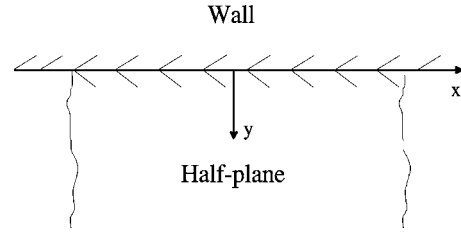


Fig. 1 Elastic half-plane sliding against a rigid wall

terials would increase the number of parameters by six (i.e., up to 12!), and then it would become almost impossible to show the coupling between thermal and dynamic effects in a significant range of all parameters, as done here, where we shall have already six independent parameters. In contrast, the TEI case by Dow and Burton [25] had only one independent parameter (!), and the Martins et al. [14] case for DI had two independent constants (friction coefficient and Poisson's ratio).

The governing equations for coupled thermoelasticity (from Boley and Weiner [29], Eqs. (1.4.1–1.4.9)) are

$$K\nabla^2 T = \rho c_E \frac{\partial T}{\partial t} + m T_0 \frac{\partial e}{\partial t} \quad (1)$$

$$\frac{\partial \sigma_{ij}}{\partial x_j} = \rho \frac{\partial^2 u_i}{\partial t^2} \quad (2)$$

with the constitutive law

$$\sigma_{ij} = \lambda \delta_{ij} e + 2\mu e_{ij} - m \delta_{ij} (T - T_0) \quad (3)$$

and

$$e_{ij} = \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \quad m = (3\lambda + 2\mu)\alpha \quad (4)$$

($K, \alpha, \rho, c_E, \lambda, \mu$ are the thermal conductivity, the coefficient of thermal expansion, the density, the specific heat, and Lamé's constants of the material, respectively).

Eliminating the stresses between (2)–(4), we have

$$\mu \nabla^2 u_i + (\lambda + \mu) \frac{\partial e}{\partial x_i} - m \frac{\partial T}{\partial x_i} = \rho \frac{\partial^2 u_i}{\partial t^2} \quad (5)$$

If we neglect the coupled term in the heat equation (see Boley and Weiner [29], Eringen [30], Chadwick [31], Nowacki [32]) and consider plane strain conditions ($u_z=0$ and $e = \partial u_x / \partial x + \partial u_y / \partial y$), Eqs. (1) and (5) reduce to

$$\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} - \frac{1}{k} \frac{\partial T}{\partial t} = 0 \quad (6)$$

$$(\lambda + 2\mu) \frac{\partial^2 u_x}{\partial x^2} + \mu \frac{\partial^2 u_x}{\partial y^2} + (\lambda + \mu) \frac{\partial^2 u_y}{\partial x \partial y} - m \frac{\partial T}{\partial x} - \rho \frac{\partial^2 u_x}{\partial t^2} = 0 \quad (7)$$

$$(\lambda + \mu) \frac{\partial^2 u_x}{\partial x \partial y} + \mu \frac{\partial^2 u_y}{\partial x^2} + (\lambda + 2\mu) \frac{\partial^2 u_y}{\partial y^2} - m \frac{\partial T}{\partial y} - \rho \frac{\partial^2 u_y}{\partial t^2} = 0 \quad (8)$$

We assume that the thermoelastic body occupies the region $y > 0$ and that in this region the displacements and temperature can be written as

$$u_x = \Re\{A \exp(i\omega x - sy + bt)\} \quad (9)$$

$$u_y = \Re\{B \exp(i\omega x - sy + bt)\} \quad (10)$$

$$T = \Re\{C \exp(i\omega x - sy + bt)\} \quad (11)$$

Substituting these results into (6)–(8) and canceling the common exponential factors, we obtain the three algebraic equations

¹However, the TEI critical speed is only due to the displacement constraint in the normal direction and could be removed by considering a force control in normal direction (Afferrante and Ciavarella [27]).

$$\left(-\omega^2 + s^2 - \frac{b}{k}\right)C = 0 \quad (12)$$

$$[-\omega^2(\lambda + 2\mu) + \mu s^2 - b^2\rho]A - \iota(\lambda + \mu)\omega sB - \iota m\omega C = 0 \quad (13)$$

$$-\iota(\lambda + \mu)\omega sA + [(\lambda + 2\mu)s^2 - \mu\omega^2 - b^2\rho]B + msC = 0 \quad (14)$$

One eigenvalue for s^2 is clearly given by Eq. (12) as

$$s_3^2 = \omega^2 + \frac{b}{k} \quad (15)$$

The other two eigenvalues correspond to eigenfunctions for which $C=0$ and are solutions of the equation

$$[c_1^2(s^2 - \omega^2) - b^2][c_2^2(s^2 - \omega^2) - b^2] = 0 \quad (16)$$

where

$$c_1^2 = \frac{(\lambda + 2\mu)}{\rho}, \quad c_2^2 = \frac{\mu}{\rho} \quad (17)$$

are the dilatational and shear wave speeds, respectively. Hence, the two solutions for s_1^2, s_2^2 are

$$s_1^2 = \omega^2 + \frac{b^2}{c_1^2}, \quad s_2^2 = \omega^2 + \frac{b^2}{c_2^2} \quad (18)$$

To extract the eigenfunctions, it is convenient to rewrite Eqs. (13) and (14) in terms of the wave speeds. We have

$$[-c_1^2\omega^2 + c_2^2s^2 - b^2]A - \iota\omega(c_1^2 - c_2^2)sB = \frac{\iota m\omega C}{\rho} \quad (19)$$

$$-\iota\omega(c_1^2 - c_2^2)sA + [-c_2^2\omega^2 + c_1^2s^2 - b^2]B = -\frac{msC}{\rho} \quad (20)$$

When $s=s_1$, $C=0$, and these equations are both satisfied if

$$\iota s_1 A = \omega B \quad (21)$$

so a suitable eigenfunction set is

$$s_1 = \sqrt{\omega^2 + \frac{b^2}{c_1^2}}, \quad A_1 = \omega K_1, \quad B_1 = \iota s_1 K_1, \quad C_1 = 0 \quad (22)$$

where the arbitrary constants A_1, B_1, C_1 have been eliminated in favor of a new arbitrary constant K_1 , representing the degrees of freedom associated with the eigenvalue s_1 .

With $s=s_2$ and $C=0$, Eqs. (19) and (20) are both satisfied if

$$A\omega = -\iota s_2 B \quad (23)$$

so a suitable eigenfunction set is

$$s_2 = \sqrt{\omega^2 + \frac{b^2}{c_2^2}}, \quad A_2 = -\iota s_2 K_2, \quad B_2 = \omega K_2, \quad C_2 = 0 \quad (24)$$

where the arbitrary constant K_2 represents the degrees of freedom associated with the eigenvalue s_2 .

With $s=s_3$, $C \neq 0$, and we must solve the inhomogeneous equations (19) and (20). We obtain

$$A = \frac{\iota m\omega C}{\rho b(c_1^2/k - b)}, \quad B = -\frac{ms_3 C}{\rho b(c_1^2/k - b)}, \quad (25)$$

so a suitable eigenfunction set is

$$s_3 = \sqrt{\omega^2 + \frac{b}{k}}, \quad A_3 = \iota m\omega K_3, \quad B_3 = -ms_3 K_3, \quad C_3 = \rho b \left(\frac{c_1^2}{k} - b \right) K_3 \quad (26)$$

where the arbitrary constant K_3 represents the degrees of freedom associated with the eigenvalue s_3 .

Therefore, a general solution for the temperature and displacements fields can be written as

$$u_x = \Re\{A_1 e^{(\iota\omega x - s_1 y + bt)} + A_2 e^{(\iota\omega x - s_2 y + bt)} + A_3 e^{(\iota\omega x - s_3 y + bt)}\} \quad (27)$$

$$u_y = \Re\{B_1 e^{(\iota\omega x - s_1 y + bt)} + B_2 e^{(\iota\omega x - s_2 y + bt)} + B_3 e^{(\iota\omega x - s_3 y + bt)}\} \quad (28)$$

$$T = \Re\{C_1 e^{(\iota\omega x - s_1 y + bt)} + C_2 e^{(\iota\omega x - s_2 y + bt)} + C_3 e^{(\iota\omega x - s_3 y + bt)}\} \quad (29)$$

where the constants A_i, B_i, C_i are given by (22), (24), and (26).

3 Known Results

We shall start by obtaining, as special cases, the known results of a free surface, of frictionless interface waves, of the Martins-Adams Dynamic Instability for friction with no heating effects, and the Burton TEI instability when thermal effects are included, but not inertia.

3.1 Free Surface: Rayleigh Wave Equations. If the body has a free surface and therefore no contact and no heating, we have $K_3=0$, and the traction-free condition ($\sigma_{yy}=\sigma_{xy}=0$ at $y=0$) gives

$$\sigma_{yy} = \lambda \left(\frac{\partial u_x}{\partial x} + \frac{\partial u_y}{\partial y} \right) + 2\mu \frac{\partial u_y}{\partial y} = \iota\lambda\omega(A_1 + A_2) - (\lambda + 2\mu)(s_1 B_1 + s_2 B_2) \quad (30)$$

$$\sigma_{xy} = \mu \left(\frac{\partial u_y}{\partial x} + \frac{\partial u_x}{\partial y} \right) = \iota\omega\mu(B_1 + B_2) - \mu(s_1 A_1 + s_2 A_2) \quad (31)$$

Using (17), (22), and (24), the above equations can be rewritten in the following form

$$-2\omega s_1 K_1 + \iota(\omega^2 + s_2^2)K_2 = 0 \quad (32)$$

$$\iota[c_1^2(\omega^2 - s_1^2) - 2c_2^2\omega^2]K_1 - 2c_2^2\omega s_2 K_2 = 0 \quad (33)$$

For a non-trivial solution, we require

$$4\omega^2 c_2^2 s_1 s_2 + (\omega^2 + s_2^2)[c_1^2(\omega^2 - s_1^2) - 2c_2^2\omega^2] = 0 \quad (34)$$

The Rayleigh wave solution corresponds to the case where

$$\iota\omega x + bt = \iota\omega(x - c_R t) \quad (35)$$

and hence

$$b = -\iota\omega c_R \quad (36)$$

Substituting this result into (18) and (34) and canceling a factor of $\omega^4 c_2^2$, we then obtain

$$4\sqrt{(1 - M_1^2)(1 - M_2^2)} - (2 - M_2^2)^2 = 0 \quad (37)$$

where

$$M_1 = \frac{c_R}{c_1}, \quad M_2 = \frac{c_R}{c_2} \quad (38)$$

in agreement with the classical equation for Rayleigh waves.

3.2 Frictionless Contact Interface Waves. Now, suppose the boundary conditions are frictionless contact with a rigid plane, so $u_y=\sigma_{xy}=0$ at the surface. We obtain (simplifying A_i, B_i by (22) and (24))

$$\iota s_1 K_1 + \omega K_2 = 0 \quad (39)$$

$$\omega s_1 K_1 - \iota s_2^2 K_2 = 0 \quad (40)$$

and these have a non-trivial solution if

$$s_1 = 0 \quad \text{or} \quad s_2^2 = \omega^2 \quad (41)$$

The second root corresponds to $b=0$, from (18) and hence does not correspond to a traveling wave. If $s_1=0$, There is no decay with y and hence we have simply a dilatational wave propagating parallel to the interface.

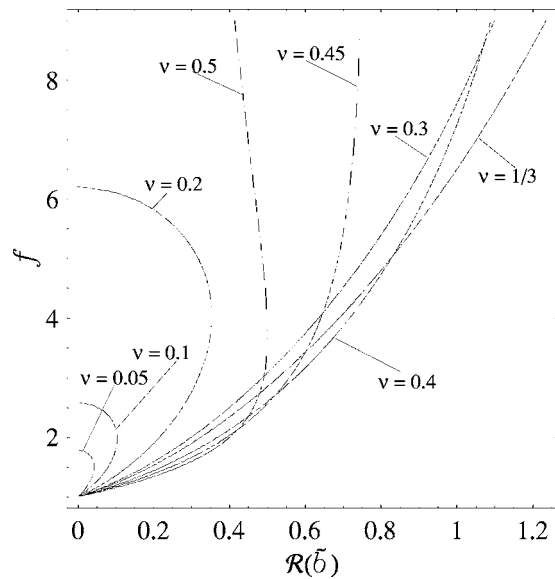


Fig. 2 Some example results for Dynamic Instability (DI): variation of the real part of the dimensionless growth rate $\tilde{b}$ with the friction coefficient f for different ν

3.3 Frictional Contact Interface Waves: DI. When considering Coulomb friction at the interface, but no frictional heating, the case treated by Martins et al. [14] is recovered. For the moment, we shall neglect frictional heating (or assume that the thermal expansion term $m=0$). The frictional tractions will act in the positive x -direction on the surface $y=0$ of $y \geq 0$, which is a “negative” x -surface, and hence σ_{xy} is negative. Thus, the mechanical boundary conditions are therefore

$$u_y(x,0,t) = 0 \quad \sigma_{xy}(x,0,t) - f\sigma_{yy}(x,0,t) = 0 \quad (42)$$

(since σ_{yy} must be compressive and hence also negative).

Substituting for A_i , B_i from Eqs. (22) and (24), we get

$$\iota s_1 K_1 + \omega K_2 = 0 \quad (43)$$

$$[-c_2^2 s_1 \omega - \iota f(c_1^2 - 2c_2^2)\omega^2 + \iota f c_1^2 s_1^2]K_1 + [\iota c_2^2 s_2^2 + 2f c_2^2 \omega s_2]K_2 = 0 \quad (44)$$

For a non-trivial solution of this equation, we require the vanishing of the determinant. Therefore, after some algebraic manipulation and by using (18), we obtain

$$s_1(\omega^2 - s_2^2) - \iota f \omega(\omega^2 + s_2^2 - 2s_1 s_2) = 0 \quad (45)$$

which is the same equation as in Martins et al. [14].

In the limit case of incompressible material ($\nu=0.5$ and $c_1 \rightarrow \infty$, $s_1 = \omega$), (45) simplifies in the following form

$$s_2 = \sqrt{\omega^2 + \frac{b^2}{c_2^2}} = \frac{\iota f - 1}{\iota f + 1} \omega \quad (46)$$

Figures 2 and 3 shows some sample results (Martins et al. [14]). In particular, the variation of the real (Fig. 2) and imaginary (Fig. 3). Part of the dimensionless growth rate $\tilde{b} = b/\omega c_2$ with the friction coefficient f is plotted for different Poisson ratios ν . We have instability only for high friction coefficients ($f > 1$). Notice the imaginary part $\tilde{b}$ represents the speed at which the disturbance moves relative to the half-plane. In particular, the disturbance moves with the sliding direction, and for low ν its speed is always larger than the speed of the shear elastic waves c_2 . At high ν and f , the disturbance moves at speed lower than c_2 .

3.4 Frictional Heating: Quasi-static Solution (TEI). We now include frictional heating, but neglect inertia terms. Hence,

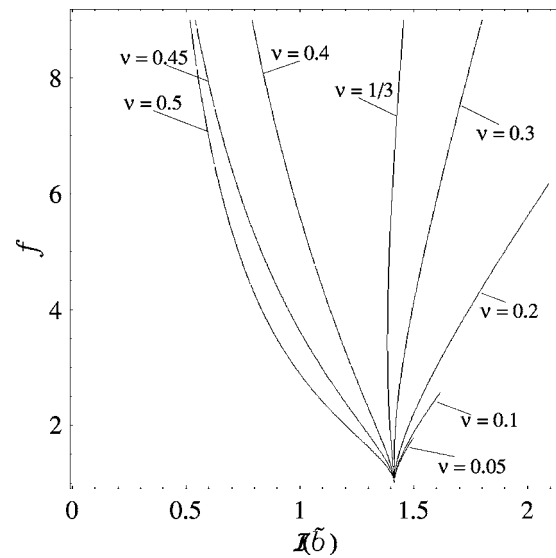


Fig. 3 Some example results for Dynamic Instability (DI): variation of the imaginary part of the dimensionless growth rate $\tilde{b}$ with the friction coefficient f for different ν

the classical TEI case is obtained. The instantaneous sliding velocity is needed to compute the amount of frictional heating. This is affected by the perturbation in the displacements, as

$$V = V_0 - \frac{\partial u_x}{\partial t}(x,0,t) \quad (47)$$

If quasi-static conditions are assumed, only thermal perturbations need to be considered. In this case, the boundary conditions are

$$u_y(x,0,t) = 0 \quad (48a)$$

$$\sigma_{xy}(x,0,t) - f\sigma_{yy}(x,0,t) = 0 \quad (48b)$$

$$-K \frac{\partial T}{\partial y}(x,0,t) + V_0 \sigma_{xy}(x,0,t) = 0 \quad (48c)$$

where the last equation gives the thermal balance between the frictional heating and the heat flow entering the half-space. Further, we are assuming that a dynamic wave instantaneously propagates, i.e., $c_1, c_2 \rightarrow \infty$, and only the eigenvalue s_3 needs to be considered. In this case Eqs. (48) give the well-known characteristic equation (see Dow and Burton [25])

$$K\omega \sqrt{1 + \frac{b}{\omega^2 k}} = \frac{2\alpha(1+\nu)\mu f V_0}{(1-\nu)\left(1 + \sqrt{1 + \frac{b}{\omega^2 k}}\right)} \quad (49)$$

and the stability boundary corresponds to the passage of a real zero through the origin $b=0$ when

$$V = V_{cr} = \frac{K\omega(1-\nu)}{f\mu\alpha(1+\nu)} \quad (50)$$

Figure 4 shows some example results. In particular, the dependence of the real part of the dimensionless growth rate $\tilde{b} = bk/\omega^2$ on the dimensionless wave number $\gamma = \omega k/c_1$ is shown for different friction coefficients f . Notice this plot is similar to Fig. 8 of Azarkhin and Barber [33], except that here a different dimensionless form of parameters is proposed in order to facilitate the comparison with the later cases of TEDI.

The domain with negative $\text{Re}(\hat{b})$ means that the waves with corresponding frequency γ decay. In addition the perturbation with largest growth rate eventually dominates the process.

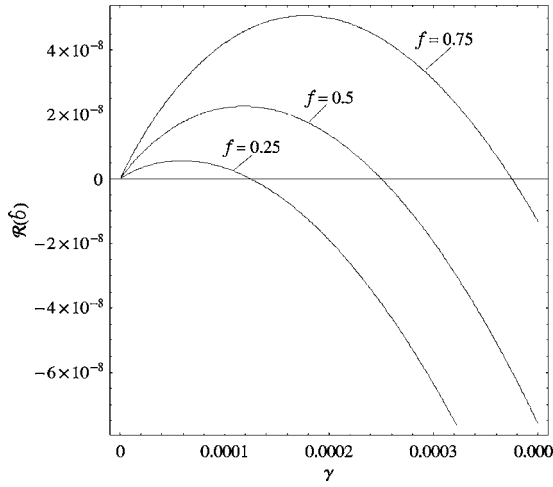


Fig. 4 Some example results for the TEI: variation of the real part of the dimensionless growth rate $\hat{b} = bk/\omega^2$ with the dimensionless wave number $\gamma = \omega k/c_1$ for different friction coefficient f ($V_0 = 10^{-3}$, $H = 1$).

4 New Results: Frictional Sliding With Heat Generation (TEDI)

If we include inertia terms and boundary condition (48c) modifies to

$$-K \frac{\partial T}{\partial y}(x, 0, t) + V \sigma_{xy}(x, 0, t) = 0 \quad (51)$$

Equation (51) is non-linear, since it contains the perturbation both in V and in σ_{xy} . We therefore need to perform a linear perturbation about the uniform solution in which

$$V = V_0 \quad \sigma_{xy}(x, 0, t) = -f p_0 \quad (52)$$

and p_0 is the (uniform) unperturbed contact pressure. We then have

$$V \sigma'_{xy}(x, 0, t) = \left[V_0 - \frac{\partial u_x}{\partial t}(x, 0, t) \right] (\sigma_{xy}(x, 0, t) - f p_0) \quad (53)$$

where σ'_{xy} is the total shear stress (sum of the unperturbed stress and the perturbation). Dropping the second-order term in this product, we have

$$V \sigma'_{xy}(x, 0, t) = -f V p_0 + V_0 \sigma_{xy}(x, 0, t) + f p_0 \frac{\partial u_x}{\partial t}(x, 0, t) \quad (54)$$

and hence the linearized form of the perturbation equation (48c) is

$$-K \frac{\partial T}{\partial y}(x, 0, t) + V_0 \sigma_{xy}(x, 0, t) + f p_0 \frac{\partial u_x}{\partial t}(x, 0, t) = 0 \quad (55)$$

In the full coupled case (when thermal and dynamic effects are considered), replacing the constants A_i , B_i , C_i using Eqs. (22), (24), and (26), the boundary conditions (48) give

$$\iota s_1 K_1 + \omega K_2 - s_3 m K_3 = 0 \quad (56)$$

$$\begin{aligned} & (-c_2^2 s_1 \omega - \iota f (c_1^2 - 2c_2^2) \omega^2 + \iota f c_1^2 s_1^2) K_1 + (\iota c_2^2 s_2^2 + 2f c_2^2 \omega s_2) K_2 \\ & + \left[-\iota c_2^2 \omega s_3 + f (c_1^2 - 2c_2^2) \omega^2 - f c_1^2 s_3^2 + f b \left(\frac{c_1^2}{k} - b \right) \right] m K_3 = 0 \end{aligned} \quad (57)$$

$$\begin{aligned} & \left(-V_0 c_2^2 s_1 \omega + f \frac{p_0}{\mu} c_2^2 b \omega \right) K_1 + \left(\iota V_0 c_2^2 s_2^2 - \iota f \frac{p_0}{\mu} c_2^2 b s_2 \right) K_2 \\ & + \left[\frac{K s_3 b}{m} \left(\frac{c_1^2}{k} - b \right) - \iota V_0 c_2^2 \omega s_3 + \iota f \frac{p_0}{\mu} c_2^2 b \omega \right] m K_3 = 0 \end{aligned} \quad (58)$$

For a non-trivial solution, we require the determinant of the coefficient matrix to be zero. Therefore, after algebraic manipulations, we obtain the following characteristic equation

$$\begin{aligned} & f V_0 m c_2^2 (\omega^2 - s_2^2) [2c_2^2 \omega^2 + c_1^2 (s_1^2 - \omega^2)] + f \frac{p_0}{\mu} m c_2^2 b \{ f c_1^2 (s_1^2 - \omega^2) s_2 \\ & + \iota c_2^2 \omega (\omega^2 - s_2^2) \} - K s_3 c_1^2 (s_1 + s_3) \{ c_2^2 (\omega^2 - s_2^2) s_1 \\ & + \iota f \omega [2c_2^2 (s_1 s_2 - \omega^2) + c_1^2 (\omega^2 - s_1^2)] \} = 0 \end{aligned} \quad (59)$$

For constant pressure p_0 and speed V_0 , the number of parameters in Eq. (59) can be reduced. In fact, if we define the dimensionless growth rate as follows

$$z = \frac{bk}{c_1^2} \quad (60)$$

the terms s_1 , s_2 , and s_3 can be rewritten as

$$s_1 = \omega \sqrt{1 + \frac{z^2}{\gamma^2}} \quad s_2 = \omega \sqrt{1 + \frac{\hat{c}^2 z^2}{\gamma^2}} \quad s_3 = \omega \sqrt{1 + \frac{z}{\gamma^2}} \quad (61)$$

where

$$\gamma = \frac{\omega k}{c_1} \quad \text{and} \quad \hat{c} = \frac{c_1}{c_2} = \sqrt{\frac{2(1-\nu)}{1-2\nu}} \quad (62)$$

With the above positions, the characteristic Eq. (59) becomes

$$\begin{aligned} & f H \frac{\hat{c}^2 z^2}{2 \gamma^3} \left[\frac{z}{\gamma} \left(f \sqrt{1 + \frac{\hat{c}^2 z^2}{\gamma^2}} - \iota \right) \hat{p}_0 - \left(2 + \frac{\hat{c}^2 z^2}{\gamma^2} \right) \hat{V}_0 \right] \\ & + \frac{\hat{c}^2 z^2}{\gamma^2} \left[\left(1 + \frac{z}{\gamma^2} \right) \sqrt{1 + \frac{z^2}{\gamma^2}} + \left(1 + \frac{z^2}{\gamma^2} \right) \sqrt{1 + \frac{z}{\gamma^2}} \right] \\ & + \iota f \left[\left(2 + \frac{\hat{c}^2 z^2}{\gamma^2} \right) \left(1 + \frac{z}{\gamma^2} + \sqrt{1 + \frac{z}{\gamma^2}} \sqrt{1 + \frac{z^2}{\gamma^2}} \right) \right. \\ & \left. - 2 \left(1 + \frac{z^2}{\gamma^2} \right) \sqrt{1 + \frac{z}{\gamma^2}} \sqrt{1 + \frac{\hat{c}^2 z^2}{\gamma^2}} \right. \\ & \left. - 2 \left(1 + \frac{z}{\gamma^2} \right) \sqrt{1 + \frac{z^2}{\gamma^2}} \sqrt{1 + \frac{\hat{c}^2 z^2}{\gamma^2}} \right] = 0 \end{aligned} \quad (63)$$

where we have introduced the following dimensionless parameters

$$\hat{V}_0 = \frac{V_0}{c_1}; \quad \hat{p}_0 = \frac{p_0}{\mu}; \quad H = \frac{2\alpha(1+\nu)\mu k}{(1-\nu)K} \quad (64)$$

Notice that practical values of $\hat{V}_0$ and $\hat{p}_0$ must be very small compared with unity, whereas typical values of H are around unity.

We study the irrational characteristic Eq. (63) by searching eigenvalues z such that

$$\text{Re}[s_i] > 0 \quad i = 1, 2, 3 \quad (65)$$

In order to do this, we transform (63) in a polynomial equation in z by squaring it six times.

In this way we are able to obtain a polynomial equation of 36th order in z . However, the only valid solutions will be those satisfying the original problem (63) and (65).

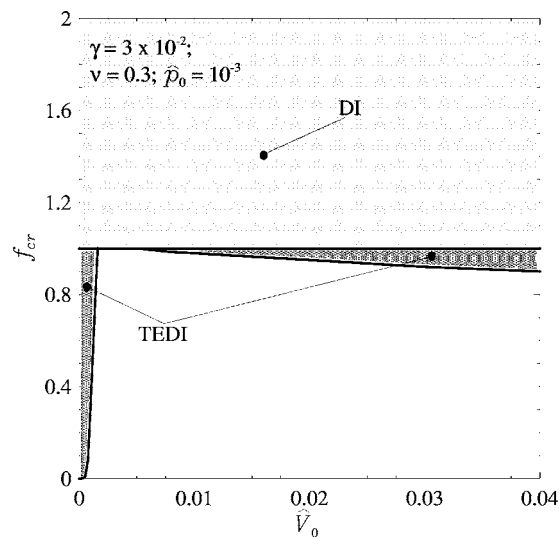


Fig. 5 Map of instability in terms of $\hat{V}_0$ for $\nu=0.3$, $\hat{p}_0=10^{-3}$ and $\gamma=3 \times 10^{-2}$

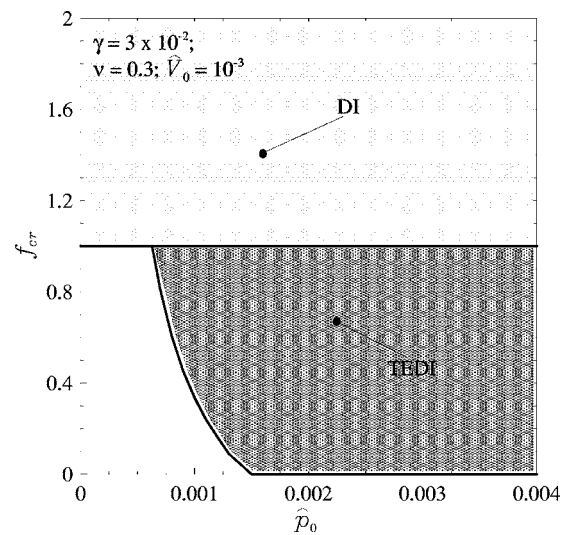


Fig. 7 Map of instability in terms of $\hat{p}_0$ for $\nu=0.3$, $\hat{V}_0=10^{-3}$ and $\gamma=3 \times 10^{-2}$

5 Results

The roots of the polynomial equation obtained from the transformation of (63) are found by using the Jenkins-Traub algorithm ([34,35]). Alternatively, and as a further check, we also used a direct iterative complex root finder, where the choice of the starting solution was, however, very critical (generally the solutions were mapped upon gradual increase of a given parameter, using a starting point was obtained from a previous case).

5.1 TEDI Maps. In this section, we present various maps of instability where the critical friction coefficient for TEDI ($f_{cr,TEDI}$), DI ($f_{cr,DI}$), and TEI ($f_{cr,TEI}$) is defined in terms of operating parameters $\hat{V}_0$ and $\hat{p}_0$.

5.1.1 Effect of the Sliding Speed $\hat{V}_0$. In Figs. 5 and 6 the maps of instability are presented in terms of $\hat{V}_0$ for different wave numbers γ . At high wave number ($\gamma > 2.9 \times 10^{-2}$, Fig. 5) instability occurs for all friction coefficients if $\hat{V}_0$ is low and, perhaps sur-

prisingly, only at high f for larger $\hat{V}_0$.

When lower values of γ are considered, the map of instability modifies as shown in Fig. 6. In particular, the region of instability enlarges and the instability zone governed by TEI eigenvalues increases. However, TEDI continues to be the only possible mechanism of instability at low speeds.

Finally, notice when $\gamma > 2.9 \times 10^{-2}$ (see Fig. 5) it is possible to define a speed range in which instability occurs only for $f > 1$, whereas when $\gamma < 2.9 \times 10^{-2}$ (see Fig. 6), this range reduces to a single value of $\hat{V}_0$.

5.1.2 Effect of the Pressure $\hat{p}_0$. Figures 7 and 8 show the maps of instability in terms of $\hat{p}_0$ for the same γ considered in Figs. 5 and 6, respectively.

Again, friction instability is governed by DI eigenvalues at high friction coefficient ($f > 1$). When $\gamma > 2.9 \times 10^{-2}$ (Fig. 7), the critical friction coefficient f_{cr} is less than one only for high pressure

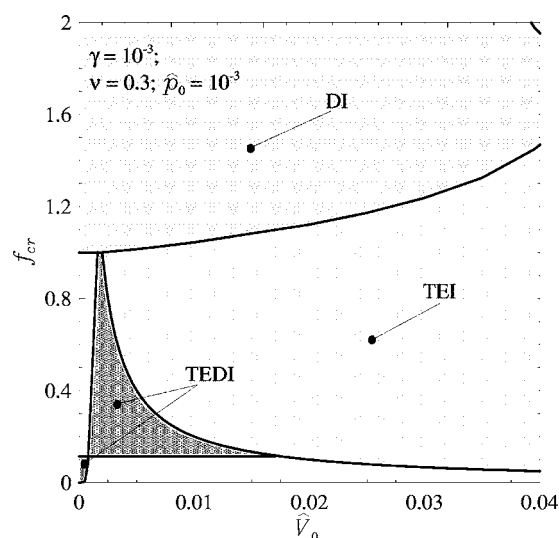


Fig. 6 Map of instability in terms of $\hat{V}_0$ for $\nu=0.3$, $\hat{p}_0=10^{-3}$ and $\gamma=10^{-3}$

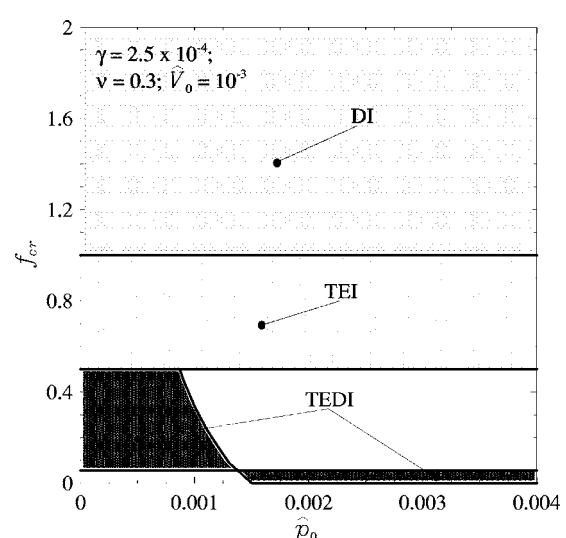


Fig. 8 Map of instability in terms of $\hat{p}_0$ for $\nu=0.3$, $\hat{V}_0=10^{-3}$ and $\gamma=2.5 \times 10^{-4}$

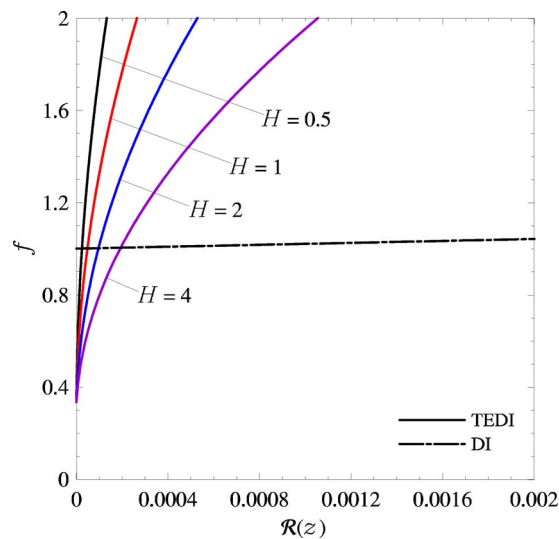


Fig. 9 Variation of the exponential growth rate $\Re(z)$ with the friction coefficient f for different value of H , $\nu=0.3$, $\hat{p}_0=\hat{V}_0=10^{-3}$, $\gamma=2$

and $f_{cr}=0$ for $\hat{p}_0 \geq 1.5 \times 10^{-3}$. In Fig. 8 the TEI critical friction coefficient is less than 1 and quasi-static eigenvalues dominate the system response for $f_{cr,TEI} < f < f_{cr,DI} = 1$.

5.2 TEDI growth rates. In this section we present results for the exponential growth rate $\Re(z)$.

5.2.1 Effect of the Parameter H . By examining equation (63), we notice the frictionless dynamic solution is recovered for $H=0$. Increasing H from zero corresponds to activating thermal effects, and new families of unstable roots appear.

In Fig. 9, the variation of the exponential growth rate $\Re(z)$ with the friction coefficient f is shown for different values of H . Dynamic solutions are plotted with dashed-dotted line (we have instability only for very high friction coefficient $f > 1$ and hardly any further change occurs in the plot when we increase thermal effects).

The classical quasi-static (TEI) solutions give negative growth rates for the parameters of Fig. 9 and, hence, are not plotted here.

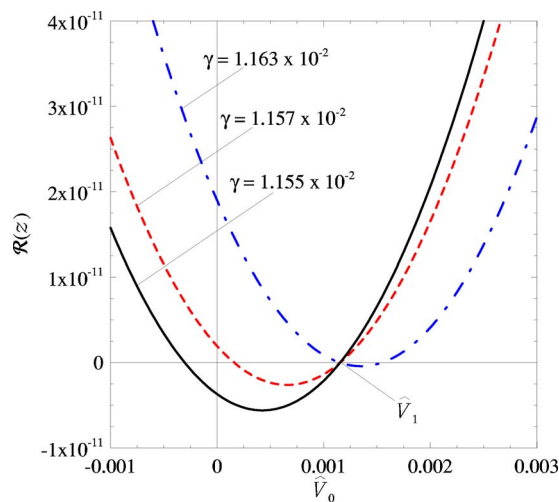


Fig. 10 Variation of the exponential growth rate $\Re(z)$ with the sliding speed $\hat{V}_0$ for different wave numbers γ and $H=1$, $f=0.5$, $\nu=0.3$, $\hat{p}_0=10^{-3}$

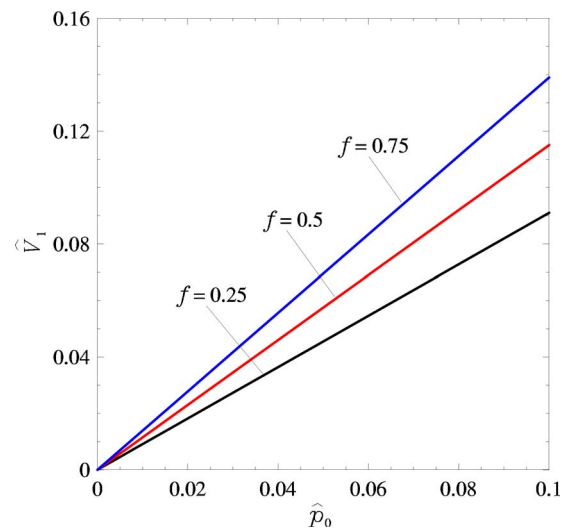


Fig. 11 Variation of the characteristic speed $\hat{V}_1$ with the pressure $\hat{p}_0$ for different friction coefficients ($\nu=0.3$)

When thermal effects, even if very small, are added, a new family of unstable roots appear (named TEDI here), and the system can be unstable even for friction coefficients less than 1. In particular, a critical friction coefficient exists and it is a function of the Poisson's ratio, speed, and pressure, as shown in the above maps. However, H affects the growth rate but not the stability boundary. In other words, f_{cr} is independent of H .

In addition, the growth factor is generally much smaller than the frequency of oscillation (i.e., the real part of the eigenvalue is much smaller than the imaginary part), and hence we expect that its effect can only be seen if sliding continues for a long term. Physically, this is to give time essentially for the frictional heating to produce significant departure from the quasi-static solution.

5.2.2 Effect of the Dimensionless Wave Number γ . Figure 10 shows the variation of the exponential TEDI growth rate $\Re(z)$ with the sliding speed $\hat{V}_0$ for different wave numbers γ . Notice we can define a characteristic speed $\hat{V}_1$ and a particular value of $\gamma = \gamma_{cr}$ (corresponding to the curve with minimum in $\hat{V}_1$) for which we have unconditional instability, i.e., the system is unstable for all speeds.

The characteristic speed $\hat{V}_1$ linearly increases with $\hat{p}_0$ and it does not depend on material parameter H and dimensionless wave number γ , as shown in Fig. 11.

Figure 12 shows the exponential growth rate $\Re(z)$ as a function of the dimensionless wave number γ for different friction coefficients and $\hat{V}_0 < \hat{V}_1$. TEDI growth rates are plotted with solid line, TEI ones with a dashed line, and DI ones with a dashed-dotted line. For small γ , a zoom of the plot was done to better show the behavior of the various eigenvalues. In fact, the regions of $\gamma > 1$ shown in the original figure, are limited to the case of nanometer wavelengths for most materials. Hence, we shall concentrate in the region of the zoomed figure on the left, for small $\gamma < 1$.

Here we find that, when $f < 1$, we do not have any DI, and we therefore find only TEDI and TEI solutions,² which are of comparable size in the region of the zoom. Clearly, while the TEI growth rates are positive only for very small wave numbers, they become negative for γ larger than a certain limit value γ_{lim} . Further, TEI eigenvalues show a maximum at $\gamma = \gamma_{lim}/2$, whereas TEDI eigenvalues do not show a maximum, and continue to in-

²These curves remind the single curve in the dimensionless plot of Azarkhin and Barber [33].

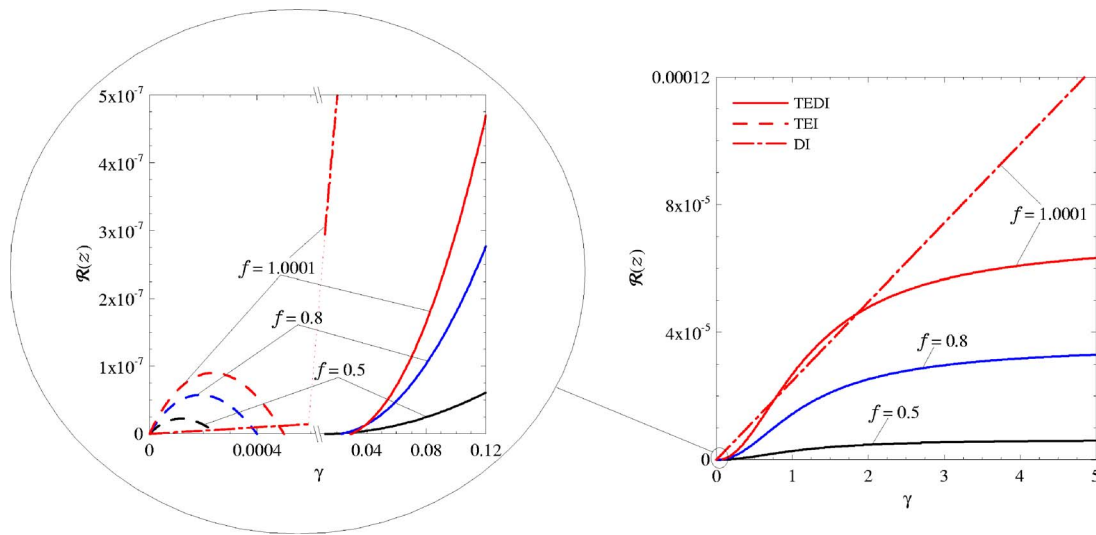


Fig. 12 Variation of the exponential growth rate $\Re(z)$ with γ for different values of the friction coefficient f , $\hat{\rho}_0=10^{-3}$, $\hat{V}_0=10^{-3} < \hat{V}_1$, $H=1$ and $\nu=0.3$

crease for larger and larger wave numbers, i.e., smaller and smaller wavelengths. TEDI eigenvalues increase monotonically with wave number and they tend to an asymptotic value, for which the smaller wavelengths (i.e., larger wave numbers) are most unstable. It is possible, however, that, as in the case of DI roots (Ranjith and Rice [18]), these high growth factors render the response of a material interface with Coulomb friction ill-posed. Physically, this implies that during sliding, energy is transferred to shorter wavelengths, leading to pulse sharpening and splitting, and creates an inherent grid-size dependence. Ranjith and Rice [18] showed that an experimentally based rate- and state-dependent friction law, in which the shear strength in response to an abrupt change in normal stress evolves continuously with time, regularizes the problem.

Turning back to the TEDI growth rates, they are positive only when γ is greater than a certain value γ_{cr} . This value γ_{cr} increases with the friction coefficient moving, for example, from 1.161×10^{-2} for $f=0.5$ to 2.9×10^{-2} for $f=0.8$. Figures 13 and 14 show that γ_{cr} is an increasing linear function of $\hat{V}_0$ and decreasing of $\hat{\rho}_0$, respectively. In particular, for low pressure instability is for $\gamma < \gamma_{cr}$. In fact, the characteristic speed $\hat{V}_1$ reduces by decreasing $\hat{\rho}_0$ and, as shown in Fig. 15, for $\hat{V}_0 > \hat{V}_1$ the TEDI eigenvalues have a positive real part when $\gamma < \gamma_{cr}$. In addition, the effect of H on γ_{cr} is negligible.

Finally, Fig. 15 shows the variation of the exponential growth rate $\Re(z)$ with γ for different f and $\hat{V}_0 > \hat{V}_1$.

Here, we find, again, that TEI roots dominate the system behavior at low γ , but now TEDI eigenvalues show a maximum and for high wave numbers (small wavelengths) the system is stable. However, the growth factors seem to be three orders of magnitude smaller than the TEI eigenvalues (see the cut in the scales of the figure), and hence it is unlikely that these new roots at high speed have a significant effect.

6 Conclusions

In the classical dynamic instability (DI) in the simplest case of a single half-plane sliding against a rigid non-conductor (studied by Martins et al. [14]), instability occurs at sufficiently high friction f , in which case all wavelengths are made unstable. Vice versa, when inertia effects are neglected and frictional heating is considered, TEI instability occurs at all wavelengths bigger than a certain value, depending on the sliding speed. The coupling between dynamic and thermoelastic effects also makes unstable a

new family of solutions corresponding to frictionless contact interface waves, i.e., the dilatational waves propagating parallel to the interface, which acquire a positive growth factor with thermal effects.

Various maps of instability are presented in terms of material properties (Poisson's ratio ν) and operating parameters (sliding speed and applied pressure) for different ranges of the wave number of the perturbation.

At high wave numbers (small wavelengths) TEI disappears and instability will be governed by DI at high friction coefficients, by TEDI at low friction. At low wave numbers (large wavelengths), the three forms of instability are possible. In particular, when the sliding speed is very high, the mechanism of instability will be dominated by TEI also at high friction coefficients. At low speeds, DI is the dominant mechanism when $f > 1$, whereas TEDI occurs for $f < 1$. The critical friction coefficient for TEDI is zero for low

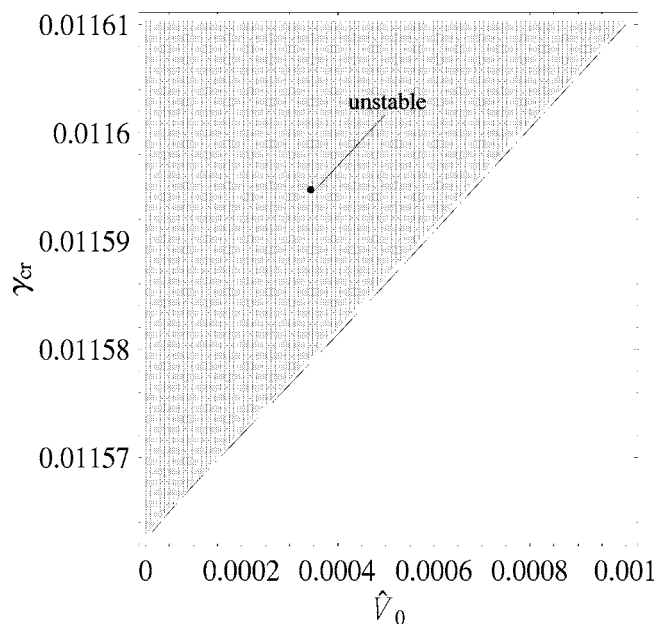


Fig. 13 Dependence of γ_{cr} on $\hat{V}_0$ for $\hat{\rho}_0=10^{-3}$, $H=1$, $f=0.5$, and $\nu=0.3$

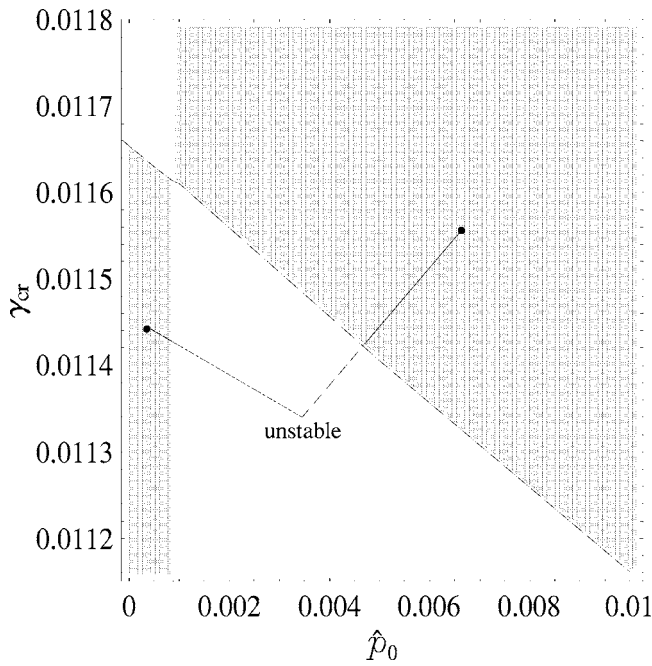


Fig. 14 Dependence of γ_{cr} on $\hat{p}_0$ for $\hat{V}_0=10^{-3}$, $H=1$, $f=0.5$, and $\nu=0.3$

Poisson ratios ($\nu \leq 0.2$) and it increases with ν . In particular, for high Poisson ratios, the system is always stable when $f < 1$.

New TEDI eigenvalues are found at low speeds, similar to the DI ones and in particular, they seem to be ill-conditioned in the sense that they give rise to unbounded growth factors at large wave numbers. A regularized friction law is required to make sensible predictions, like the Rice-Ruina law. Vice versa, at large speeds, the instability roots are more similar to the TEI roots, and hence the regularized law is not needed. However, in this range the growth factors seem to be particularly small.

Acknowledgment

Useful discussions with Prof. J. R. Barber and G. Demelio are acknowledged.

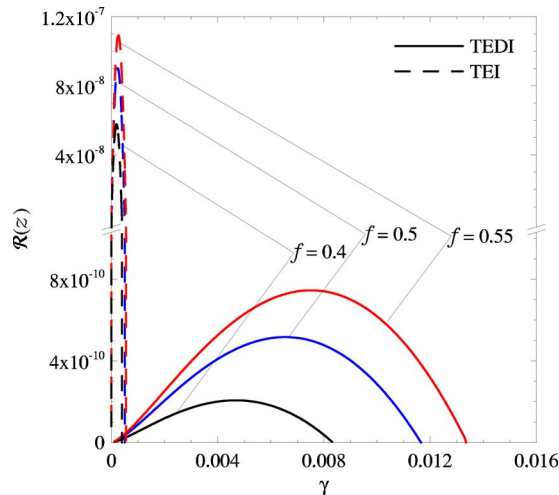


Fig. 15 Variation of the exponential growth rate $\Re(z)$ with γ for different values of the friction coefficient f , $\hat{p}_0=10^{-3}$, $\hat{V}_0=2 \times 10^{-3} > \hat{V}_1$, $H=1$ and $\nu=0.3$

Nomenclature

- b = growth rate (s^{-1})
- c_1 = dilatational wave speed (m/s)
- c_2 = shear wave speed (m/s)
- c_E = specific heat ($J/kg^\circ C$)
- c_R = Rayleigh wave speed (m/s)
- E = Young's modulus (N/m^2)
- f = friction coefficient
- k = diffusivity (m^2/s)
- K = thermal conductivity ($W/m^\circ C$)
- ω = wave number (m^{-1})
- p = contact pressure (N/m^2)
- p_0 = unperturbed contact pressure (N/m^2)
- q = heat flux per unit area (W/m^2)
- t = time (s)
- T = temperature field ($^\circ C$)
- u = displacement field (m)
- V_0 = sliding speed (m/s)
- V = instantaneous sliding speed (m/s)
- α = thermal expansion coefficient ($^\circ C^{-1}$)
- ι = imaginary unit
- λ = Lamé's constant (N/m^2)
- μ = shear modulus (N/m^2)
- ν = Poisson's coefficient
- ρ = density (Kg/m^3)
- σ = stress field (N/m^2)

References

- [1] Dowson, D., 1998, *History of Tribology*, 2nd ed., John Wiley & Sons, New York.
- [2] Duffour, P., and Woodhouse, J., 2004, "Instability of Systems With a Frictional Point Contact. Part 1: Basic Modelling," *J. Sound Vib.*, **271**, pp. 365–390.
- [3] Ben-Zion, Y., 2001, "Dynamic Ruptures in Recent Models of Earthquake Faults," *J. Mech. Phys. Solids*, **49**, pp. 2209–2244.
- [4] Rice, J. R., Lapusta, N., and Ranjith, K., 2001, "Rate and State Dependent Friction and the Stability of Sliding Between Elastically Deformable Solids," *J. Mech. Phys. Solids*, **49**(9), pp. 1865–1898.
- [5] Barquins, M., Koudine, A. A., and Vallet, D., 1996, "On the Existence of Schallamach Waves During Surface Friction in Rubber-Like Material as a Function of Thickness," *C. R. Acad. Sci., Ser. IIb Mec.*, **323**, pp. 433–438.
- [6] Kinkaid, N. M., O'Reilly, O. M., and Papadopoulos, P., 2003, "Automotive Disc Brake Squeal," *J. Sound Vib.*, **267**, pp. 105–166.
- [7] Kennedy, F. E., and Ling, F. F., 1974, "A Thermal, Thermoelastic and Wear Simulation of a High Energy Sliding Contact Problem," *ASME J. Lubr. Technol.*, **96**, pp. 497–507.
- [8] Anderson, A. E., and Knapp, R. A., 1989, "Hot Spotting in Automotive Friction Systems," *Intl. Conf. on Wear of Materials*, Denver, CO, April 9–13, Vol. 2, pp. 673–680.
- [9] Popp, K., and Rudolph, M., 2004, "Vibration Control to Avoid Stick-Slip Motion," *J. Vib. Control*, **10**, pp. 1585–1600.
- [10] Johnson, K. L., 2001, "Dynamic Friction," in *Tribology Research: From Model Experiment to Industrial Problem*, G. Dalmaz, A. A. Lubrecht, D. Dowson, and M. Priest, eds., Elsevier, New York.
- [11] Stoneley, R., 1924, "Elastic Waves at the Surface of Separation of Two Solids," *Proc. R. Soc. London*, **A106**, pp. 416–428.
- [12] Achenbach, J. D., and Epstein, H. I., 1967, "Dynamic Interaction of a Layer and a Half-Space," *J. Engrg. Mech. Div.*, **EM5**, pp. 27–42.
- [13] Murty, G. S., 1975, "Wave Propagation at an Unbounded Interface Between Two Elastic Half-Spaces," *J. Acoust. Soc. Am.*, **58**, pp. 1094–1095.
- [14] Martins, J. A. C., Guimaraes, J., and Faria, L. O., 1995, "Dynamic Surface Solutions in Linear Elasticity and Viscoelasticity With Frictional Boundary Conditions," *J. Vib. Acoust.*, **117**, pp. 445–451.
- [15] Adams, G. G., 1995, "Self-Excited Oscillations of Two Elastic Half-Spaces Sliding With a Constant Coefficient of Friction," *ASME J. Appl. Mech.*, **62**, pp. 867–872.
- [16] Adams, G. G., 1998, "Dynamic Instabilities in the Sliding of Two Layered Elastic Half-Spaces," *ASME J. Tribol.*, **120**(2), pp. 289–295.
- [17] Ranjith, K., and Rice, J. R., 1999, "Stability of Quasi-Static Slip in a Single Degree of Freedom Elastic System With Rate and State Dependent Friction," *J. Mech. Phys. Solids*, **47**(6), pp. 1207–1218.
- [18] Ranjith, K., Rice, J. R., 2001, "Slip Dynamics at an Interface Between Dissimilar Materials," *J. Mech. Phys. Solids*, **49**(2), pp. 341–361.
- [19] Barber, J. R., Dundurs, J., and Comninou, M., 1980, "Stability Consideration in Thermoelastic Contact," *ASME J. Appl. Mech.*, **47**, pp. 871–874.
- [20] Clausing, A. M., and Chao, B. T., 1965, "Thermal Contact Resistance in a Vacuum Environment," *ASME J. Heat Transfer*, **87**, pp. 243–251.

- [21] Richmond, O., and Huang, N. C., 1977, "Interface Stability During Unidirectional Solidification of a Pure Metal," *Proc. 6th Canadian Congress of Applied Mechanics*, Vancouver, pp. 453–454.
- [22] Yigit, F., and Barber, J. R., 1994, "Effect of Stefan Number on Thermoelastic Instability in Unidirectional Solidification," *Int. J. Mech. Sci.*, **36**, pp. 707–723.
- [23] Yigit, F., 1998, "Effect of Mold Properties of Thermo-Elastic Instability in Unidirectional Planar Solidification," *J. Therm. Stresses*, **21**, pp. 55–81.
- [24] Barber, J. R., 1969, "Thermoelastic Instabilities in the Sliding of Conforming Solids," *Proc. R. Soc. London, Ser. A*, **312**, pp. 381–394.
- [25] Dow, T. A., and Burton, R. A., 1972, "Thermoelastic Instability of Sliding Contact in the Absence of Wear," *Wear*, **19**, pp. 315–328.
- [26] Afferrante, L., Ciavarella, M., and Barber, J. R., 2007, "Sliding Thermoelastic Instability," *Proc. R. Soc. London, Ser. A*, **462**, pp. 2161–2176.
- [27] Afferrante, L., and Ciavarella, M., 2007, "A Note on ThermoElastoDynamic Instability (TEDI) for a 1D Elastic Layer: Force Control," *Int. J. Solids Struct.*, **44**, pp. 1380–1390.
- [28] Burton, R. A., Nerlikar, V., and Kilaparti, S. R., 1973, "Thermoelastic Instability in a Seal-Like Configuration," *Wear*, **24**, pp. 177–188.
- [29] Boley, B. A., and Weiner, J. H., 1985, *Theory of Thermal Stresses*, Krieger, Malabar, FL.
- [30] Eringen, A. C., 1967, *Mechanics of Continua*, Wiley, New York.
- [31] Chadwick, P., 1960, "Thermoelasticity: The Dynamical Theory," in *Progress in Solids Mechanics* Vol. 1 I. N. Sneddon and R. Hill eds., North Holland, Amsterdam.
- [32] Nowacki, W., 1962, *Thermoelasticity*, Addison-Wesley, Reading, MA.
- [33] Azarkhin, A., and Barber, J. R., 1985, "Transient Thermoelastic Contact Problem of Two Sliding Half-Planes," *Wear*, **102**, pp. 1–13.
- [34] Jenkins, M. A., and Traub, J. F., 1969/1970, "A Three-Stage Variable-Shift Iteration for Polynomial Zeros and its Relation to Generalized Rayleigh Iteration," *Numer. Math.*, **14**, pp. 252–263.
- [35] Jenkins, M. A., and Traub, J. F., 1970, "A Three-Stage Algorithm for Real Polynomials Using Quadratic Iteration," *SIAM (Soc. Ind. Appl. Math.) J. Numer. Anal.*, **7**, pp. 545–566.

The Girsanov Linearization Method for Stochastically Driven Nonlinear Oscillators

Nilanjan Saha

D. Roy¹

e-mail: royd@civil.iisc.ernet.in

Department of Civil Engineering,
Indian Institute of Science,
Bangalore, Karnataka 560 012, India

For most practical purposes, the focus is often on obtaining statistical moments of the response of stochastically driven oscillators than on the determination of pathwise response histories. In the absence of analytical solutions of most nonlinear and higher-dimensional systems, Monte Carlo simulations with the aid of direct numerical integration remain the only viable route to estimate the statistical moments. Unfortunately, unlike the case of deterministic oscillators, available numerical integration schemes for stochastically driven oscillators have significantly poorer numerical accuracy. These schemes are generally derived through stochastic Taylor expansions and the limited accuracy results from difficulties in evaluating the multiple stochastic integrals. As a numerically superior and semi-analytic alternative, a weak linearization technique based on Girsanov transformation of probability measures is proposed for nonlinear oscillators driven by additive white-noise processes. The nonlinear part of the drift vector is appropriately decomposed and replaced, resulting in an exactly solvable linear system. The error in replacing the nonlinear terms is then corrected through the Radon-Nikodym derivative following a Girsanov transformation of probability measures. Since the Radon-Nikodym derivative is expressible in terms of a stochastic exponential of the linearized solution and computable with high accuracy, one can potentially achieve a remarkably high numerical accuracy. Although the Girsanov linearization method is applicable to a large class of oscillators, including those with nondifferentiable vector fields, the method is presently illustrated through applications to a few single- and multi-degree-of-freedom oscillators with polynomial nonlinearity. [DOI: 10.1115/1.2712234]

Keywords: Girsanov transformation, Ito-Taylor expansion, multiple stochastic integrals, Monte Carlo simulation, Wiener processes, nonlinear oscillators

1 Introduction

The available approximate analytical methods for determination of transitional probability density or statistical moments of the response of stochastically driven nonlinear oscillators generally suffer from the dimensionality curse and/or restrictions on the type and number of stochastic excitations [1]. A much more versatile and popularly adopted route is a Monte Carlo simulation (MCS) aided by direct numerical integration of the associated stochastic differential equations (SDEs). Direct MCS, while often regarded as less elegant, can be used to solve problems of significantly higher complexity and dimensions. Even then, its implementation is hindered by the intensively repetitive computation over a possibly large ensemble and the accuracy of the direct integration scheme. Indeed, as compared to numerical integration schemes for deterministic DEs, most of the available ones for SDEs, often derived via stochastic Taylor expansions, have considerably lower orders of accuracy. This happens due to difficulties in computing the multiple stochastic integrals (MSIs) [2,3]. A few such typical schemes include the Euler-Maruyama (Maruyama [4]), stochastic Heun (Gard [5]), stochastic Runge-Kutta [6–8], and stochastic Newmark [9,10]. Moreover, there are also implicit variations of these methods [10–12] that have a higher stochastic numerical stability than their explicit counterparts. Although the implicit methods work well for even stiff SDEs, solutions tend to become inexact for very low step sizes [10]. These stochastic integration schemes may also be catego-

rized as strong and weak. Although strong schemes are designed to simulate the pathwise unique response corresponding to a particular realization of the stochastic excitation [13], weak schemes only compute statistical moments of functions of the response by replacing the MSIs with random variables having simpler distributions [3]. Accordingly, a weak algorithm is computationally faster and the preferred choice in most engineering applications, wherein it suffices to obtain approximations to certain statistical moments or distribution functions. Nevertheless, among all these integration schemes—weak or strong, implicit or explicit—the ones with the highest accuracy are the stochastic Runge-Kutta (an explicit method of weak order 2 [8]) and the stochastic Newmark (an implicit method of weak order 3 [10]).

Among the analytical or numeric-analytic techniques, linearization methods (the equivalent linearization method in particular) have found a certain measure of acceptability. Recently, Socha [14,15] has provided a review of equivalent and statistical linearization in the analysis of nonlinear oscillators. Equivalent linearization methods essentially try to replace the nonlinear function by an equivalent linear one such that the error in the replacement is minimized in a sense. The approximating linear function may be obtained through several criteria, e.g., mean square approximation [16,17], errors in response moments [18], energy [19], frequency [20], probability density [21], and an improved version of Gaussian equivalent linearization [22]. These methods generally use iterative procedures for the determination of the equivalent linear function and, subsequently, response statistics. Such statistical linearization methods have been applied to such applications as filtering, control [23] and computation of response statistics of hysteretic oscillators [24]. It is however known that the class of global linearization methods, although very efficient when it works, is unacceptably inaccurate for most nonlinear single- and

¹Corresponding author.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received April 10, 2006; final manuscript received November 3, 2006. Review conducted by N. Sri Namachchivaya.

multi-degree-of-freedom (SDOF and MDOF) systems. Other than globally defined equivalent linearization schemes, mention may also be made of a few other forms of stochastic linearizations. These include conditional linearization [25] and local forms of linearizations, viz. the phase space linearization [26] and the locally transversal linearization [27,28]. Although conditional linearization lacks universal applicability, the others are limited in numerical accuracy. Indeed, the observation on the lack of general applicability is valid for several other analytical tools, such as the method of moment closures [29], stochastic averaging [30], maximum entropy [31] etc. Moreover, the quality of approximation via all these methods in the general context of nonlinear oscillators is very often poor.

A weak and numeric-analytic method of stochastic linearization, referred to as the Girsanov linearization method (GLM), is presently proposed for nonlinear oscillators under additive stochastic excitations. The essence of the GLM is an appropriate linearized replacement of the nonlinear drift terms followed by an appeal to the Girsanov transformation of probability measures to weakly correct for the error caused due to this replacement. This correction is multiplicative in nature, satisfies a scalar SDE that is "exactly" solvable in terms of the linearized solution and is referred to as the Radon-Nikodym derivative (also called the likelihood ratio). In particular, the solution for the likelihood ratio is in the form of a stochastic exponential whose argument is in the form of integrals of the linearized solution. However, the numerical computation of the stochastic exponential is generally prone to considerable numerical instability and forms the most crucial part of the development of the GLM. The present framework developed to arrest the numerical instability requires, as a first step, that the time interval of interest be discretized into smaller subintervals and the linearization performed over each such subinterval. Two versions of the GLM then evolve depending on how the stochastic exponential is computed over each subinterval. Although version I uses a stochastic Taylor expansion for an approximate evaluation of the exponential, version II does it in such a way that the error is introduced only due to the finiteness of the ensemble size. Other than the semi-analyticity and unparalleled numerical accuracy, yet another useful feature of the GLM is its ability to treat oscillators even with nondifferentiable nonlinearity. Nevertheless, only a limited numerical illustration of the GLM is presently provided through its application to a few single- and two-degree-of-freedom (2-DOF) oscillators with polynomial nonlinearity.

2 Girsanov Linearization

As mentioned earlier, present study restricts itself to the derivation of the GLM in the context of nonlinear (mechanical) oscillators while keeping in mind that the method may be readily extended to a more general class of nonlinear SDEs. Thus consider the equations of motion of an n -DOF oscillator excited by white noise processes in the following form:

$$[M]\ddot{X} + C\dot{X} + KX + \chi(X, \dot{X}, t) = \sum_{r=1}^q G_r(t)\dot{W}_r + F(t) \quad (1)$$

where $X = \{x^{(1)}, x^{(2)}, \dots, x^{(n)}\}^T \in \mathbf{R}^n$ is the displacement vector. Equation (1) is subject to the initial condition $X(t=0) := X_0 = \{x_0^{(1)}, x_0^{(2)}, \dots, x_0^{(n)}\}^T$, $[M], [C], [K] \in \mathbf{R}^{n \times n}$ are constant mass, damping, and stiffness matrices, respectively, $\chi(X, \dot{X}, t)$ is any vector function (not necessarily smooth) that ensures a unique solution of Eq. (1) at least in the weak sense, $\{G_r(t): \mathbf{R} \rightarrow \mathbf{R}^n\}$ is the set of n diffusion vectors (additive), $\{W_r(t) | r \in [1, q]\}$ is a q -dimensional vector of independently evolving zero-mean Wiener processes with $W_r(0)=0$, $\mathbf{E}[W_r(t)-W_r(s)]^2 = (t-s)$, $t > s \forall r \in [1, q]$ and $F(t) = \{F^{(j)}(t) | j=1, \dots, n\}$ is the externally applied deterministic force vector. $\mathbf{E}$ denotes the expectation operator with respect to the underlying probability measure P . The description

of the oscillator as in Eq. (1) being entirely formal (due to non-differentiability of Wiener processes, which implies that $\dot{W}_r(t)$ exists merely as a valid measure, but not as a mathematical function), they may more appropriately be recast as a system of $2n$ first order SDEs in the following incremental form:

$$dx_1^{(j)} = x_2^{(j)} dt$$

$$dx_2^{(j)} = a^{(j)}(X, \dot{X}, t) dt + \sum_{r=1}^q b_r^{(j)}(t) dW_r(t), \quad j=1, 2, \dots, n \quad (2)$$

where

$$a^{(j)}(X, \dot{X}, t) = - \sum_{k=1}^n \hat{C}_{jk} \dot{x}^{(k)} - \sum_{k=1}^n \hat{K}_{jk} x^{(k)} - \hat{\chi}^{(j)}(X, \dot{X}, t) + \hat{F}^{(j)}(t)$$

$$[\hat{C}] = [M^{-1}][C], \quad [\hat{K}] = [M^{-1}][K],$$

$$\{\hat{F}\} = [M^{-1}]\{F\}, \quad \{\hat{\chi}\} = [M^{-1}]\{\chi\}$$

$$B_r(t) = \{b_r^{(j)}(t) | j=1, \dots, n\} = [M^{-1}]G_r(t) \quad (3)$$

Moreover, define $\bar{X}_0 = \{x_{1,0}^{(1)}, x_{2,0}^{(1)}, x_{1,0}^{(2)}, x_{2,0}^{(2)}, \dots, x_{1,0}^{(n)}, x_{2,0}^{(n)}\}^T \in \mathbf{R}^{2n}$. Now we assume that the drift vector $A = \{a^{(j)} | j=1, \dots, n\}$ (coefficient of dt term) can be decomposed into two constituent parts as $a^{(j)} = a_l^{(j)} + a_{nl}^{(j)}$ (i.e., $A = A_l + A_{nl}$), where A_l denotes the linear part of the vector field that should remain in the linearized equations following GLM and A_{nl} denotes the nonlinear part. To ensure boundedness of the solution vector $\bar{X} \triangleq \{x_1^{(1)}, x_2^{(1)}, x_1^{(2)}, x_2^{(2)}, \dots, x_1^{(n)}, x_2^{(n)}\}^T \in \mathbf{R}^{2n} \forall t$ (such that $X_1 := X = \{x_1^{(j)}\}^T \in \mathbf{R}^n$, $X_2 := \dot{X} = \{x_2^{(j)}\}^T \in \mathbf{R}^n$) and uniqueness in a weak sense, it is assumed that the drift and diffusion vectors, $A = \{a^{(j)}\}$ and $B_r = \{b_r^{(j)}\}$ are measurable and satisfy the following bounds:

$$\|A(\bar{X}, t)\| + \left\| \sum_{r=1}^q B_r(t) \right\| \leq Q_1(1 + \|\bar{X}\|) \quad (4a)$$

$$\|A(\bar{X}, t) - A(\bar{Y}, t)\| \leq Q_2 \|\bar{X} - \bar{Y}\| \quad (4b)$$

where $\bar{Y} \in \mathbf{R}^{2n}$, $Q_1, Q_2 \in \mathbf{R}^+$ and $\|\cdot\|$ denotes the Euclidean norm (i.e., the $L_2(P)$ norm in the associated probability space $[\Omega, \mathcal{F}_t, P]$). Let the initial conditions be mean square bounded, i.e., $\mathbf{E}\|\bar{X}(t_0)\|^2 < \infty$ (without a loss of generality, the initial condition vector is presently treated as deterministic). The time interval $[0, T]$ of interest is so ordered (discretized) that $0 = t_0 < t_1 < \dots < t_L = T$, $h_i = t_i - t_{i-1}$, and $T_i = (t_{i-1}, t_i]$, where $i \in \mathbf{Z}^+$. For further simplification of the rest of the presentation without a loss of focus on the main issues, we assume a uniform time step $h_i = h \forall i$ unless otherwise stated.

We are interested in computing the weak response, i.e., expectations of the form

$$v = \mathbf{E}[f(\bar{X}, t)] \quad (5)$$

where $f(\cdot)$ denotes some real-valued function. When we utilize Monte Carlo simulation to evaluate the expectation, v is replaced by its sample-mean formula [32]

$$v \approx \hat{v} = \frac{1}{N} \sum_{m=1}^N [f(\bar{X}^{(m)}, t)] \quad \text{for large } N \quad (6)$$

where $\bar{X}^{(m)}(t)$ denotes the m th sample path of $\bar{X}(t)$ and N is the number of sample paths (i.e., the ensemble size). Since it suffices to evaluate $\hat{v}$ using weak solutions of Eq. (1), we intend to use a measure transformation based on Girsanov's theorem [2] to locally linearize Eq. (1) and thus (hopefully) obtain highly accurate

estimates of the required expectation based on the known analytical solutions of the linear SDEs. By “local linearization,” we mean that the governing SDEs would be replaced by a system of linearized ones so that the i th such linearized SDE could serve as an accurate replacement for the nonlinear SDE over T_i .

Briefly, the Girsanov theorem implies that if the drift coefficient of an Ito process (with a nondegenerate diffusion coefficient) is altered to an extent (such that the original and altered drift coefficients satisfy the Novikov condition, as defined below), then the law of the process would not be altered drastically. Indeed, the law of the drift-modified process will remain absolutely continuous with respect to that of the original process and we can explicitly compute the Radon-Nikodym derivative (see the Radon-Nikodym theorem below). Before going into the formulation of GLM we state a few definitions and theorems that would be useful during the formulation.

DEFINITION (ABSOLUTE CONTINUITY OF MEASURES). Let $[\Omega, \mathcal{F}, \{F_t\}_{t \geq 0}]$ be a filtered probability space (i.e., $[\Omega, \mathcal{F}, P]$ is a probability space and $\{F_t\}_{t \geq 0}$ is a filtration on $[\Omega, P]$; presently, it is the filtration generated by the Brownian motion processes). If we fix $T > 0$ and let Q be another probability measure on $\{F_t\}_T$, then we say that Q is absolutely continuous with respect to $P|_{F_T}$ (the restriction of P to F_T) and we write $Q \ll P$ if $P(H) = 0 \Rightarrow Q(H) = 0$ for all $H \in F_T$.

RADON-NIKODYM THEOREM. Let the probability measure Q be absolutely continuous with respect to probability measure P , then there exists a random variable $\Lambda(\omega) \geq 0$; such that $\Lambda(\omega)$ is F_T measurable, $E_P[\Lambda(\omega)] = 1$ and $Q(H) = \int_H \Lambda(\omega) dP$ for any measurable set H . In that case we write $dQ/dP = \Lambda(\omega)$ where, $\Lambda(\omega)$ is the Radon-Nikodym derivative or the likelihood ratio.

DEFINITION (A MARTINGALE). A stochastic process $\{M_t\}_{t \geq 0}$ on $[\Omega, \mathcal{F}, P]$ of dimension n is called a martingale with respect to filtration $\{F_t\}_{t \geq 0}$ (and with respect to P) if

- M_t is $\{F_s\}_{s \geq 0}$ measurable for all t
- $E[|M_t|] < \infty$ for all t
- $E[M_s | F_t] = M_t$ for all $s \geq t$

The Martingale property. Suppose $Q \ll P$ with $dQ/dP = \Lambda(\omega)$ on F_T (filtration generated by the Brownian motion processes). Then $Q \ll P$ for all $t \in [0, T]$ and if we define $M_t = d(Q)/d(P)$, then M_t is a martingale with respect to F_t and P .

Although a linearization of nonlinear oscillators using the Girsanov theorem could be nonuniquely performed, an obvious and rather tempting choice is to arrive at the linearized equations by simply removing A_{nl} from Eq. (2). If this route is adopted, then the transformed (linearized) SDEs over T_i take the form

$$\begin{aligned} d\tilde{x}_1^{(j),i} &= \tilde{x}_2^{(j),i} dt \\ d\tilde{x}_2^{(j),i} &= a_1^{(j)}(\tilde{X}^i, \dot{\tilde{X}}^i, t) dt \\ &+ \sum_{r=1}^q b_r^{(j)}(t) \underbrace{\{[b_r^{(j)}(t)]^{-1} a_{nl}^{(j)}(\tilde{X}^i, \dot{\tilde{X}}^i, t) dt + dW_r(t)\}}_{d\hat{W}_r^i(t)} \\ j &= 1, 2, \dots, n \end{aligned} \quad (7)$$

where $\{\hat{W}_r(t) | r \in [1, q]\}$ is another vector of q -dimensional Brownian motion process (that absorbs the nonlinear part, A_{nl} , of the drift field) by the Levy characterization of Brownian motion [13] under a new probability measure Q . Note that $\{\hat{W}_r^i(t)\}$ as well as the linearized displacement and velocity processes, $\tilde{X}^i(t) := \{\tilde{x}_1^{(j),i} | j \in [1, n]\}$ and $\tilde{X}_2^i(t) := \{\tilde{x}_2^{(j),i} | j \in [1, n]\}$, are restricted to the interval T_i as indicated by the superscript i . Equation (7) is subject to initial conditions $\tilde{X}^i(t_{i-1}) = \tilde{X}^{i-1}(t_{i-1})$, $i > 1$ so that

$\tilde{X}^1(t_0) = \bar{X}(t_0)$ for $i = 1$. Moreover, we define $\tilde{X}^i = \{\tilde{x}_1^{(1),i}, \tilde{x}_2^{(1),i}, \dots, \tilde{x}_1^{(n),i}, \tilde{x}_2^{(n),i}\}^T$. Now, to compensate for the removal of nonlinear drift terms, we need to augment the linearized SDEs with the one-dimensional correction process $y^i(t)$ (i.e., the i th Radon-Nikodym derivative) that satisfies the following scalar SDE:

$$dy^i = \sum_{j=1}^n \sum_{r=1}^q [b_r^{(j)}(t)]^{-1} a_{nl}^{(j)}(\tilde{X}^i, \dot{\tilde{X}}^i, t) y^i d\hat{W}_r^i(t) \quad (8)$$

subject to the known initial condition $y^i(t_{i-1}) = y^{i-1}(t_{i-1})$ so that $y^1(t_0) = 1.0$ for $i = 1$. Toward a simpler presentation, we henceforth remove the superscript i provided that the interval T_i is unambiguously clear from the context. Note that an exact analytical solution to the linearized SDEs corresponding to Eq. (7) is readily available [2] irrespective of the dimension $2n$, and this solution is independent of $y(t)$. Indeed, a very attractive feature of the GLM is that it can exploit the available information on the solutions of linear SDEs—an area that is well researched and developed. If one writes Eqs. (7) and (8) in terms of the drift-modified Brownian motion processes $\hat{W}_r(t)$ for each r , then it is clear that the resulting SDEs are linear in $\tilde{X}$. Equation (7) may then be written in the new probability law ($\tilde{Q}$) as

$$d\tilde{X} = \tilde{A}\tilde{X} dt + \tilde{F}(t) dt + \sum_{r=1}^q [B_r(t)] d\hat{W}_r(t) \quad (9)$$

where

$$\tilde{A} = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & \cdots & 0 & 0 \\ -\hat{K}_{11} & -\hat{C}_{11} & -\hat{K}_{12} & -\hat{C}_{12} & \cdots & \cdots & -\hat{K}_{1n} & -\hat{C}_{1n} \\ 0 & 0 & 0 & 1 & \cdots & \cdots & 0 & 0 \\ -\hat{K}_{21} & -\hat{C}_{21} & -\hat{K}_{22} & -\hat{C}_{22} & \cdots & \cdots & -\hat{K}_{2n} & -\hat{C}_{2n} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \cdots & 0 & 1 \\ -\hat{K}_{n1} & -\hat{C}_{n1} & -\hat{K}_{n2} & -\hat{C}_{n2} & \cdots & \cdots & -\hat{K}_{nn} & -\hat{C}_{nn} \end{bmatrix}$$

is the system matrix of the GLM-based linearized Eq. (7). Note that $\hat{K}_{jk}$ and $\hat{C}_{jk}$ are the elements of stiffness and damping matrices and $\tilde{F}(t) \in \mathbf{R}^{2n}$ is obtainable from the normalized forcing vector $\hat{F}(t) \in \mathbf{R}^n$ (Eq. (3)) by augmenting the latter with zeros at locations of the displacement components. The solution of Eq. (9) with the initial condition vector $\tilde{X}_{t_{i-1}}^i := \tilde{X}^{i-1}(t_{i-1})$ is of the form

$$\begin{aligned} \tilde{X}(t) &= \exp\left(\tilde{A} \int_{t_{i-1}}^t ds\right) \tilde{X}_{t_{i-1}}^i + \exp\left(\tilde{A} \int_{t_{i-1}}^t ds\right) \\ &\times \int_{t_{i-1}}^t \left[\exp\left(-\tilde{A} \int_{t_{i-1}}^s ds_1\right) \tilde{F}(s_1) \right] ds + \exp\left(\tilde{A} \int_{t_{i-1}}^t ds\right) \\ &\times \int_{t_{i-1}}^t \left[\exp\left(-\tilde{A} \int_{t_{i-1}}^s ds_1\right) \sum_{j=1}^n \sum_{r=1}^q [b_r^{(j)}(t)] d\hat{W}_r(s) \right] ds \end{aligned} \quad (10)$$

The Radon-Nikodym derivative $y(t)$ is a strictly positive random process (an exponential martingale) computable as

$$y(t) = y_{i-1} \exp \left\{ \int_{t_{i-1}}^t \sum_{j=1}^n \sum_{r=1}^q [b_r^{(j)}(t)]^{-1} a_{nl}^{(j)}(\tilde{X}_1, \tilde{X}_2, t) dW_r(s) - \frac{1}{2} \int_{t_{i-1}}^t [b_r^{(j)}(t)]^{-2} [a_{nl}^{(j)}(\tilde{X}_1, \tilde{X}_2, t)]^2 ds \right\} \quad (11)$$

Computations of required expectations may now be accomplished in terms of the GLM-based linearized solution as: $\mathbf{E}[f(\tilde{X}_i, t_i)] = \mathbf{E}[y_i f(\tilde{X}_i, t_i)]$, subject to a localized form of Novikov's restrictions on A_{nl} (i.e., $\mathbf{E}[\exp(\frac{1}{2} \int_{t_{i-1}}^t A_{nl}^2 ds)] < \infty$) for $t \in T_i$.

Unfortunately, even if one chooses a very small step size, it is generally not possible to develop the above idea into a numerically stable algorithm. One may expect numerical instabilities in the form of the computed moment histories underflowing or overflowing along with widely varying oscillations. Reasons behind this may be traced to the fact that, for the Girsanov theorem to be valid, expectations of the strictly positive martingales $\{y^i(t)\}$ should be exactly equal to 1.0 $\forall i \in \mathbf{Z}^+$ [3]. For instance, if the drift nonlinearity is such that it affects the solution significantly away from that corresponding to the linear SDE (i.e., the SDE without A_{nl}), then sample paths of $\{y^i(t)\}$ would oscillate considerably away from 1. These fluctuations imply that the computed expectation of $y^i(t)$ for a given i would generally be quite different from 1.0 for finite (possibly small) ensemble sizes. Moreover, these errors would also propagate very fast in time, as explained in the following. Restricting attention to the interval T_i , we note that the argument of $y^i(t)$ may be expanded in an Ito-Taylor's series (each term of which is a zero-mean Gaussian Ito integral) so that $y^i(t)$ essentially admits an expression in the form of an infinite product of log-normal (exponentiated Gaussian) random variables. However, owing to a truncation of the Ito-Taylor expansion after a certain number of terms, this product would remain finite for all practical purposes. Moreover, as we progress over successive time intervals, $y(t)$ is updated once more through exponentiation, i.e., according to the identity: $y^i(t) = \zeta(t) \prod_{k=1}^{i-1} y_k^k$, where $\zeta(t)$ is the stochastic exponential that appears as the coefficient of y_{i-1}^{i-1} in Eq. (11). Although the construction of $y(t)$ and the linear solution through exponential functions allows the GLM to be developed into a Lie group method, evaluations of such functions require more floating point operations thereby leading to higher errors. Thus, it is nearly impossible to precisely satisfy the requirement $\mathbf{E}[y^i(t)] = 1$ especially for sufficiently large i as the errors also accumulate exponentially. In other words, given that $y(t)$ is composed purely of exponential functions, there are just two possibilities, viz. either $\lim_{i \rightarrow \infty} \mathbf{E}[y^i(t)] = 0$ or $\lim_{i \rightarrow \infty} \mathbf{E}[y^i(t)] = \infty$.

Figure 1(a) shows typically zero-mean Gaussian and (the corresponding) log-normal probability density functions (PDF) and Fig. 1(b) shows two sample paths of typically cumulative products of log-normal random variables—while one of these paths overflows, the other underflows. It is evident (Fig. 1(a)), unlike the Gaussian PDF (G-PDF) that is symmetric ~ 0 , the log-normal density is not symmetric ~ 1.0 . Indeed, clustering of the realizations of such a log-normal variable (generated from a zero-mean Gaussian variable) is more toward the left of 1 than to the right. Thus, it is this asymmetry of the log-normal PDF that causes paths of $y^i(t)$ underflow more often than overflow as $i \rightarrow \infty$. The same reasoning also convinces us about the near impossibility of arriving at a GLM-based linearized drift field uniformly valid over the entire time interval of interest because this would create very large fluctuations in the paths of $y(t)$ away from 1.0. In other words, the GLM has to be applied along with a discretization of the time axis and it is thus a numeric-analytic procedure for SDEs. These are

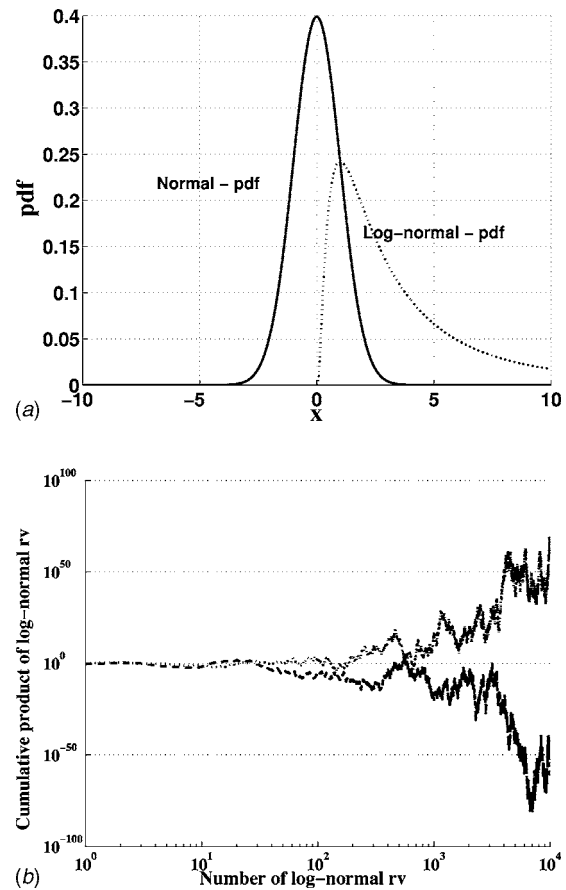


Fig. 1 (a) Typical PDFs of Gaussian and log-normal distributions and (b) the cumulative product of lognormal random variables

probably the reasons why methods based on Girsanov's measure transformation have hardly been applied for stochastically driven nonlinear oscillators of engineering interest.

In order to circumvent such numerical difficulties in the GLM, it is then required to modify the linear SDEs that provide a solution "close" to that of the nonlinear SDEs, at least locally (i.e., within a given subinterval T_i). This would, in turn, reduce the oscillations (away from 1.0) in the sample paths of $y^i(t)$ to a considerable extent. This is presently achieved by locally introducing a "subtracting set" S^i , whose elements are simply various terms in the stochastic Taylor (Ito-Taylor) expansion of the nonlinear vector function $A_{nl}^i(\tilde{X}_1, \tilde{X}_2, t)$ (restricted to T_i) about $A_{nl}^{i-1}(t_{i-1}) := A_{nl,i-1}$. Moreover, these elements are hierarchically arranged in the same way as they appear in the Ito-Taylor series. Taking, for instance, the j th scalar component $a_{nl}^{(j),i}(\tilde{X}_1, \tilde{X}_2, t)$ of $A_{nl}^i(\tilde{X}_1, \tilde{X}_2, t)$, its Ito-Taylor expansion based at $a_{nl,i-1}^{(j)}$ is of the form

$$a_{nl}^{(j),i} = a_{nl,i-1}^{(j)} + \sum_{k=1}^n \sum_{r=1}^q b_r^{(k)} \frac{\partial a_{nl,i-1}^{(j)}}{\partial x_2^{(k)}} (\sqrt{t - t_{i-1}}) + \left[\frac{\partial}{\partial t} a_{nl,i-1}^{(j)} + \sum_{k=1}^n \left(x_2^{(k)} \frac{\partial}{\partial x_1^{(k)}} + a_{i-1}^{(k)} \frac{\partial}{\partial x_2^{(k)}} \right) a_{nl,i-1}^{(j)} + \frac{1}{2} \sum_{k,l=1}^n \sum_{r,s=1}^q b_r^{(k)} b_s^{(l)} \frac{\partial^2 a_{nl,i-1}^{(j)}}{\partial x_2^{(k)} \partial x_2^{(l)}} \right] (t - t_{i-1}) + \dots \quad (12)$$

The subtracting set is of the form: $S^i = \{S_1^i, S_2^i, S_3^i, \dots\}$, where, S_1^i

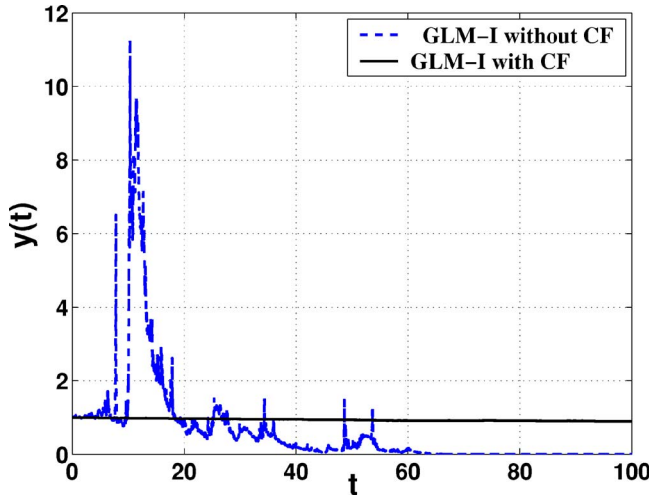


Fig. 2 A realization of the process $y(t)$ with and without the CF

$= A_{nl,i-1}$, $S_2^i = \sum_{k=1}^n \sum_{r=1}^q b_r^{(k)} \partial a_{nl,i-1}^{(j)} / \partial x_2^{(k)} (\sqrt{t-t_{i-1}})$, and so on. Thus, with the incorporation of the elements of the subtracting set as hierarchically corrective forcing terms we rewrite Eq. (2) as

$$d\tilde{X} = \tilde{A}\tilde{X} dt + \tilde{F}(t)dt + \tilde{\gamma}_v dt + \sum_{r=1}^q B_r(t) \{B_r^{-1}(t)[A_{nl}(X, \dot{X}, t) - \tilde{\gamma}_v]dt + d\hat{W}_r(t)\}$$

Then, we replace the term $\{B_r^{-1}(t)[A_{nl}(X, \dot{X}, t) - \tilde{\gamma}_v]dt + d\hat{W}_r(t)\}$ with $d\tilde{W}_r(t)$. Therefore, the linearized SDEs in $(\tilde{X}_1, \tilde{X}_2)$ are obtainable as

$$d\tilde{X} = \tilde{A}\tilde{X} dt + \tilde{F}(t)dt + \tilde{\gamma}_v dt + \sum_{r=1}^q B_r(t)d\tilde{W}_r(t) \quad (13)$$

where, $\tilde{\gamma}_v(t) \in \mathbf{R}^{2n}$ is obtained from the normalized forcing vector $\gamma_v = \sum_{i=1}^v S_i$; $\gamma_v \in \mathbf{R}^n$, $v \in \mathbf{Z}^+$ (just as $\tilde{F}(t) \in \mathbf{R}^{2n}$ is obtainable from the normalized forcing vector $\hat{F}(t) \in \mathbf{R}^n$) by augmenting the latter with zeros at locations of the displacement components. The Radon-Nikodym derivative is given by

$$y(t) = y_{i-1} \exp \left\{ \int_{t_{i-1}}^t \sum_{j=1}^n \sum_{r=1}^q [b_r^{(j)}(t)]^{-1} [a_{nl}^{(j)}(\tilde{X}_1, \tilde{X}_2, t) - \gamma_v^{(j)}] d\tilde{W}_r(s) - \frac{1}{2} \int_{t_{i-1}}^t [b_r^{(j)}(t)]^{-2} [a_{nl}^{(j)}(\tilde{X}_1, \tilde{X}_2, t) - \gamma_v^{(j)}]^2 ds \right\} \quad (14)$$

From limited numerical experiments, it has been observed that it generally suffices to use only the first element S_1 of the subtracting set and thus just employ $\tilde{\gamma}_1$ as the corrective forcing vector to construct the linearized SDEs (Eq. (13)). We may consequently expect an adequate reduction in the spurious oscillations of $y(t)$ in the process. Figure 2 shows the variation of stochastic process $y(t)$ with and without application of γ_1 . As is indeed verifiable from this figure, by using γ_v (Eq. (14)) we are in fact trying to force the expectation of $y(t)$ close to 1.0. Subsequently, the functional ν of Eq. (6) can be approximately computed through its Monte Carlo estimator

$$\hat{\nu}(t) = \frac{1}{N} \sum_{i=1}^N y^{(i)}(t) [f(\tilde{X}^{(i)}, \dot{\tilde{X}}^{(i)}, t)]; \quad t \in T_i \quad (15)$$

We capture the essence of the idea behind the GLM through the following theorem.

Theorem on Girsanov Linearization. Consider the system of SDEs (Eq. (2)) in terms of the process $\tilde{X}(t): \mathbf{R} \rightarrow \mathbf{R}^{2n}$ with $\tilde{X}(t=0) = \tilde{X}_0$ and $W_r(t)$ being a q -dimensional Wiener process under the filtration F_t . Let the velocity drift vector $A(\tilde{X}, t)$ be decomposed additively as $A = A_l + A_{nl}$, where A_l and A_{nl} are respectively linear and nonlinear in $\tilde{X}$. Now consider the linearized SDEs (Eq. (13)) in $\tilde{X} \in \mathbf{R}^{2n}$ with the initial condition $\tilde{X}(t=0) = \tilde{X}_0$. Under a local form of Novikov restriction for $t \in T_i = (t_{i-1}, t_i]$ (i.e., $\mathbf{E}[\exp(\frac{1}{2} \int_{t_{i-1}}^t (A_{nl} - \tilde{\gamma}_v)^2 ds)] < \infty$), the Radon-Nikodym derivative $y(t)$ in Eq. (14) is a positive martingale with respect to F_t and measure P . Then the measure $\tilde{Q}$ for the linearized process $\tilde{X}$ is given by $d\tilde{Q} = y(t)dP$ so that the $\tilde{Q}$ -law of $\tilde{X}$ is same as the P -law of $\tilde{X}$. Thus, if $f(\tilde{X}, t): \mathbf{R}^{2n} \times \mathbf{R} \rightarrow \mathbf{R}$ is a (weakly) bounded and measurable function, then the following identity holds: $\mathbf{E}_P[f(\tilde{X}, t)] = \mathbf{E}_{\tilde{Q}}[y(t)f(\tilde{X}, t)]$.

An "Order Infinity" Algorithm. We have thus far been focusing only on issues related to how to check large oscillations in $y(t)$ and in the process preventing overflows and underflows. However, use of Ito-Taylor expansions for the evaluation of $y(t)$ still remains a stumbling block in a versatile and efficient algorithmic implementation of GLM. Computing MSIs not only requires substantial computational effort, the associated Ito-Taylor expansion also demands that the nonlinear part of the vector field be sufficiently differentiable—a condition that is generally not met in impact and elastoplastic problems. Even though the GLM is a numeric-analytical technique, its implementation as outlined above requires the evaluation of a large number of MSIs for computing $y(t)$ over every time step (see the Appendix). For instance, an $O(h^4)$ time-marching algorithm (map) based on a direct Ito-Taylor expansion for an SDOF Duffing oscillator under an additive white noise requires 5 MSIs as compared to 10 MSIs required in the GLM. Indeed, adopting the above form of the GLM requires about twice the number of Ito integrals than that required for solutions of SDEs through a direct Ito-Taylor expansion. Worse, we cannot use such expansions for nondifferentiable nonlinear functions appearing in the original SDEs that are not sufficiently differentiable.

We now propose an alternative procedure for computing $y(t)$ without taking recourse to stochastic Taylor expansion. In particular, we focus on a different procedure for the evaluation of integrals of the forms (within a subinterval T_i) $\int_{t_{i-1}}^{t_i} a_{nl}^{(j)}(\tilde{X}_1, \tilde{X}_2, t) d\tilde{W}_r(s)$ and $\int_{t_{i-1}}^{t_i} [a_{nl}^{(j)}(\tilde{X}_1, \tilde{X}_2, t)]^2 ds$. Recall that computation of $y(t)$ (Eq. (14)) involves only these two typical integrals. The first may be interpreted as a zero-mean Gaussian random variable with variance $\int_{t_{i-1}}^{t_i} \mathbf{E}[a_{nl}^{(j)}(\tilde{X}_1, \tilde{X}_2, t)]^2 ds$, i.e., the integral of the second type. It therefore suffices to restrict our attention to the latter integral alone. We use a Monte Carlo approach for evaluation of this integral [32]. Thus, we treat $\int_{t_{i-1}}^{t_i} \mathbf{E}[a_{nl}^{(j)}(\tilde{X}_1, \tilde{X}_2, t)]^2 ds$ as a mathematical expectation with respect to a new probability measure $\tilde{P}$, whose density is uniformly distributed in $[t_{i-1}, t_i]$. In other words, we have the identity

$$\int_{t_{i-1}}^{t_i} \mathbf{E}[a_{nl}^{(j)}(\tilde{X}_1, \tilde{X}_2, t)]^2 ds = h_i \mathbf{E}_{\tilde{P}}[\mathbf{E}[a_{nl}^{(j)}(\tilde{X}_1, \tilde{X}_2, t)]^2] \quad (16)$$

where the expectation inside the bracket is with respect to the Girsanov-transformed measure $\tilde{Q}$. Accordingly, we generate a set of random time instants $\{t(\omega_k) \in T_i | k \in \mathbf{Z}^+\}$ that are uniformly distributed in T_i . Thus, based on the known analytical solution of the linearized SDEs (Eq. (13)), we can evaluate $\mathbf{E}[a_{nl}^{(j)}(\tilde{X}_1, \tilde{X}_2, t)]^2$ as

a continuous function of $t \in T_i$ in closed form and then find its values at each of $t(\omega_k)$. Now, to evaluate the expectation on the right-hand side of Eq. (15), we again employ the Monte Carlo estimator

$$\mathbf{E}_{\tilde{P}}\{\mathbf{E}[a_{nl}^{(j)}(\tilde{X}_1, \tilde{X}_2, t)]^2\} \cong \frac{1}{N} \sum_{k=1}^N [\mathbf{E}(a_{nl}^{(j)}\{\tilde{X}_1[t(\omega_k)], \tilde{X}_2[t(\omega_k)], t(\omega_k)\})^2] \quad (17)$$

Given the finite precision arithmetic used in all computing machines, it is generally not feasible to increase the accuracy of numerical results indefinitely. For instance, if 10^{-16} is the lowest absolute value of a real number that a computer can use in (floating point) algebraic operations, even a so-called “ $O(h^\infty)$ method” would only be able to achieve an approximation accuracy of order $<10^{-16}$. If we denote the local order of accuracy in computing the left-hand side of Eq. (16) to be $O(h_i^{\alpha+1})$, then this is the highest possible order if the identity $h^{\alpha+1}=10^{-16}$ holds. Solving this identity for α and assuming $\mathbf{E}\{[a_{nl}^{(j)}(\tilde{X}_1, \tilde{X}_2, t)]^2\}$ to be sufficiently smooth, it is possible to represent this expectation locally over T_i in terms of an interpolating polynomial of order α with $\alpha+1$ interpolation points in T_i . If such an arrangement is adopted, then one may compute $\mathbf{E}\{[a_{nl}^{(j)}(\tilde{X}_1, \tilde{X}_2, t)]^2\}$ anywhere in T_i based on its known discrete values at just $\alpha+1$ points. Such an arrangement makes sense as the determination of $\mathbf{E}\{[a_{nl}^{(j)}(\tilde{X}_1, \tilde{X}_2, t)]^2\}$ from the known linearized solution is quite computationally intensive. However, in order to avoid numerical instabilities (such as Runge’s phenomenon [33]) associated with Lagrangian interpolating polynomials, we prescribe using spline interpolations of the appropriate order. Thus, we subdiscretize the interval T_i into α subintervals (of equal lengths) with $(\alpha+1)$ equispaced points (t_{i-1} and t_i being two boundary points), obtain the solution $\mathbf{E}\{[a_{nl}^{(j)}(\tilde{X}_1, \tilde{X}_2, t)]^2\}$ only at those points and pass the spline interpolant through these points.

3 Numerical Illustrations

For purposes of illustration, we presently choose two SDOF and one 2-DOF nonlinear oscillators with cubic nonlinearity. The two SDOF oscillators are the hardening Duffing (HD) and the two-well Duffing-Holmes (DH) oscillators. Taken together, these two SDOF oscillators exhibit a large class of phenomena characteristic of many higher-dimensional nonlinear dynamical systems. The purpose behind choosing the 2-DOF example is to demonstrate the ready applicability of the GLM to nonlinear MDOF problems.

An SDOF Hardening Duffing Oscillator (HD) Under Additive Noise. First consider an HD oscillator only under additive white noise. The governing equation is of the form

$$\ddot{x} + C\dot{x} + K_1x + K_2x^3 = \sigma\dot{W}(t) \quad (18)$$

where K_1 , K_2 , and C are positive real constants representing the stiffness, the “strength” of nonlinearity, and the coefficient of viscous damping, respectively, $W(t)$ is a standard Brownian motion process, and σ denotes the intensity of noise. Equation (18) is subject to the initial condition $X_0 = \{x_0, \dot{x}_0\}^T$. We also denote $X := \{x_1, x_2\}^T$. Writing Eq. (18) in the Girsanov-transformed form of Eq. (7), the linearized SDEs and the scalar SDE corresponding to the exponential correction process over T_i may be written as

$$d\tilde{x}_1(t) = \tilde{x}_2(t)dt$$

$$d\tilde{x}_2(t) = (-C\tilde{x}_2 - K_1\tilde{x}_1)dt + \underbrace{\sigma\{-\sigma^{-1}K_2\tilde{x}_1^3dt + dW(t)\}}_{d\hat{W}(t)}$$

$$dy(t) = -\sigma^{-1}K_2\tilde{x}_1^3y d\hat{W}(t) \quad (19)$$

Solutions for the linearized state $\tilde{X} := \{\tilde{x}_1, \tilde{x}_2\}^T$ and the correction process $y(t)$ presently take the form

$$\tilde{X}(t) = \exp \tilde{A}(t - t_{i-1})X_0 + \exp \tilde{A}(t - t_{i-1}) \int_{t_{i-1}}^t \exp(-\tilde{A}(s - t_{i-1})) \times \begin{bmatrix} 0 & 0 \\ 0 & \sigma \end{bmatrix} \begin{Bmatrix} 0 \\ d\hat{W}_1(s) \end{Bmatrix} \quad (20a)$$

$$y(t) = \exp \left[\int_{t_{i-1}}^t (-\sigma^{-1}K_2\tilde{x}_1^3) d\hat{W}(t) - \frac{1}{2} \int_{t_{i-1}}^t (-\sigma^{-1}K_2\tilde{x}_1^3)^2 ds \right] \quad (20b)$$

where $\tilde{A} = \begin{bmatrix} 0 & 1 \\ -K_1 & -C \end{bmatrix}$ and $B_r(t) = \begin{Bmatrix} 0 \\ \sigma \end{Bmatrix}$. The exponential solution for $y(t)$ has integral arguments that may be expanded in terms of MSIs through the Ito-Taylor expansion (see Appendix). By using the estimator Eq. (15) we may readily obtain the expectation of a given function $f(X, t)$.

Now we turn attention to the linearization method using the subtracting set S^i corresponding to the Ito-Taylor expansion of the nonlinear vector function $K_2x_1^3$. Consider first element $S_1^i = -K_2x_{1,i-1}^3$ of S^i and add it to the linear part of the drift a_l while subtracting the same from the nonlinear part a_{nl} . The modified drift term in the linearized velocity equation is then $a_l(\tilde{X}, t) = -C\tilde{x}_2 - K_1\tilde{x}_1 - K_2\tilde{x}_{1,i-1}^3$. The term that goes into the transformed Brownian motion is $a_{nl}(\tilde{X}, t) = -K_2\tilde{x}_1^3 + K_2\tilde{x}_{1,i-1}^3$. We then have the following incremental form for the Girsanov-linearized SDEs:

$$d \begin{Bmatrix} \tilde{x}_1 \\ \tilde{x}_2 \end{Bmatrix} = \begin{bmatrix} 0 & 1 \\ -K_1 & -C \end{bmatrix} dt + \begin{Bmatrix} 0 \\ -K_2\tilde{x}_{1,i-1}^3 \end{Bmatrix} dt + \sigma \left\{ \underbrace{-\sigma^{-1}[K_2\tilde{x}_1^3 - K_2\tilde{x}_{1,i-1}^3]dt + dW(t)}_{d\tilde{W}(t)} \right\} \quad (21)$$

The additional scalar SDE for the correction term is

$$dy(t) = -\sigma^{-1}[K_2\tilde{x}_1^3 - K_2\tilde{x}_{1,i-1}^3]y d\tilde{W}(t) \quad (22)$$

This yields the solution for $y(t)$ as

$$y(t) = y_{i-1} \exp \left[\int_{t_{i-1}}^t (-\sigma^{-1}[K_2\tilde{x}_1^3 - K_2\tilde{x}_{1,i-1}^3]) d\tilde{W}(t) - \frac{1}{2} \int_{t_{i-1}}^t (-\sigma^{-1}[K_2\tilde{x}_1^3 - K_2\tilde{x}_{1,i-1}^3])^2 ds \right] \quad (23)$$

According to the notation in Eq. (13), we have $\tilde{\gamma} = \begin{Bmatrix} 0 \\ -K_2\tilde{x}_{1,i-1}^3 \end{Bmatrix}$ and $B_r(t) = \begin{Bmatrix} 0 \\ \sigma \end{Bmatrix}$. Note that both the integrals in the right-hand side of Eq. (23) are stochastic and computed using both the versions of the GLM as described earlier. Version 1 (henceforth referred to as GLM-I), which uses an Ito-Taylor expansion, is presently done to $O(h^4)$ and the modeling details of the associated MSIs are described in the Appendix. Recall that, unlike version 1, version 2 (GLM-II) provides a so-called $O(h^\infty)$ solution (i.e., the accuracy of solution is only affected by the finiteness of ensemble sizes) and does not require computing any derivatives and MSIs.

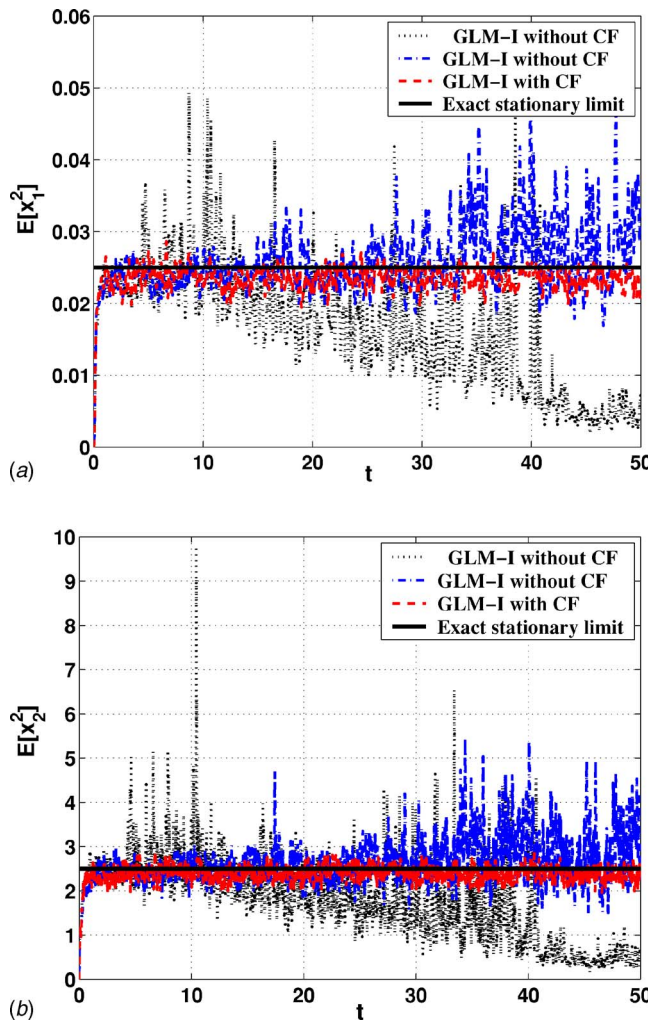


Fig. 3 The HD oscillator (Eq. (18))—histories of second moments: (a) $E[x_1^2]$ and (b) $E[x_2^2]$; $C=5.0$, $K_1=K_2=100.0$, $\sigma=5.0$, $h=0.01$, $X_0=\{0,0\}^T$; solid black lines indicate exact stationary limits

In Figs. 3(a) and 3(b), we show time histories of $E(x_1^2)$ and $E(x_2^2)$ with the two variants of GLM-I (with and without the first element S_1^i of the subtracting set leading to the corrective forcing vector $\bar{\gamma}_1$) as well as the exact stationary limits. Note that the exact stationary (joint) density function is available in this case through the solution of the reduced Fokker-Planck equation and is given by [34]

$$p(x, \dot{x}) = R \exp\left(-\frac{C\dot{x}^2}{\sigma^2} - \int_0^x C \frac{K_1 q + K_2 q^3}{0.5\sigma^2} dq\right) \quad (24)$$

where the real constant R has to be so found as to satisfy the normalization constraint $\int_{-\infty}^{\infty} p(x, \dot{x}) dx d\dot{x} = 1$. The statistical moment of any deterministic function $\Phi(x, \dot{x})$ may then be found as $E[\Phi(x, \dot{x})] = \int_{-\infty}^{\infty} \Phi(x, \dot{x}) p(x, \dot{x}) dx d\dot{x}$. The integrations required to find the stationary limits are presently computed using the symbolic manipulator MAPLE®. In Fig. 3, we see two instances of solutions via GLM-I without the usage of the corrective forcing (CF). While, in one such instance, we have an overflowing trend, the other instance has a typical solution that tends to underflow. The GLM-I with the CF however appears to work without numerical instabilities, and the resulting solutions match well the exact stationary limit as time progresses. Thus, unless otherwise specified, we apply GLM-I and II with the CF from now onwards.

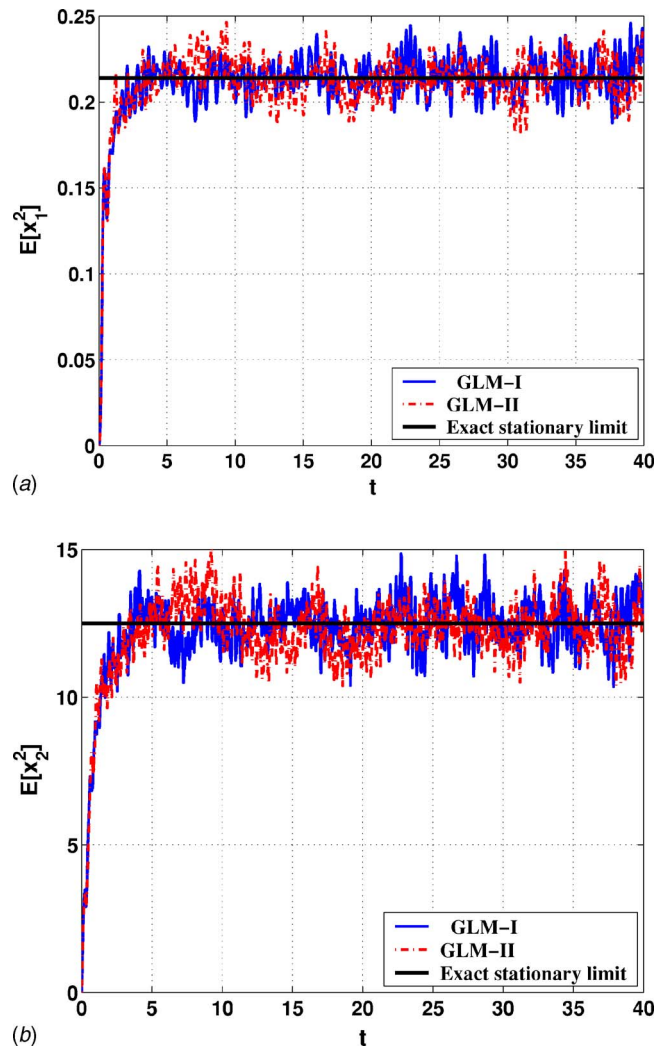


Fig. 4 The HD oscillator (Eq. (18))—histories of second moments through GLM-I and II: (a) $E[x_1^2]$ and (b) $E[x_2^2]$; $C=1.0$, $K_1=10.0$, $K_2=100.0$, $\sigma=5.0$, $h=0.01$, $X_0=\{0,0\}^T$; solid black lines indicate exact stationary limits

Figure 4 shows second moment histories of displacement and velocity components via GLM-I and II when the nonlinearity parameter (K_2) is ten times the stiffness (K_1) and the additive noise intensity ($\sigma=5.0$) is very high. Also shown in Fig. 4 are the corresponding stationary limits. In this and many other such cases, both the versions of GLM work accurately, even though variance of the estimate appears to be less with GLM-II for the same ensemble size. Thus, in all such cases, results would henceforth be plotted with any one of the versions. It is interesting to observe (see Eq. (24)) that K_2 only affects the stationary limit for $E[x_1^2(t)]$ but not that for $E[x_2^2(t)]$. Figure 5 shows the variations of $E[x_2^2(t)]$, as obtained through GLM-I, for different values of K_2 as well as the associated stationary limits (black and solid lines). Time histories of $E[x_1^2(t)]$ are plotted in Fig. 6 using GLM-II for different values of the time step-size, h . Incidentally, for $h=0.05$, GLM-I fails to work (as indicated in the inset) and numerical instabilities set in for $t \geq 0.8$ s or so.

An HD Oscillator Under Combined Deterministic Excitation and Additive Noise. The governing equation is presently taken to be in the form [9]

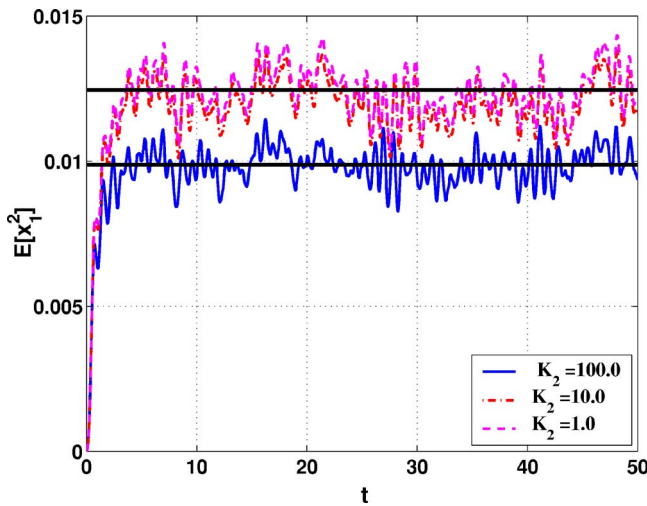
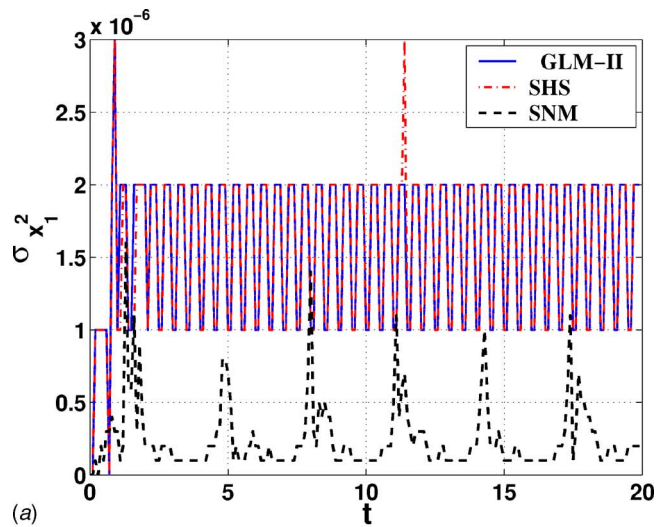


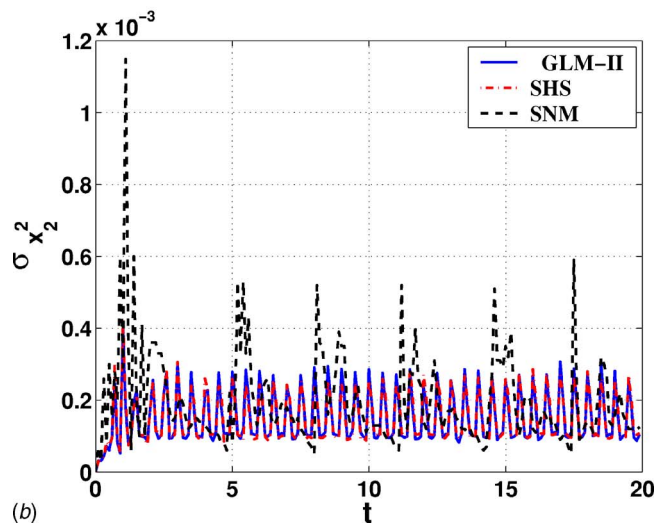
Fig. 5 The HD oscillator (Eq. (18))— $E[x_1^2]$ for different values of K_2 using GLM-II; $C=1.0$, $K_1=10.0$, $\sigma=0.50$, $h=0.01$, $X_0=\{0,0\}^T$; solid black lines indicate exact stationary limits

$$\ddot{x} + 2\pi\epsilon_1\dot{x} + 4\pi^2\epsilon_2x(1+x^2) = 4\pi^2\epsilon_3\cos(2\pi t) + 4\pi^2\epsilon_4\dot{W}_1(t) \quad (25)$$

A look at the solution for $y(t)$ (see Eq. (13)) shows that it contains the inverse of the diffusion coefficient (noise intensity). Since division by a small number could cause numerical problems, it would be interesting to see if the GLM had any instability for very low noise intensity. Figure 7 shows plots of variance histories of displacement and via GLM-II corresponding to the HD oscillator of Eq. (25) with $\epsilon_4=0.0005$ (i.e., very low noise intensity). Here, the sinusoidal forcing parameter $\epsilon_3=1.0$ is so chosen that the oscillator exhibit a dumb-bell-shaped one-periodic solution for $\epsilon_4=0$ and such a structure is expected to be preserved under “small” additive noise. Since no exact solutions are presently available, we choose the implicit stochastic Newmark method (SNM) [35,36] and the explicit stochastic Heun method (SHS) [5] for purposes of comparison. Although the displacement updates via SNM and SHS are comparable, the velocity update via SNM is



(a)



(b)

Fig. 7 The HD oscillator (Eq. (25))— $\epsilon_1=0.25$, $\epsilon_2=1.0$, $\epsilon_3=1.0$, $\epsilon_4=0.0005$, $h=0.01$, $X_0=\{0,0\}^T$; histories of (a) variance of displacement and (b) variance of velocity, using GLM-II, SHS, and SNM

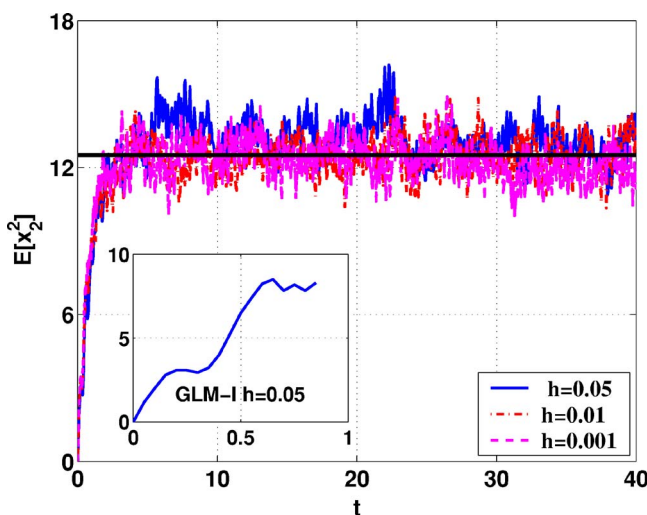


Fig. 6 The HD oscillator (Eq. (18))—histories of $E[x_2^2]$ for different values of h using GLM-II; $C=1.0$, $K_1=10.0$, $K_2=100.0$, $F(t)=0$, $\sigma=5.0$, $X_0=\{0,0\}^T$; the solid black line indicates the exact stationary limit. Inset shows that GLM-I terminates unsuccessfully before 1 s for $h=0.05$.

just as good as a drift-implicit Euler method, which is worse than the SHS. As is seen from Fig. 7, while solutions via GLM and SHS match reasonably, solutions via the SNM are too far off. However, over most of the parameter ranges, the results through SNM indeed match closely with those via GLM. A typical such case is reported in Fig. 8.

A Duffing-Holmes (DH) Oscillator Under Combined Deterministic Excitation and Additive Noise. We now take up a DH oscillator subject to a sinusoidal force and an additive white-noise excitation. For $\epsilon_3=\epsilon_4=0$, the system has two potential wells (basins of attraction) corresponding to the two stable fixed points at $(x,\dot{x})=(1,0)$ and $(x,\dot{x})=(-1,0)$. This oscillator may, for instance, be a useful model for the nonlinear dynamics of a periodically forced buckled beam. The governing equation is

$$\ddot{x} + 2\pi\epsilon_1\dot{x} + 4\pi^2\epsilon_2x(-1+x^2) = 4\pi^2\epsilon_3\cos(2\pi t) + 4\pi^2\epsilon_4\dot{W}_1(t) \quad (26)$$

In the absence of the diffusion term, this oscillator exhibits a wide variety of multiperiodic and chaotic solutions. Figure 9 shows displacement variance histories as well as phase plots of $E[x_1(t)]$ versus $E[x_2(t)]$ as determined via GLM-I, SHS, and SNM. The

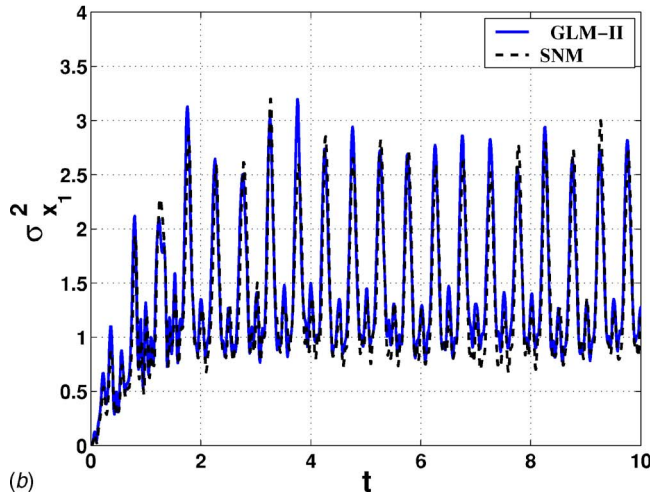
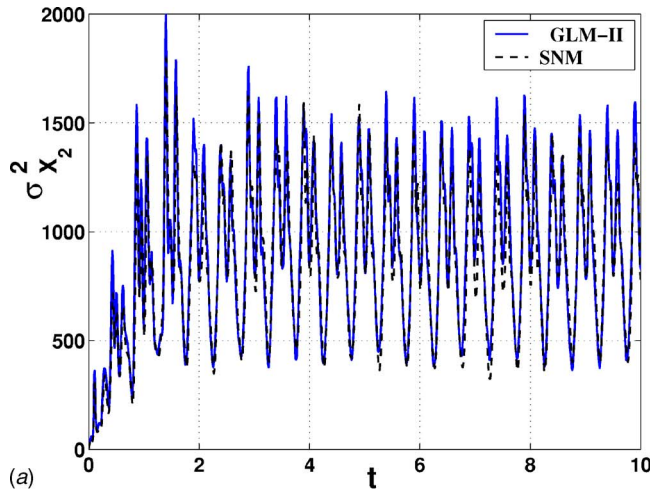


Fig. 8 The HD oscillator (Eq. (25))—histories of (a) variance of displacement and (b) variance of velocity using GLM-II and SNM; $\varepsilon_1=0.25$, $\varepsilon_2=1.0$, $\varepsilon_3=42.0$, $\varepsilon_4=1.0$, $h=0.01$, $X_0=\{0,0\}^T$

parameters $\varepsilon_1, \varepsilon_2, \varepsilon_3$ are so chosen that they belong to the chaotic regime of the oscillator when $\varepsilon_4=0$. When the noise intensity is small (presently $\varepsilon_4=0.05$), the SNM again fails to work accurately with SHS and GLM still providing a close match (Fig. 9). However, as ε_4 becomes higher, SNM-based results match closely with those via GLM and SHS (Fig. 10).

A 2-DOF Nonlinear Oscillator Under Additive Noises. In order to numerically demonstrate the applicability of the GLM to higher dimensional oscillators, we choose a 2-DOF nonlinear oscillator driven by additive white noises and described by

$$\begin{aligned} \ddot{x}_1 + C_1 \dot{x}_1 + (K_1 + K_2)x_1 - K_2 x_3 + \alpha x_1^3 &= \sigma_1 \dot{W}_1(t) \\ \ddot{x}_3 + C_2 \dot{x}_3 + (K_2 + K_3)x_3 - K_2 x_1 &= \sigma_2 \dot{W}_2(t) \end{aligned} \quad (27)$$

The derivation of the GLM for this oscillator follows precisely from the general formulation recorded in the previous section. However, for completeness, we retrace the main steps briefly. Thus, introducing the state space vector $\bar{X} := \{x_1^{(1)}, x_2^{(1)}, x_1^{(2)}, x_2^{(2)}\}^T$, the displacement vector $X_1 := \{x_1^{(1)}, x_1^{(2)}\}^T$ and the velocity vector $X_2 := \{x_2^{(1)}, x_2^{(2)}\}^T$, the incremental form of the above SDE-s is written as

$$dx_1^{(1)} = x_2^{(1)} dt$$

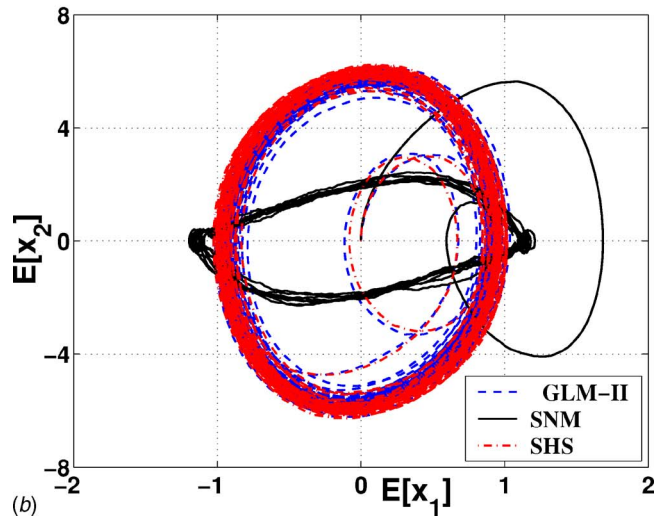
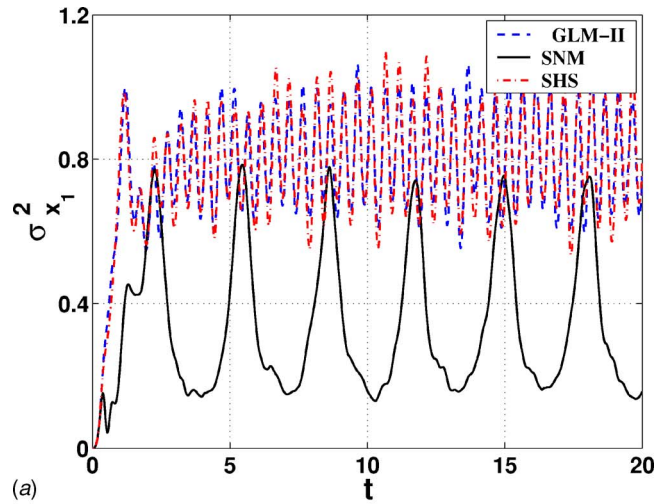


Fig. 9 The DH oscillator (Eq. (26))—histories of (a) variance of displacement and (b) phase plot of $E[x_2]$ and $E[x_1]$ using GLM-II, SNM and SHS; $\varepsilon_1=0.25$, $\varepsilon_2=0.5$, $\varepsilon_3=0.5$, $\varepsilon_4=0.10$, $h=0.01$, $X_0=\{0,0\}^T$

$$dx_2^{(1)} = [-C_1 x_2^{(1)} - (K_1 + K_2)x_1^{(1)} - \alpha(x_1^{(1)})^3 + K_2 x_1^{(2)}]dt + \sigma_1 dW_1(t)$$

$$dx_1^{(2)} = x_2^{(2)} dt$$

$$dx_2^{(2)} = [-C_2 x_2^{(2)} - (K_2 + K_3)x_1^{(2)} + K_2 x_1^{(1)}]dt + \sigma_2 dW_2(t) \quad (28)$$

These SDEs are subject to the initial condition vector $\bar{X}_0 = \{x_{10}^{(1)}, x_{20}^{(1)}, x_{10}^{(2)}, x_{20}^{(2)}\}^T$. Now in order to derive the GLM-based linearized form (as in Eq. (13)) over the i th interval T_i , we need to construct the corrective forcing $\tilde{\gamma}_1^{(i)}$ by using the first element S_1^i of the subtracting set S^i . Accordingly, the linearized and residual (nonlinear) parts of the velocity drift corresponding to the second scalar SDE in Eq. (28) are given by

$$a_l^{(2)}(\tilde{X}, t) = -C_1 \tilde{x}_2^{(1)} - (K_1 + K_2)\tilde{x}_1^{(1)} + K_2 \tilde{x}_1^{(2)} - \alpha(\tilde{x}_{1,i-1}^{(1)})^3$$

$$\text{and } a_{nl}^{(2)}(\tilde{X}, t) = -\alpha(\tilde{x}_1^{(1)})^3 + \alpha(\tilde{x}_{1,i-1}^{(1)})^3 \quad (29)$$

Then, the GLM-based linearized SDE-s are given by

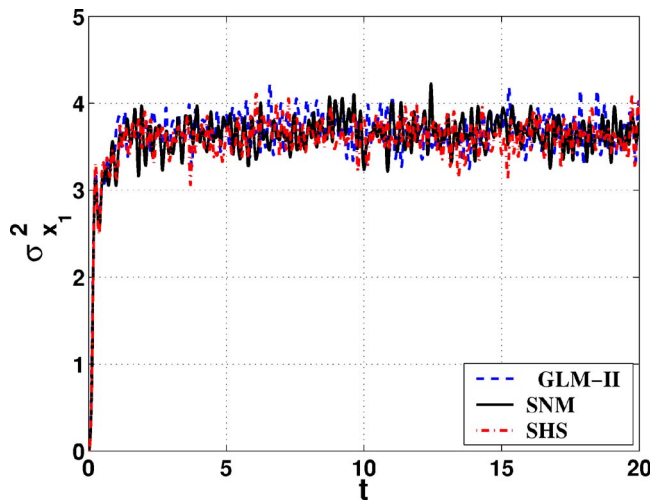


Fig. 10 The DH oscillator (Eq. (26))—histories of variance of displacement using GLM-II, SNM and SHS; $\varepsilon_1=0.25$, $\varepsilon_2=0.5$, $\varepsilon_3=0.5$, $\varepsilon_4=1.0$, $h=0.01$, $X_0=\{0,0\}^T$

$$d\tilde{x}_1^{(1)}(t) = \tilde{x}_2^{(1)}(t)dt$$

$$d\tilde{x}_2^{(1)}(t) = [-C_1\tilde{x}_2^{(1)} - (K_1 + K_2)\tilde{x}_1^{(1)} + K_2\tilde{x}_1^{(2)} - \alpha(\tilde{x}_{1,i-1}^{(1)})^3]dt + \sigma_1 \underbrace{\{-\sigma_1^{-1}[\alpha(\tilde{x}_1^{(1)})^3 - \alpha(\tilde{x}_{1,i-1}^{(1)})^3]dt + dW_1(t)\}}_{d\tilde{W}_1(t)}$$

$$d\tilde{x}_1^{(2)}(t) = \tilde{x}_2^{(2)}(t)dt$$

$$d\tilde{x}_2^{(2)}(t) = [-C_2\tilde{x}_2^{(2)} - (K_2 + K_3)\tilde{x}_1^{(2)} + K_2\tilde{x}_1^{(1)}]dt + \sigma_2 d\tilde{W}_2(t) \quad (30a)$$

where $\tilde{W}_2(t)=W_2(t)$. The scalar SDE governing the Radon-Nikodym derivative is

$$dy(t) = -\sigma_1^{-1}[\alpha(\tilde{x}_1^{(1)})^3 - \alpha(\tilde{x}_{1,i-1}^{(1)})^3]y d\tilde{W}_1(t) \quad (30b)$$

Solutions for $\tilde{X}(t)$ and the Radon-Nikodym derivative turn out to be

$$\begin{aligned} \tilde{X} = & \exp\left(\tilde{A} \int_{t_{i-1}}^t ds\right) \tilde{X}_0 + \exp\left(\tilde{A} \int_{t_{i-1}}^t ds\right) \int_{t_{i-1}}^t \exp\left(-\tilde{A} \int_{t_{i-1}}^s ds_1\right) \\ & \times \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & -\alpha(\tilde{x}_{1,i-1}^{(1)})^3 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} ds + \exp\left(\tilde{A} \int_{t_{i-1}}^t ds\right) \\ & \times \int_{t_{i-1}}^t \exp\left(-\tilde{A} \int_{t_{i-1}}^s ds_1\right) \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \sigma_1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_2 \end{bmatrix} \begin{Bmatrix} 0 \\ d\tilde{W}_1(s) \\ 0 \\ d\tilde{W}_2(s) \end{Bmatrix} \end{aligned} \quad (31a)$$

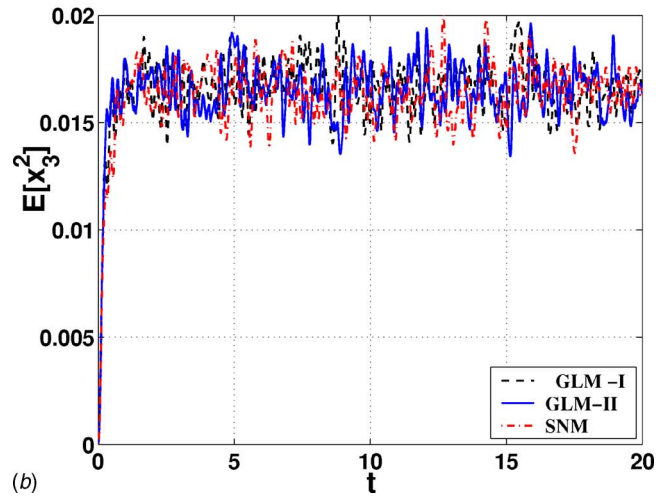
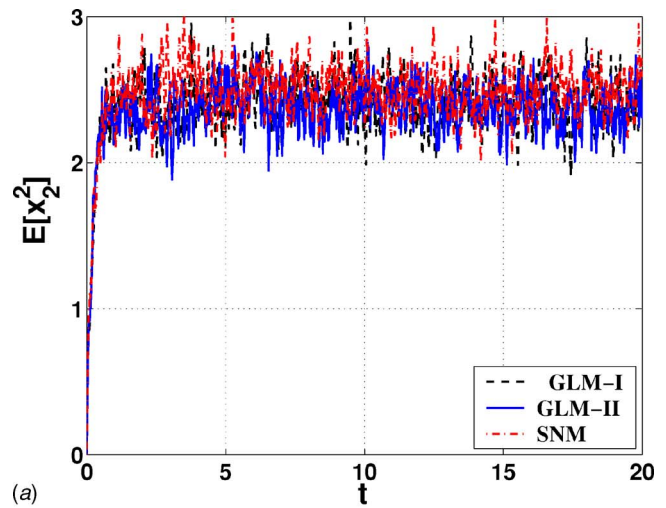


Fig. 11 The 2-DOF oscillator (Eq. (27))—second moment histories using GLM-I, GLM-II, and SNM: (a) $E[x_2^2]$ and (b) $E[x_3^2]$; $C_1=C_2=5.0$, $K_1=K_2=K_3=100.0$, $\alpha=100.0$, $\sigma_1=\sigma_2=5.0$, $h=0.01$, $\tilde{X}_0=\{0,0,0,0\}^T$

$$\begin{aligned} y(t) = & \exp\left(\int_{t_{i-1}}^t \{-\sigma_1^{-1}[\alpha(\tilde{x}_1^{(1)})^3 - \alpha(\tilde{x}_{1,i-1}^{(1)})^3]\} d\tilde{W}_1(s)\right. \\ & \left. - \frac{1}{2} \int_{t_{i-1}}^t \{\sigma_1^{-1}[\alpha(\tilde{x}_1^{(1)})^3 - \alpha(\tilde{x}_{1,i-1}^{(1)})^3]\}^2 ds\right) \end{aligned} \quad (31b)$$

For $i=1$, we have the initial condition $\tilde{X}_0=\{x_{1,0}^{(1)}, x_{2,0}^{(1)}, x_{1,0}^{(2)}, x_{2,0}^{(2)}\}^T$ and the system coefficient matrix for the linearized vector SDE is given by

$$\tilde{A} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -(K_1 + K_2) & -C_1 & K_2 & 0 \\ 0 & 0 & 0 & 1 \\ K_2 & 0 & -(K_2 + K_3) & -C_2 \end{bmatrix}$$

Recall that the expectation of a given scalar function $f(\tilde{X}, t)$ may be computed as $E[f(\tilde{X}, t)] = E[y(t)f(\tilde{X}, t)]$.

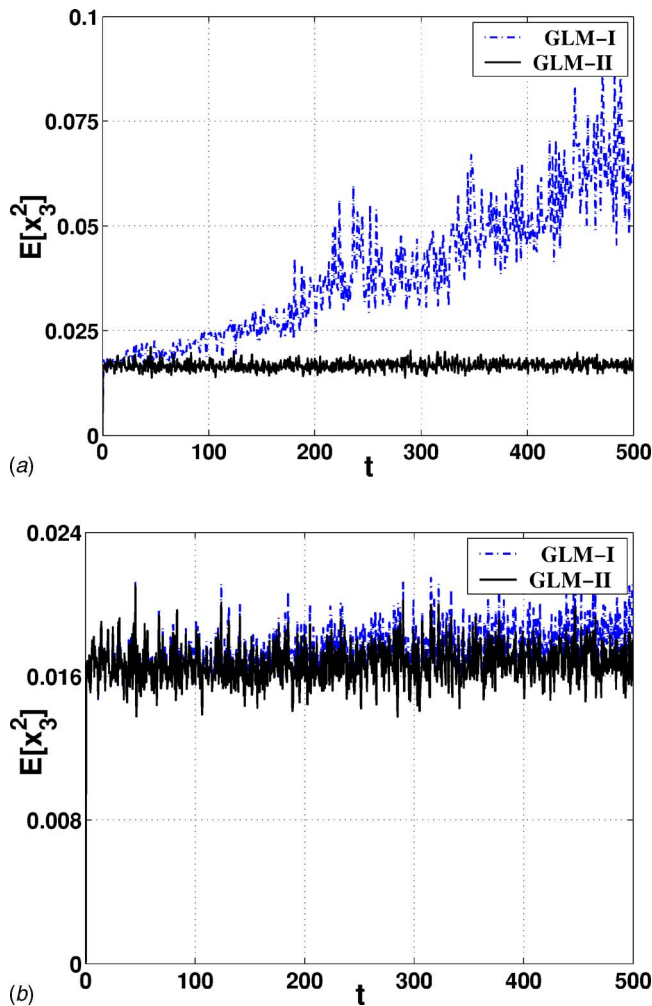


Fig. 12 The 2-DOF oscillator (Eq. (27))—plots of $E[x_3^2]$ using GLM-I and GLM-II for step sizes (a) $h=0.01$ and (b) $h=0.005$; $C_1=C_2=5.0$, $K_1=K_2=K_3=100.0$, $\alpha=100.0$, $\sigma_1=\sigma_2=5.0$, $\tilde{X}_0=\{0,0,0,0\}^T$

Figure 11 shows a couple of second moment histories via GLM-I and GLM-II over 0–20 s with a step size $h=0.01$. While the closeness of solutions through the two versions of the GLM appears to be impeccable over this interval, the power of GLM-II over GLM-I is brought out if the same histories are plotted over a much longer interval, as in Fig. 12(a). Here, we have used the same step size and plotted the histories of $E[x_3^2(t)]$ over $t \in [0, 500]$ via the two versions of the GLM. While GLM-II shows no numerical instability, the solution through GLM-I clearly exhibits an overflowing trend and becomes unacceptably inaccurate. However, as shown in Fig. 12(b), GLM-I works well even over such a long interval if the step size is reduced (in this case halved).

4 Conclusions

A weak form of linearization method, applicable to nonlinear oscillators under additive stochastic excitations, is developed based on a Girsanov transformation of probability measures. The proposed GLM suffers from no dimensionality curse and has the potential to achieve an accuracy of approximation that is unparalleled by any other existing method to the authors' knowledge. However, a major problem lies with handling the exponential correction term (martingale), which shows high sensitivity to how the

linearization is being attempted and often exhibits numerically spurious oscillations. As a result, the GLM-based solution often tends to overflow or underflow as time progresses. In order that such numerical obstacles are virtually removed, a modification is effected in the linearized form and the process of linearization is employed over small intervals of time, as in a numerical method. It is this latter strategy that imparts a numeric-analytic character to the GLM. Two different versions of the GLM are presented in this study. The first version requires that the argument of the exponential martingale be evaluated through a stochastic Taylor expansion, which in turn necessitates the evaluation of multiple stochastic integrals along with the attendant requirement that the nonlinear terms be sufficiently differentiable. The second version, on the other hand, employs a new Monte Carlo-based strategy and evaluates the exponential correction to (virtually) any desired order or accuracy. Moreover, it requires no Taylor expansion or differentiability assumptions on the nonlinear terms. The GLM is numerically illustrated for a few SDOF and MDOF nonlinear oscillators with polynomial nonlinearity. Comparisons of the GLM-based solutions with exact stationary solutions, whenever available, are also provided.

Since the GLM propagates the solution purely through exponential transformations, it is readily exploitable as a Lie group method to help preserve any dynamical invariants that may be of interest. The GLM is also modifiable to achieve importance sampling and variance reduction with high precision. These issues constitute interesting elements of a future study.

Acknowledgment

The authors thankfully acknowledge the financial support from the Board of Research in Nuclear Sciences (BRNS), Government of India, through Grant No. DAEO/MCV/DR/112.

Appendix

Note that the first integral $\int_0^t (-\sigma^{-1}[K_2\tilde{x}_1^3 - K_2\tilde{x}_{1,i}^3])d\hat{W}_1(s)$ on the right-hand side of Eq. (23) is an Ito integral and we will expand it using an Ito-Taylor expansion in terms of MSIs. By application of linearity property of Ito integral, we get

$$\begin{aligned} \int_0^t (-\sigma^{-1}[K_2\tilde{x}_1^3 - K_2\tilde{x}_{1,i}^3])d\hat{W}_1(s) &= -\sigma_1^{-1}K_2 \int_0^t (\tilde{x}_1^3)d\hat{W}_1(s) \\ &\quad + \int_0^t (\sigma^{-1}K_2\tilde{x}_{1,i}^3)d\hat{W}_1(s) \end{aligned} \quad (A1)$$

Now we write the first term on the right-hand side of Eq. (A1) as

$$-\sigma_1^{-1}K_2 \int_0^t (\tilde{x}_1^3)d\hat{W}_1(s) = -\sigma_1^{-1}K_2 \times SI_1$$

Applying Ito-Taylor expansion to SI_1 repeatedly, we get

$$\begin{aligned}
\Rightarrow SI_1 = & \int_0^t \tilde{x}_{10}^3 d\hat{W}_1(s) + 3 \int_0^t \int_0^s \tilde{x}_{10}^2 \tilde{x}_{20} ds_1 d\hat{W}_1(s) + 3 \int_0^t \int_0^s (-K_2) \tilde{x}_{10,i}^3 ds_1 d\hat{W}_1(s) + 6 \int_0^t \int_0^s \int_0^{s_1} \left[\tilde{x}_{10} \tilde{x}_{20}^2 + \int_0^{s_2} (\tilde{x}_{20}^3 + 2\tilde{x}_{10} \tilde{x}_{20} \{-K_1 \tilde{x}_{10} - C_1 \tilde{x}_{20} - K_2 \tilde{x}_{10,i}^3\}) \right. \\
& - C_1 \tilde{x}_{20} - K_2 \tilde{x}_{10,i}^3 \} + \tilde{x}_{10} \sigma_1^2) ds_3 + 2\sigma_1 \int_0^{s_2} \left(\tilde{x}_{10} \tilde{x}_{20} + \int_0^{s_3} (\tilde{x}_{20}^3 + \tilde{x}_{10} \{-K_1 \tilde{x}_{10} - C_1 \tilde{x}_{20} - K_2 \tilde{x}_{10,i}^3\}) ds_4 \right. \\
& \left. \left. + \tilde{x}_{10} \int_0^{s_3} \sigma_1 d\hat{W}_1(s_4) \right) d\hat{W}_1(s_3) \right] ds_2 ds_1 d\hat{W}_1(s) - 3K_1 \int_0^t \int_0^s \int_0^{s_1} \left[\tilde{x}_{10}^3 + 3 \int_0^{s_2} \tilde{x}_{10}^2 \tilde{x}_{20} ds_3 \right] ds_2 ds_1 d\hat{W}_1(s) \\
& - 3C_1 \int_0^t \int_0^s \int_0^{s_1} \left[\tilde{x}_{10}^2 \tilde{x}_{20} + \int_0^{s_2} [2\tilde{x}_{10} \tilde{x}_{20}^2 + \tilde{x}_{10}^2 \{-K_1 \tilde{x}_{10} - C_1 \tilde{x}_{20} - K_2 \tilde{x}_{10,i}^3\}] ds_3 + \sigma_1 \int_0^{s_2} \left(\tilde{x}_{10}^2 \right. \right. \\
& \left. \left. + 2 \int_0^{s_3} \tilde{x}_{10} \tilde{x}_{20} ds_4 \right) d\hat{W}_1(s_3) \right] ds_2 ds_1 d\hat{W}_1(s) + 3\sigma_1 \int_0^t \int_0^s \int_0^{s_1} \left[\tilde{x}_{10}^2 + 2 \int_0^{s_2} \left(\tilde{x}_{10} \tilde{x}_{20} + \int_0^{s_3} (\tilde{x}_{20}^3 + \tilde{x}_{10} \{-K_1 \tilde{x}_{10} - C_1 \tilde{x}_{20} - K_2 \tilde{x}_{10,i}^3\}) ds_4 \right. \right. \\
& \left. \left. + \tilde{x}_{10} \int_0^{s_3} \sigma_1 d\hat{W}_1(s_4) \right) ds_3 \right] d\hat{W}_1(s_2) ds_1 d\hat{W}_1(s) + \dots
\end{aligned} \tag{A2}$$

Now, the other integral on the right-hand side of Eq. (23) may also be expanded in a similar way as

$$\begin{aligned}
\frac{1}{2} \int_0^t (-\sigma^{-1} [K_2 \tilde{x}_1^3 - K_2 \tilde{x}_{1,i}^3])^2 ds = & \frac{1}{2} \sigma_1^{-2} K_2^2 \int_0^t \tilde{x}_1^6 ds + \frac{1}{2} \sigma_1^{-2} K_2^2 \tilde{x}_{1,i}^6 \int_0^t ds - \sigma_1^{-2} K_2^2 \tilde{x}_{1,i}^3 \int_0^t \tilde{x}_1^3 ds = \frac{1}{2} \sigma_1^{-2} K_2^2 \times SI_2 + \frac{1}{2} \sigma_1^{-2} K_2^2 \tilde{x}_{1,i}^6 t - \sigma_1^{-2} K_2^2 \tilde{x}_{1,i}^3 \\
& \times SI_3
\end{aligned}$$

Applying Ito-Taylor expansions to SI_2 and SI_3 repeatedly, we get

$$\begin{aligned}
SI_2 = & \int_0^t \tilde{x}_{10}^6 ds + 6 \int_0^t \int_0^s \tilde{x}_{10}^5 \tilde{x}_{20} ds_1 ds + 6 \int_0^t \int_0^s (-K_2) \tilde{x}_{10,i}^3 ds_1 ds + 30 \int_0^t \int_0^s \int_0^{s_1} \left[\tilde{x}_{10}^4 \tilde{x}_{20}^2 + \int_0^{s_2} (4\tilde{x}_{10}^3 \tilde{x}_{20}^3 + 2\tilde{x}_{10}^4 \tilde{x}_{20} \{-K_1 \tilde{x}_{10} - C_1 \tilde{x}_{20} \right. \\
& \left. - K_2 \tilde{x}_{10,i}^3\} + \tilde{x}_{10}^4 \sigma_1^2) ds_3 + \int_0^{s_2} 2\tilde{x}_{10}^4 \tilde{x}_{20} \sigma_1 d\hat{W}_1(s_3) \right] ds_2 ds_1 ds - 6K_1 \int_0^t \int_0^s \int_0^{s_1} \left[\tilde{x}_{10}^6 + 6 \int_0^{s_2} \tilde{x}_{10}^5 \tilde{x}_{20} ds_3 \right] ds_2 ds_1 ds \\
& - 6C_1 \int_0^t \int_0^s \int_0^{s_1} \left[\tilde{x}_{10}^5 \tilde{x}_{20} + \int_0^{s_2} [5\tilde{x}_{10}^4 \tilde{x}_{20}^2 + \tilde{x}_{10}^4 \{-K_1 \tilde{x}_{10} - C_1 \tilde{x}_{20} - K_2 \tilde{x}_{10,i}^3\}] ds_3 + \int_0^{s_2} \tilde{x}_{10}^5 \sigma_1 d\hat{W}_1(s_3) \right] ds_2 ds_1 ds + 3\sigma_1 \int_0^t \int_0^s \int_0^{s_1} \left[\tilde{x}_{10}^2 \right. \\
& \left. + \int_0^{s_2} 2\tilde{x}_{10} \tilde{x}_{20} ds_3 \right] d\hat{W}_1(s_2) ds_1 ds + \dots
\end{aligned} \tag{A3}$$

$$\begin{aligned}
SI_3 = & \int_0^t \tilde{x}_{10}^3 ds + 3 \int_0^t \int_0^s \tilde{x}_{10}^2 \tilde{x}_{20} ds_1 ds + 3 \int_0^t \int_0^s (-K_2) \tilde{x}_{10,i}^3 ds_1 ds + 6 \int_0^t \int_0^s \int_0^{s_1} \left[\tilde{x}_{10} \tilde{x}_{20}^2 + \int_0^{s_2} (\tilde{x}_{20}^3 + 2\tilde{x}_{10} \tilde{x}_{20} \{-K_1 \tilde{x}_{10} - C_1 \tilde{x}_{20} - K_2 \tilde{x}_{10,i}^3\} \right. \\
& \left. + \tilde{x}_{10} \sigma_1^2) ds_3 + \int_0^{s_2} 2\tilde{x}_{10} \tilde{x}_{20} \sigma_1 d\hat{W}_1(s_3) \right] ds_2 ds_1 ds - 3K_1 \int_0^t \int_0^s \int_0^{s_1} \left[\tilde{x}_{10}^3 + 3 \int_0^{s_2} \tilde{x}_{10}^2 \tilde{x}_{20} ds_3 \right] ds_2 ds_1 ds - 3C_1 \int_0^t \int_0^s \int_0^{s_1} \left[\tilde{x}_{10}^2 \tilde{x}_{20} \right. \\
& \left. + \int_0^{s_2} [2\tilde{x}_{10} \tilde{x}_{20}^2 + \tilde{x}_{10}^2 \{-K_1 \tilde{x}_{10} - C_1 \tilde{x}_{20} - K_2 \tilde{x}_{10,i}^3\}] ds_3 + \int_0^{s_2} \tilde{x}_{10}^2 \sigma_1 d\hat{W}_1(s_3) \right] ds_2 ds_1 ds + 3\sigma_1 \int_0^t \int_0^s \int_0^{s_1} \left[\tilde{x}_{10}^2 \right. \\
& \left. + \int_0^{s_2} 2\tilde{x}_{10} \tilde{x}_{20} ds_3 \right] d\hat{W}_1(s_2) ds_1 ds + \dots
\end{aligned} \tag{A4}$$

The MSIs present in SI_3 are the same as those in SI_2 . The MSIs are zero-mean Gaussian variables. See Roy and Dash [9] for details of evaluating the covariance matrix for MSIs.

References

- [1] Lin, Y. K., and Cai, G. Q., 1988, "Exact Stationary Response Solution for Second Order Nonlinear Systems Under Parametric and External White Noise Excitations: Part II," *ASME J. Appl. Mech.*, **55**, pp. 702–705.
- [2] Kloeden, P. E., and Platen, E., 1999, *Numerical Solution of Stochastic Differential Equations*, Springer, New York.
- [3] Milstein, G. N., 1995, *Numerical Integration of Stochastic Differential Equations*, Kluwer, Dordrecht.
- [4] Maruyama, G., 1955, "Continuous Markov Processes and Stochastic Equations," *Rend. Circ. Mat. Palermo*, **4**, pp. 48–90.
- [5] Gard, T. C., 1988, *Introduction to Stochastic Differential Equations*, Marcel Dekker, New York.
- [6] Rumelin, W., 1982, "Numerical Treatment of Stochastic Differential Equations," *SIAM (Soc. Ind. Appl. Math.) J. Numer. Anal.*, **19**(3), pp. 604–613.
- [7] Burrage, K., Burrage, P., and Tian, T., 2004, "Numerical Methods for Strong Solutions of Stochastic Differential Equations: An Overview," *Proc. R. Soc. London, Ser. A*, **460**(2041), pp. 373–402.
- [8] Tocino, A., and Vigo-Aguiar, J., 2002, "Weak Second Order Conditions for

- Stochastic Runge-Kutta Methods,” SIAM J. Sci. Comput. (USA), **24**(2), pp. 507–523.
- [9] Roy, D., and Dash, M. K., 2005, “Explorations of a Family of Stochastic Newmark Methods in Engineering Dynamics,” *Comput. Methods Appl. Mech. Eng.*, **194**(45–47), pp. 4758–4796.
- [10] Roy, D., 2006, “A Family of Weak Stochastic Newmark Methods for Simplified and Efficient Monte Carlo Simulations of Oscillators,” *Int. J. Numer. Methods Eng.*, **67**(3), pp. 364–399.
- [11] Burrage, K., and Tian, T., 2004, “Implicit Stochastic Runge-Kutta Methods for Stochastic Differential Equations,” *BIT Numer. Math.*, **44**(1), pp. 21–39.
- [12] Milstein, G. N., Platen, E., and Schurz, H., 1998, “Balanced Implicit Methods for Stiff Stochastic Systems,” *SIAM (Soc. Ind. Appl. Math.) J. Numer. Anal.*, **35**(3), pp. 1010–1019.
- [13] Oksendal, B., 2004, *Stochastic Differential Equations—An Introduction With Applications*, 6th ed., Springer, New York.
- [14] Socha, L., 2005, “Linearization in Analysis of Nonlinear Stochastic Systems: Recent Results—Part I: Theory,” *Appl. Mech. Rev.*, **58**(3), pp. 178–205.
- [15] Socha, L., 2005, “Linearization in Analysis of Nonlinear Stochastic Systems: Recent Results—Part II: Applications,” *Appl. Mech. Rev.*, **58**(5), pp. 303–353.
- [16] Socha, L., and Pawleta, M., 1994, “Corrected Equivalent Linearization,” *Mach. Dyn. Probl.*, **7**, pp. 149–161.
- [17] Elishakoff, I., and Colojani, P., 1997, “Stochastic Linearization Critically Re-examined,” *Chaos, Solitons Fractals*, **8**(12), pp. 1957–1972.
- [18] Crandall, S. H., 2001, “Is Stochastic Equivalent Linearization a Subtly Flawed Procedure,” *Probab. Eng. Mech.*, **16**(2), pp. 169–176.
- [19] Falsone, G., and Elishakoff, I., 1994, “Modified Stochastic Linearization Technique for Coloured Noise Excitation of Duffing Oscillator,” *Int. J. Non-Linear Mech.*, **29**(1), pp. 65–69.
- [20] Apetaur, M., and Opicka, F., 1983, “Linearization of Nonlinear Stochastically Excited Dynamic Systems,” *J. Sound Vib.*, **86**(4), pp. (563–585).
- [21] Socha, L., 1999, “Statistical and Equivalent Linearization Techniques With Probability Density Criteria,” *J. Theor. Appl. Mech.*, **37**, pp. 369–382.
- [22] Anh, N. D., and Hung, L. X., 2003, “An Improved Criterion of Gaussian Equivalent Linearization for Analysis of Nonlinear Stochastic Systems,” *J. Sound Vib.*, **268**(1), pp. 177–200.
- [23] Kazakov, I. E., 1998, “An Extension of the Method of Statistical Linearization,” *Avtom. Telemekh.*, **59**, pp. 220–224.
- [24] Grundmann, H., Hartmann, C., and Waubke, H., 1998, “Structures Subjected to Stationary Stochastic Loadings. Preliminary Assessment by Statistical Linearization Combined With an Evolutionary Algorithm,” *Comput. Struct.*, **67**(1–3), pp. 53–64.
- [25] Iyengar, R. N., and Roy, D., 1996, “Conditional Linearization in Nonlinear Random Vibration,” *J. Eng. Mech.*, **122**(3), pp. 197–200.
- [26] Roy, D., 2000, “Exploration of the Phase-Space Linearization Method for Deterministic and Stochastic Nonlinear Dynamical Systems,” *Nonlinear Dyn.*, **23**(3), pp. 225–258.
- [27] Roy, D., 2001, “A New Numeric-Analytical Principle for Nonlinear Deterministic and Stochastic Dynamical Systems,” *Proc. R. Soc. London, Ser. A*, **457**(2007), pp. 539–566.
- [28] Roy, D., 2004, “A Family of Lower- and Higher-Order Transversal Linearization Techniques in Non-Linear Stochastic Engineering Dynamics,” *Int. J. Numer. Methods Eng.*, **61**(5), pp. 764–790.
- [29] Ibrahim, R. A., 1978, “Stationary Response of a Randomly Parametric Excited Nonlinear System,” *ASME J. Appl. Mech.*, **45**, pp. 910–916.
- [30] Zhu, W. Q., 1988, “Stochastic Averaging Methods in Random Vibration” *Appl. Mech. Rev.*, **41**, pp. 189–199.
- [31] Bernard, P., and Wu, L., 1998, “Stochastic Linearization: The Theory,” *J. Appl. Probab.*, **35**(3), pp. 718–730.
- [32] Rubinstein, R. Y., 1981, *Simulation and the Monte Carlo Method*, Wiley, New York.
- [33] Yazici, A., Atlas, I., and Ergenc, T., 2005, “2d Polynomial Interpolation: A Symbolic Approach with Mathematica,” *Lect. Notes Comput. Sci.*, **3482**, pp. 463–471.
- [34] Wang, R., and Zhang, Z., 2000, “Exact Stationary Solutions of the Fokker-Planck Equation for Nonlinear Oscillators Under Stochastic Parametric and External Excitations,” *Nonlinearity*, **13**(3), pp. 907–920.
- [35] Roy, D., and Dash, M. K., 2002, “A Stochastic Newmark Method for Engineering Dynamical Systems” *J. Sound Vib.*, **249**(1), pp. 83–100.
- [36] Roy, D., 2003, “A Weak Form of Stochastic Newmark Method With Applications to Engineering Dynamical Systems,” *Appl. Math. Model.*, **27**(6), pp. 421–436.

D. Karagiozova
Department of Mechanical Engineering,
Hong Kong University of Science and
Technology,
Clear Water Bay,
Kowloon, Hong Kong, China;
Institute of Mechanics,
Bulgarian Academy of Sciences,
Acad. G. Bonchev Street,
Block 4,
Sofia 1113, Bulgaria

T. X. Yu¹
e-mail: metxyu@ust.hk

Z. Y. Gao

Department of Mechanical Engineering,
Hong Kong University of Science and
Technology,
Clear Water Bay,
Kowloon, Hong Kong, China

Stress–Strain Relationship for Metal Hollow Sphere Materials as a Function of Their Relative Density

The stress–strain relationship for uniaxial compression of a metal hollow sphere material in large strains is obtained using a simplified model for the spheres' deformation within a 3D block assuming a hexagonal packing pattern. The yield strength and material strain hardening are obtained as functions of the relative density in two characteristic loading directions. The expression for the stress–strain relationship consisting of quadratic and linear terms with respect to the relative density is linked to the partitioning of the deformation energy during compression. The theoretical predictions are compared with limited test results on mild steel hollow sphere material and finite element simulation results obtained by our group. [DOI: 10.1115/1.2712235]

Keywords: metallic foam, plastic deformation, uniaxial loading, yield strength, strain hardening, relative density

1 Introduction

The properties of metallic cellular materials (honeycomb, foam, hollow sphere agglomerate) have been studied extensively in recent years [1] owing to the wider engineering applications in lightweight structures for impact energy absorption [2], acoustic wave attenuation, etc. A large number of experimental, numerical, and analytical studies on the behavior of cellular materials have been reported in the literature (e.g., see Refs. [3,4]). However, relatively little is published on the prediction of the properties of metal hollow sphere (MHS) agglomerates, especially with respect to their stress–strain behavior in large strains due to the complexity of the spheres' deformation within a material block.

Metal foams made from hollow spheres offer low density with reasonably good energy absorption and high strength to weight ratios. Sintered metal hollow sphere foams have a certain volume fraction of enclosed pore space inside the spheres, but also have interstitial porosity between the sintered spheres appearing as a mixed open/closed-cell cellular solid. The key physical attributes of the MHS material, which are relevant to their mechanical behavior, include the characteristic sphere dimensions (diameter and wall thickness), wall material properties, and the relative density, ρ^*/ρ_s . The latter is affected by both the relative density of individual spheres as well as the packing pattern of spheres in the cellular solid. Unlike foams that are processed using expansion methods or injection of molten metal, hollow sphere metal foams may be processed either as random or ordered [5], leading to isotropic or anisotropic macroscopic elastic–plastic behavior, respectively. In other words, it is likely that a higher control on the mechanical properties of the MHS materials can be achieved in comparison to the open or closed cell foams.

Due to the particular deformation mechanism of the spheres within a material block, a stress enhancement was experimentally observed [6,7] when the deformation progresses, which is not typical for the other types of open or closed cell metal foams. Among the few studies on the characterization of MHS materials,

most thorough analyses of the material properties are published in Refs. [8–11] but they are exclusively concerned with the elastic modulus and the initial yield strength despite the fact that these materials manifest a noticeable strain hardening [6,7] before densification. The hardening is an inherent property of the metal hollow sphere material due to the deformation mechanism of the spheres and reflects the hardening feature of the force–displacement characteristic of a single hollow sphere under uniaxial compression [12]. In this case, the assumption of a relatively flat stress plateau cannot adequately characterize the MHS material. The present study aims to obtain the stress–strain relationship in large strains and describe the material hardening as a function of the relative density. A rigid, perfect-plastic model is assumed for the base material, which is a reasonable approximation of typical mild steel in large strains.

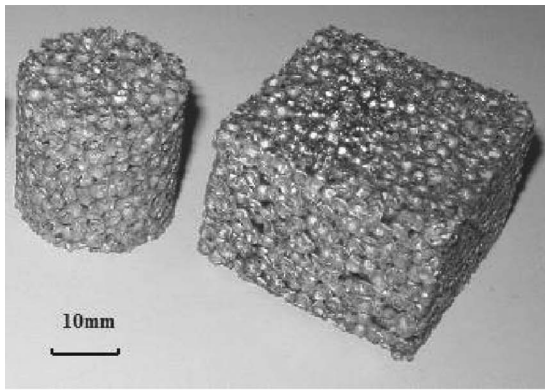
2 Simplified Analytical Model

2.1 A Model Analysis of the Crushing Mechanism Based on Representative Blocks. Although the real cellular materials have a somewhat random arrangement of the cells, it is still essential to analyze the major mechanisms of deformation in regularly packed spheres in order to understand the influence of the various parameters on the response of these materials to different loading conditions. The actual structure of a three-dimensional (3D) array of hollow spheres can be replaced by equivalent structural units with regular arrangements, which allow for analytical modeling.

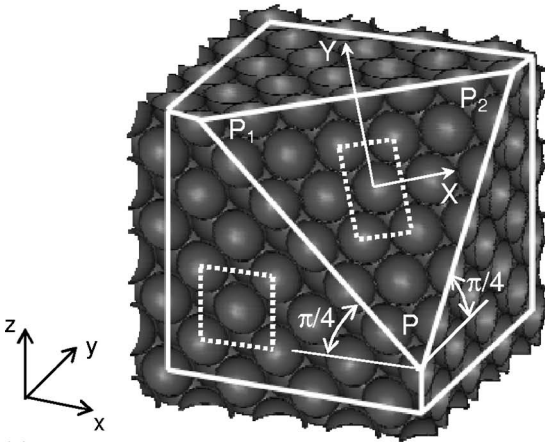
For the present purpose of an analytical formulation, it is anticipated that a hexagonal arrangement of the spheres could be a realistic idealization of a MHS material formed by sintering technology (Fig. 1(a)). Figure 1(b) shows a group of hexagonally packed spheres where two characteristic loading planes are apparent and structurally representative blocks in this material are marked in the PP_1P_2 plane and the x - z plane, respectively. Due to the symmetry, these blocks are further reduced as shown in Figs. 2(a) and 2(b). The model representing loading perpendicular to the X - Y plane (in the Z direction, FCC pattern) consists of one half rigid hemisphere on the top and three deformable 1/3 hemispheres underneath it, while symmetric boundary conditions are applied (Fig. 2(a)). The model representing loading in the z direc-

¹Corresponding author.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received April 24, 2006; final manuscript received November 3, 2006. Review conducted by Antoinette Maniatty.



(a)



(b)

Fig. 1 (a) MHS material; and (b) Hexagonally packed spheres

tion (HCP pattern) consists of 1/2 rigid hemisphere on the top and four deformable 1/4 hemispheres underneath it when assuming symmetric boundaries as well (Fig. 2(b)).

The large deformation of a thin-walled sphere or a single spherical shell, known as a typical post-buckling problem, has received attention since the 1960s; see e.g., Leckie and Penny [13], Morris and Calladine [14], Updike [15], and de Oliveira and

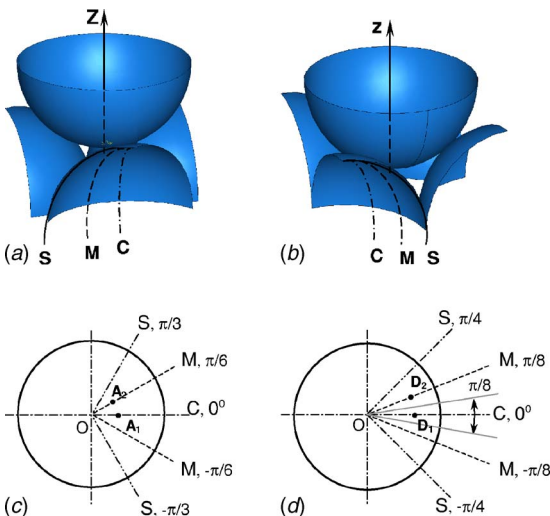


Fig. 2 Structural representative blocks: (a) compression in the Z direction; (b) compression in the z direction; and (c), (d) characteristic lines of the deformation mechanism

Wierzbicki [16]. However, although many interesting phenomena have been explored for the quasi-static and dynamic deformation of 1D and 2D ring arrays [17,18], very few attempts have been made to study the 1D or 2D sphere arrays. Recently, Ruan et al. [12] presented experimental and theoretical analyses on the compression of 1D and 2D arrays of hollow spheres using nonconnected ping-pong balls, and showed that the load–deformation curves of ball arrays with either square or hexagonal packing patterns can be predicted by superposing a number of 1D chains.

The deformation of a 3D array of hollow spheres in the MHS material obviously depends on the connections between the spheres and it is different for each packing pattern [8–11]. The solid connection between the spheres makes the response unlike the simple snap-through deformation in a single sphere or in an array of nonconnected spheres [12] by manifesting much higher initial strength. The full modeling of the deformation process certainly involves numerical simulations using large models [8–11], which makes the analysis in large strains difficult.

Some simplifying assumptions are made in Ref. [19] in order to facilitate an analytical analysis of the deformation of MHS materials with a regular structure. A structural approach using the concept of an equivalent structure is employed to describe the large plastic deformations during the post-collapse process under compression when the spheres have point connections between them. In the theoretical analysis, the deformation of the representative block is analyzed when assuming that the upper layer consists of undeformable spheres while the spheres in the lower layer are deformable.

Due to the symmetry, only one-third or one-quarter of the deformable hemisphere is considered and the deformation is then described by the deformation of three characteristic lines: central line (*C*), middle line (*M*), and symmetry line (*S*), see Figs. 2(c) and 2(d). In these figures, points D_1 and D_2 mark the positions on the lower sphere, where the upper one makes its first touch on the *C* and *M* lines, respectively. For both structural models, the compression causes deformation of the central line (*C* line) at the first instant of contact between the layers followed by the deformation of the middle line (*M* line) while the *S* line is never in contact with the upper sphere during deformation. The deformation of the *S* line is caused by the downward displacement of the sphere with fixed connection points between the neighboring spheres. Variation of the contact angles, Φ_T , and contact radii, R_T , are shown in Figs. 3(c) and 3(d) for the FCC and HCP patterns, respectively. The contact radii characterize the cross section of the upper sphere obtained by an intersection with a vertical plane at the point of contact with the bottom (deformable) sphere. The contact angles characterize the contact point on the deformable sphere in the plane of the corresponding characteristic line. The particular values for the *M* line, $R_T = R_M$ and $\Phi_T = \Phi_{0M}$, are marked in both figures.

The construction of the geometrical profiles of the deformed characteristic lines due to the vertical displacement, δ , is shown in Fig. 3. It is anticipated that the formation of rolling plastic hinges [14] for lines *C*, *M*, *S* is the basic mechanism of deformation. An additional stationary hinge is formed at the equator of the sphere at the deformed *S* line. It is assumed further that the radii of the concave and convex curves with centers O_1 and O_2 are proportional, i.e., $r_{O1} = k_r r_{O2}$. The condition for unextensional line *S* is satisfied by varying the hinge angle α while the same condition for characteristic lines *C* and *M* is satisfied by varying the contact point between the spheres (point *E* in Fig. 3(b) defined by angle ψ). The details for the calculation of the deformed shapes of these lines are given in Ref. [19].

The following dependences of the plastic energy, D , during compression of the representative blocks hold

$$D(C) = f_C(k_{rC}, \psi_C), \quad (1a)$$

$$D(M) = f_M(k_{rM}, \psi_M), \quad (1b)$$

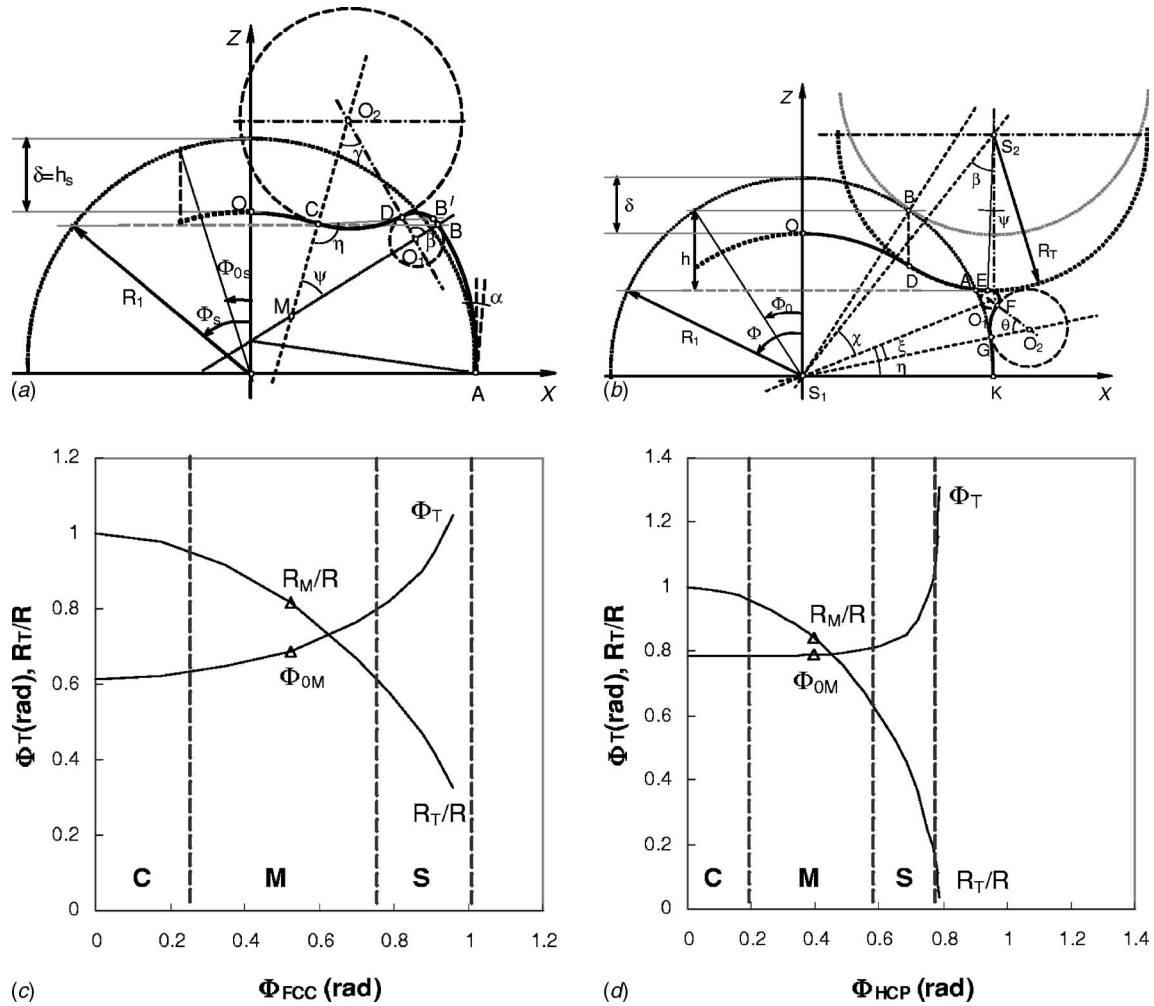


Fig. 3 Description of the deformed shapes of the characteristic lines: (a) *S* line; (b) *C* line and *M* line (Φ_0 and Φ stand for the initial and current angles associated with either *C* line or *M* line); and (c), (d) variation of the radii of the contact circles and contact angles for the FCC and HCP pattern, respectively; the sectors on a sphere associated with the *C*, *M*, and *S* line are marked

$$D(S) = f_S(k_{rS}, \alpha) \quad (1c)$$

where the free parameters for the radii of the rolling hinges, k_{rC} , k_{rM} , k_{rS} , allow for the minimization of the deformation energy under the general assumptions of the model.

The proposed model for deformation of a sphere by a snap-through mechanism at different positions on the surface simplifies the actual deformation of a sphere while retaining the characteristic features of the hollow sphere response inside a MHS material when subjected to uniaxial compression.

2.2 Shapes of the Characteristic Lines due to Uniaxial Compression of the Representative Blocks. The definitions for the engineering stress and strain of the MHS material are based on the response of a representative block to a uniaxial compression in the corresponding direction. The elastic deformations of the spheres are neglected and rigid, perfectly plastic property is assumed for the base material. Since the input energy to the system, W , (i.e., the work done by the external forces) is equal to the plastic dissipation energy, D

$$W = D, \quad (2a)$$

$$W = F\delta, \quad (2b)$$

where δ is the axial shortening of the representative block in the direction of loading. The collapse force, F , is determined as

$$F = dD/d(\delta). \quad (3)$$

It is assumed that the plastic energy dissipated due to the deformation of a hemisphere is described by the deformation of those three characteristic lines when the plastic strains are spread within segments $\pi/6$ for the representative block in Fig. 2(a) and $\pi/8$ for the representative block in Fig. 2(b), respectively. The variations of the contact angles Φ_T and contact radii, R_T , within the corresponding segments (*C*, *M*, and *S*) shown in Figs. 3(c) and 3(d) imply that the selected lines are representative for those segments. Therefore, the dissipated energy is determined as

$$D = f(D_C, D_M, D_S) = 2\pi R \frac{D_C + 2D_M + D_S}{4} = \pi R \frac{D_C + 2D_M + D_S}{2} \quad (4)$$

in either direction of uniaxial loading of the MHS material with hexagonal packing, where

$$D_C = D_{Cm} + D_{Cb}, \quad (5a)$$

$$D_M = D_{Mm} + D_{Mb}, \quad (5b)$$

$$D_S = D_{Sm} + D_{Sb} \quad (5c)$$

are the corresponding plastic energies. Subscripts *m* and *b* refer to

the bending and membrane energy, respectively.

The bending energies associated with S , C , and M lines are

$$D_{Sb}/(2\pi R) = M_0 \left[\alpha + \left(\frac{1}{k_{rS}\bar{r}_S} + 1 \right) \int_{CD} \bar{r}_c d\bar{s} + \left(\frac{1}{\bar{r}_S} - 1 \right) \int_{DB} \bar{r}_c d\bar{s} \right] \quad (6a)$$

$$D_{Cb}/(2\pi R) = M_0 \left[\frac{2}{R} \int_{DE} r_c d\bar{s} + \left(\frac{1}{k_{rC}r_C} - \frac{1}{R} \right) \int_{EF} r_c d\bar{s} + \left(\frac{1}{r_C} + \frac{1}{R} \right) \int_{FG} r_c d\bar{s} \right], \quad (6b)$$

$$D_{Mb}/(2\pi R) = M_0 \left[\left(\frac{1}{R} + \frac{1}{R_2} \right) \int_{DE} r_c d\bar{s} + \left(\frac{1}{k_{rM}r_M} - \frac{1}{R} \right) \int_{EF} r_c d\bar{s} + \left(\frac{1}{r_M} + \frac{1}{R} \right) \int_{FG} r_c d\bar{s} \right] \quad (6c)$$

where r_S , k_{rS} , r_S are the curvature radii of the segments $\hat{C}\hat{D}$ and $\hat{D}\hat{B}$ in Fig. 3(a); r_c is the current radius of the deformable sphere, $\bar{r}_c = r_c/R$; and $d\bar{s}$ denotes an infinitesimal arc length, $\bar{s} = s/R$. The membrane energies are calculated as

$$D_m/(2\pi R) = M_0 \frac{4R}{t} \int_S |\bar{r}_c - \bar{r}_0| d\bar{s} \quad (7)$$

where $\bar{r}_0 = r_0/R$, r_0 is the initial radius of the deformable sphere; and D_m stands for the energies $\hat{D}_{Sm}$, $\hat{D}_{Cm}$, or $\hat{D}_{Mm}$ associated with the corresponding segments $\hat{S}$.

Since the total deformation energy D is not uniquely defined but depends on the parameters k_{rC} , k_{rM} , k_{rS} , the collapse force is defined as

$$F = \left(\frac{dW}{d(\delta)} \right)_{\min} = \left[\frac{dD(k_{rC}, k_{rM}, k_{rS})}{d(\delta)} \right]_{\min} = \left(\left\{ \left[\frac{dD(k_{rC}, k_{rM}, k_{rS})}{d(\delta)} \right]_{\min|k_{rC}} \right\}_{\min|k_{rM}} \right)_{\min|k_{rS}} \quad (8)$$

when varying parameters k_r consecutively.

The hollow spheres within the representative block are constrained by the configuration of the neighboring spheres, so that the optimization of the crushing force has to be done under some restrictive conditions for the displacements associated with the characteristic lines. Due to the repeatability of the representative blocks, purely geometrical conditions should be satisfied

$$\eta - \xi \geq 0 \quad (9a)$$

$$(R + r) \cos \eta + r \leq R + g \quad \text{for } \eta = \xi \quad (9b)$$

for C and M lines (angles η and ξ are shown in Fig. 3(b)), where g is the gap between the spheres (Fig. 3(b)), namely

$$g_{CZ} = R(2 \sin \Phi_{0C} - 1), \quad (10a)$$

$$g_{MZ} = 0, \quad (10b)$$

$$g_{Cz} = R(1 - 1/\sqrt{2}), \quad (10c)$$

$$g_{Mz} = R[1 - \cos(\pi/8)] \quad (10d)$$

for the representative blocks shown in Figs. 2(a) and 2(b). The condition associated with S line (Fig. 3(a)) is

$$2R \sin \left(\frac{\pi}{4} - \frac{\Phi_S}{2} \right) \cos \left(\frac{\pi}{4} + \frac{\Phi_S}{2} + \alpha \right) \geq g \quad (11)$$

where

$$g_{SZ} = R(2 \sin \Phi_{0C} - 1), \quad (12a)$$

$$g_{Sz} = 0. \quad (12b)$$

The optimal values for the proportionality coefficients k_r are obtained as [19]

$$k_{rC} = 1/(0.01 + 0.6h_C/R) \quad (13a)$$

$$k_{rM} = 1/(0.01 + 0.4h_M/R) \quad (13b)$$

$$k_{rS} = 1/(0.01 + 0.2h_S/R) \quad (13c)$$

$$k_{rC} = 1/(0.01 + 0.4h_C/R) \quad (14a)$$

$$k_{rM} = 1/(0.01 + 0.4h_M/R) \quad (14b)$$

$$k_{rS} = 1/(0.01 + 1.6h_S/R) \quad (14c)$$

for compression in the Z and z directions, respectively, when satisfying conditions (10)–(13) and where

$$h_C = R(\cos \Phi_{0C} - \cos \Phi_C) \quad (15a)$$

$$h_M = R(\cos \Phi_{0M} - \cos \Phi_M) \quad (15b)$$

$$h_S = R(\cos \Phi_{0S} - \cos \Phi_S) \quad (15c)$$

are the displacements of the contact points (point D in Fig. 3(b) for C and M lines and point C in Fig. 3(a) for S line); Φ_0 and Φ are the initial and current angles associated with the lines.

The engineering stress for the MHS material is defined using the force–displacement behavior of the representative blocks

$$\sigma_{FCC}^* = F_{FCC}/A_{FCC} \quad (16a)$$

$$\epsilon_{FCC} = \delta_{FCC}/d_{0,FCC} \quad (16b)$$

$$\sigma_{HCP}^* = F_{HCP}/A_{HCP} \quad (17a)$$

$$\epsilon_{HCP} = \delta_{HCP}/d_{0,HCP} \quad (17b)$$

where $d_{0,FCC} = 2R\sqrt{2/3}$ is the initial vertical distance between the centers of two neighboring layers of spheres and $A_{FCC} = 2R^2\sqrt{3}$ is the area occupied by the representative block in Fig. 2(a), but $d_{0,HCP} = R\sqrt{2}$, $A_{HCP} = 4R^2$ for the representative block in Fig. 2(b). Deformed characteristic lines obtained for the optimal values of k_r are shown in Figs. 4(a) and 4(b) for 25% strain when compression is in the Z and the z directions, respectively.

2.3 FE Model for the Verification of the Crushing Mechanism. The finite element model shown in Fig. 2(a) is constructed to verify the analytically obtained shapes of the characteristic lines. The finite element code ABAQUS 6.4 is used and an automatic meshing with the triangular S3R shell elements is applied assuming a small bonding angle (1 deg approximately) between the spheres. (The bonding angle θ is referred to as a half of the connecting neck.) 5400–6800 nodes in one hollow sphere are used after the mesh convergence analysis. For the purpose of the analysis, it is assumed that the upper hemisphere of the model (Fig. 2(a)) is almost undeformable by prescribing a large elastic modulus, while the other parts of the block have elastic–plastic material properties. The representative block is placed between two rigid planes: the bottom one serves as a fixed boundary whereas the top one serves as the punch, i.e., only the displacement along the vertical direction is allowed, which is achieved by prescribing a constant low velocity to a large mass. The “hard contact” algorithm in ABAQUS is adopted to describe the normal properties of all the contacts (including the contacts between the

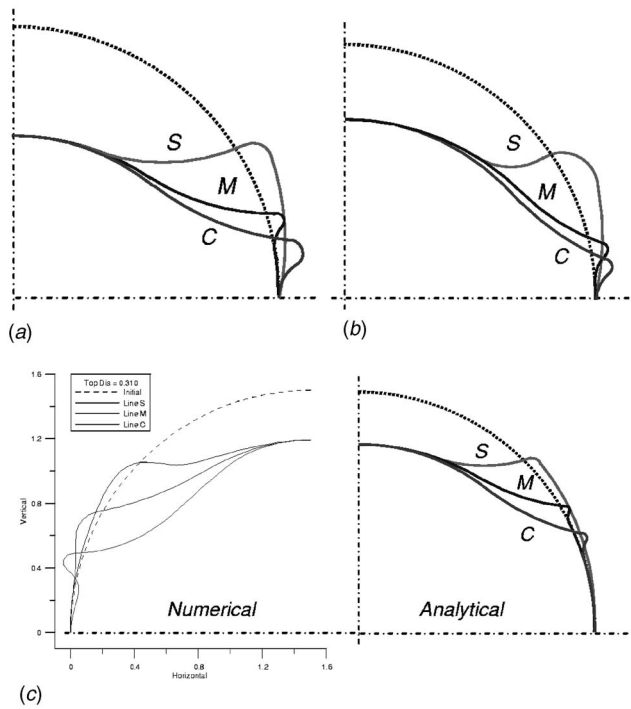


Fig. 4 Predicted shapes of the deformed characteristic lines at 25% strain: (a) compression in the Z direction (FCC); (b) compression in the z direction (HCP); and (c) comparison between the analytical model and finite element simulation (compression in the Z direction)

rigid plates and the spherical shells, the contacts between the spherical shell parts themselves), whereas the tangential behavior is defined as frictionless. Symmetry boundary conditions are applied with respect to the vertical side faces, which constrained the transverse expansion (or shrink) of the representative block.

The radius and wall thickness of the spheres are 1.5 mm and 0.049 mm, respectively. The numerically obtained shapes of the characteristic lines for loading in the Z direction and the model predictions are compared in Fig. 4(c) for equal displacements of the top of the deformable sphere, where a reasonable overall agreement is observed, although a somewhat stiffer response is analytically predicted. The stiffer response is mainly due to the assumption of inextensibility of the C line and the no slippage constrained for the M line in the model. During the actual contact between the spheres some slippage between the spheres does occur at the location of the M line.

2.4 Relationship Between the Nominal Stress and Material Density. The structural models shown in Fig. 3 are used to calculate the stress-strain response for different ratios t/R between 0.01 and 0.073; i.e., the relative density

$$\frac{\rho^*}{\rho_s} = 0.74 \left(\frac{t}{R} \right) \left[3 - 3 \left(\frac{t}{R} \right) + \left(\frac{t}{R} \right)^2 \right] \quad (18)$$

is between 0.021 and 0.15 for a hexagonal packing arrangement. Nondimensional stress-strain curves for several densities are shown in Fig. 5.

As an upper bound approach, the present model cannot predict correctly the stresses for small strains, so that the initial parts of these curves should be excluded from the analysis. As an illustration, the stress-strain curve obtained from the numerical simulation of the deformation of a sphere due to the compression by an undeformable sphere in the Z direction (Fig. 4(c)) is shown in Fig. 5(b). This figure demonstrates the agreement between the analytical model and numerical results for strains larger than about 0.03

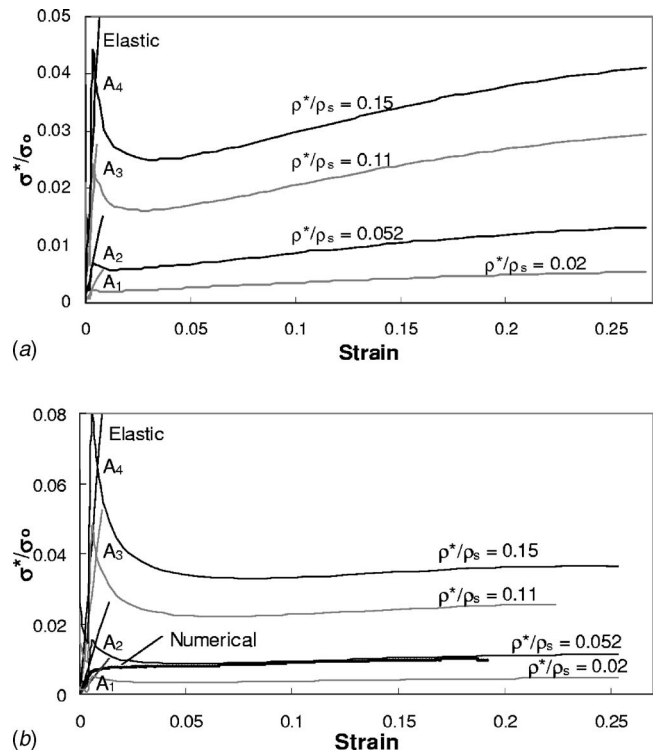


Fig. 5 Stress-strain curves: (a) compression in the z direction; and (b) compression in the Z direction; the numerical results (as described in Sec. 2.3) are shown for $\rho^*/\rho_c = 0.052$

for the particular case of $\rho^*/\rho_c = 0.052$. The elastic response given by Gasser et al. [11] is superposed in Figs. 5(a) and 5(b) on the predicted plastic stresses in order to estimate the validity of the present model. The elastic modulus is calculated as [11]

$$\frac{E^*}{E} = \left(0.118 + 0.826 \frac{t}{R} \right) \frac{t}{R} \quad (19)$$

for the present model with a point connection between the spheres. The elastic slopes in Figs. 5(a) and 5(b) are plotted for $E/\sigma_0 = 582.4$. In these figures, points A1–A4 mark the stresses corresponding to the initial strains, from which the rigid-plastic analysis begins to be valid. Similar to solid materials, the elastic strains due to uniaxial compression are proportional to the ratio E/σ_0 and one can see that the present model overestimates the stresses for small strains; however, the stress level, which corresponds to strains around 0.05, can be considered to give reliable stress estimates. In the following, the stress associated with the above strain is regarded as the yield strength, which corresponds to the yield strength for open or closed cell foams.

Since a structural approach is used to obtain the stress-strain response of the idealized material model, it is worthwhile to relate the nominal stress and initial material density using the characteristics of the deformation mechanism. The work done by the external forces is $F\delta = \sigma^* \epsilon R^3 C_1$ and hence the energy balance equation can be rearranged as

$$\sigma^* \epsilon = \frac{1}{C_1 R^3} \{ D_b + D_m \} = \frac{1}{C_1 R^3} \{ \sigma_0 t^2 R C_2(\epsilon) + \sigma_0 t R^2 C_3(\epsilon) \} \quad (20)$$

where C_1 is a characteristic constant of the particular representative block, ϵ is the uniaxial strain, and D_b and D_m are obtained from Eqs. (5) and (6). Equations (20) can be expressed in an incremental form

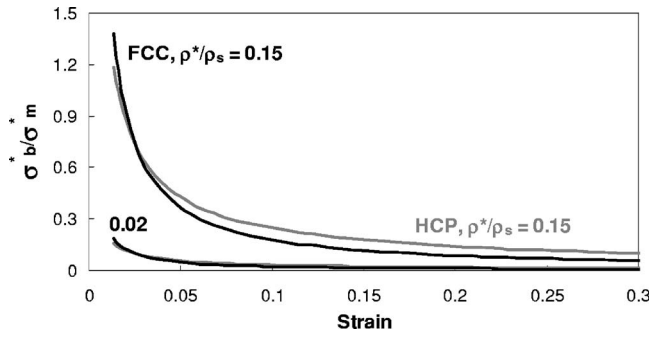


Fig. 6 Ratios between the (bending and membrane) stresses associated with the energy partitioning, σ_b^*/σ_m^* , in the two characteristic directions of loading

$$\sigma^* \Delta \epsilon = \frac{1}{R^3 C_1} \Delta \{ \sigma_0 t^2 R C_2(\epsilon) + \sigma_0 t R^2 C_3(\epsilon) \} \quad (21a)$$

and the stress-strain relationship can be decomposed into two terms

$$\frac{\sigma^*(\epsilon)}{\sigma_0} = \frac{d}{d\epsilon} \left[\frac{C_2(\epsilon)}{C_1} \left(\frac{t}{R} \right)^2 + \frac{C_3(\epsilon)}{C_1} \left(\frac{t}{R} \right) \right] \propto \left[P_1(\epsilon) \left(\frac{t}{R} \right)^2 + P_2(\epsilon) \left(\frac{t}{R} \right) \right] \quad (21b)$$

One can see that the term with $(t/R)^2$ corresponds to the contribution of the bending energy while the term with (t/R) is associated with the membrane energy.

It is further concluded that the ratio t/R can be approximated with sufficient accuracy by a linear function of the relative density as

$$t/R = 0.4983 \rho^*/\rho_s \quad (22)$$

where ρ^*/ρ_s is defined by Eq. (18). Then the stress can be decomposed into two parts proportional to $(\rho^*/\rho_s)^2$ and (ρ^*/ρ_s) when substituting Eq. (22) into Eq. (21b). The particular stress strain relationships for the two characteristic directions of loading are obtained as

$$\frac{\sigma_{FCC}^*(\epsilon)}{\sigma_0} = 0.071 \epsilon^{-0.6295} \left(\frac{\rho^*}{\rho_s} \right)^2 + 0.2674 \epsilon^{0.1608} \left(\frac{\rho^*}{\rho_s} \right), \quad \epsilon > 0.03 \quad (23a)$$

in the Z direction (FCC) and

$$\frac{\sigma_{HCP}^*(\epsilon)}{\sigma_0} = 0.0519 \epsilon^{-0.5958} \left(\frac{\rho^*}{\rho_s} \right)^2 + 0.4652 \epsilon^{0.4318} \left(\frac{\rho^*}{\rho_s} \right), \quad \epsilon > 0.03 \quad (23b)$$

in the z direction (HCP).

The constrained deformation of the spheres within the MHS material leads to an increase of the contribution of the membrane deformations while the deformation progresses. This is particularly true for spheres with smaller ratios t/R (smaller relative density). Ratios between the (bending and membrane) stresses associated with the energy partitioning, σ_b^*/σ_m^* , in the two characteristic directions of loading are shown in Fig. 6 for two material densities.

The proposed stress-strain relationships given by Eqs. (23) can also be used to verify the yield strength of tested MHS materials when substituting strain values between $0.03 < \epsilon < 0.1$. The stresses at $\epsilon=0.05$ for compression in the Z direction (FCC) and z direction (HCP) are depicted in Fig. 7 as functions of the relative density representing the yield strength in the characteristic directions of loading. Numerical simulations reported in Ref. [20] also established a relationship between the sphere characteristic ratio

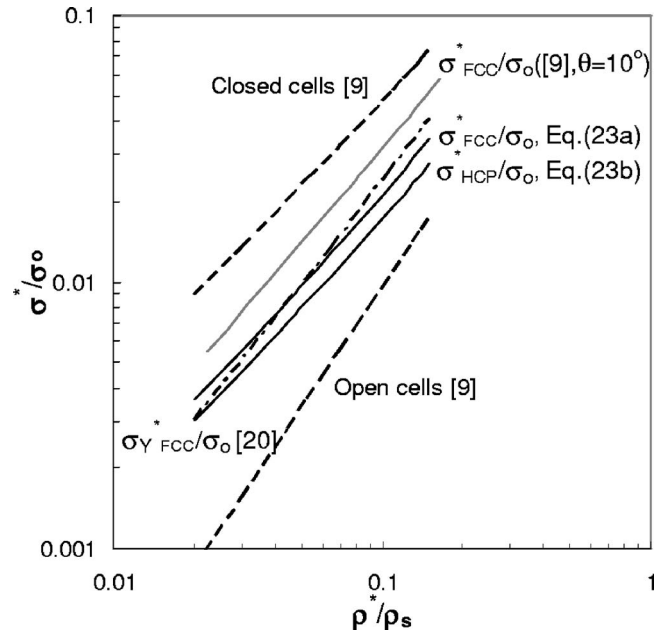


Fig. 7 Yield strength for various cellular solids; $\sigma_Y^* = \sigma^*$ ($\epsilon = 0.05$) for the MHS material

t/R and the yield strength as $\sigma_Y^*/\sigma_0 = 1.08(t/R)^{1.25}$ for loading in the Z direction (FCC) when modeling a finite connection of 5 deg between the spheres. It is anticipated that if the bonding neck is small and if the contribution of the bonding material to the relative density is negligible, the relation between the t/R ratio and the relative density, represented by Eq. (18), is valid. The yield strength according to Ref. [20] but as a function of (ρ^*/ρ_s) is shown in Fig. 7. The yield strength as a function of the relative density is also plotted for open and closed cell foam materials as stated in Ref. [9]. As expected, the yield strength of the MHS material in both directions of loading is between the theoretical predictions for the open and closed cells.

3 Comparison With Experimental Results and Discussion

3.1 Actual Material Density Versus Model Parameters.

Some quasi-static experimental results on MHS materials reported in Ref. [7] are used here to verify the consistency of the predictions for the stress-relative density dependences. The experimental stress-strain curve presented in Fig. 8(a) pertains to a quasi-static compression of a MHS material made of steel hollow spheres ($\sigma_0 = 344$ MPa, $R = 1.5$ mm, $t = 0.049$ mm) connected by sintering process. The measured averaged bonding angle was approximately 5 deg, which makes this material suitable for an analysis by the present model. A nearly hexagonal packing of the spheres is likely to be achieved by the sintering technique, so the theoretical relative density, calculated from Eq. (18), is obtained as $\rho^*/\rho_s = 0.071$. However, the actual relative material density is smaller, namely 0.052, due to some larger gaps between the spheres (Fig. 9(a)).

The reduced density of the MHS material due to the existing gaps/voids can be modeled by introducing "regular gaps" between the spheres as shown in Fig. 9(b) for compression in the Z direction (fcc). In this way, an adequate model for the actual material

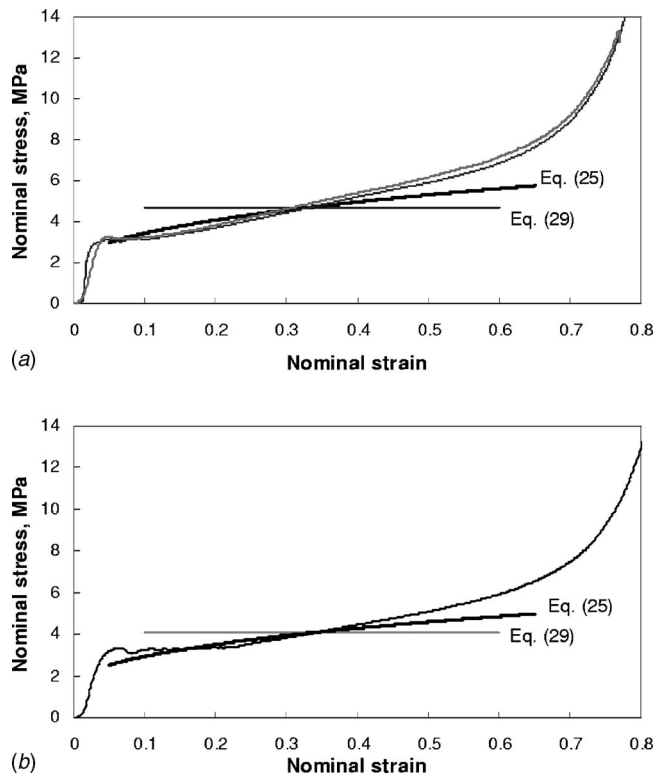


Fig. 8 Comparison between quasi-static experimental stress-strain curves and the theoretical predictions (Eq. (25)): (a) MHS material with $\rho^*/\rho_s=0.052$; and (b) MHS material with $\rho^*/\rho_s=0.045$

density can be achieved by controlling the distance between the spheres, Γ . The constraints for the deformation of the spheres become weaker, due to the increased gaps, while Eqs. (10a), (10b), and (12a) become

$$g_{CZ} = (R + \Gamma/2)(2 \sin \Phi_{0C} - 1), \quad (24a)$$

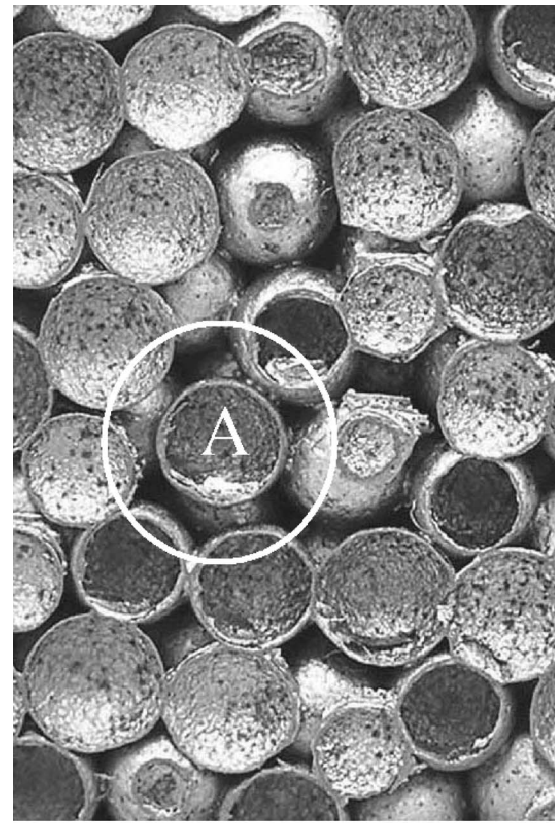
$$g_{MZ} = \Gamma/2, \quad (24b)$$

$$g_{SZ} = (R + \Gamma/2)(2 \sin \Phi_{0C} - 1). \quad (24c)$$

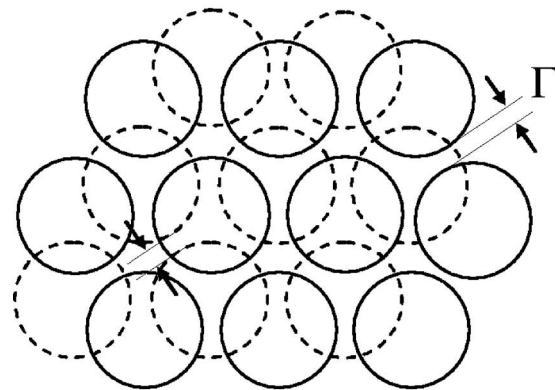
Variation of the gaps between the spheres with the reduction of the material density for a MHS material with original $\rho/\rho^*=0.071$ ($t/R=0.033$) is shown in Fig. 9(c) (the rising curve). The falling curve in this figure shows the increase of the contact radius associated with the M line ($R_T=R_M$) upon the sphere's compaction. (See the Appendix for the calculation of the contact radius R_T and contact angles.) A regular pattern for the arrangement of the spheres is assumed, so that the analysis based on a representative block remains valid when applying the optimization procedure given by Eq. (8). Note that the characteristic radius of a representative block for the stress definition is now $R + \Gamma/2$. Stress-strain curves corresponding to the initial and two reduced relative densities are shown in Fig. 10(a) for compression in the Z direction (fcc).

It should be noted, however, that the increased distance between the spheres can lead to inaccurate stress predictions, particularly for compression in the z direction (hcp) since the hollow spheres in the actual material will be less prone to deform according to the assumed mechanism. Therefore, it is worthwhile to use another method to reduce the density of the MHS model material, e.g., by using thinner model spheres (in a regular packing pattern) to achieve the actual density of the MHS material.

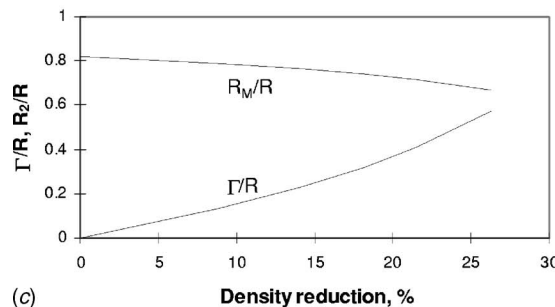
3.2 Comparison With Some Test Results. In order to compare the theoretical predictions with the test results for ρ^*/ρ_s



(a)



(b)



(c)

Fig. 9 MHS material with voids: (a) actual MHS material; (b) assumed regular packing of the spheres with gaps (FCC pattern); and (c) variation of the gap's size and contacting radius $R_T=R_M$ for reduced relative densities ($t/R=0.033$)

$=0.052$ but with spheres having $t/R=0.033$, a regular pattern with gaps of $0.54R$ is used. On the other hand, "model hollow spheres" with a reduced t/R ratio, which results in $\rho^*/\rho_s=0.052$, i.e., t/R

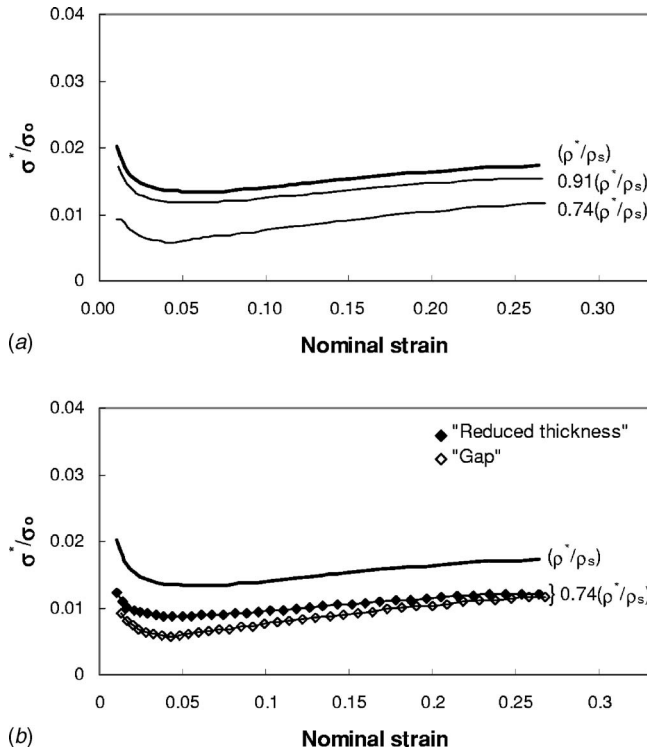


Fig. 10 Stress–strain curves for MHS materials with reduced densities (original density $\rho^*/\rho_s=0.071$); (a) reduction by increasing the gaps; and (b) stress–strain curves for MHS materials using two methods for density reduction

$=0.024$ instead of 0.033 , is considered as an alternative model for the density reduction. Comparison between the theoretical predictions based on the two models for the density reduction is presented in Fig. 10(b) for compression in the Z direction (fcc). One can see that both models predict similar trends for the stress–strain dependence although the “gap” model determines somewhat lower stresses, but the predicted stresses by both material models are relatively close for strains larger than 0.15 .

Since the two characteristic planes of loading identified for the model material are not clearly defined in the actual MHS material, it is appropriate to take the average from the stresses in both directions of loading in order to compare with the actual stress–strain curve. In this case, the second approach (using model hollow spheres with reduced thickness) gives more accurate results since the stress variation based on the mechanism of deformation described in Sec. 2.2 is accurate for both loading directions. Thus, it is suggested that the stress variation under uniaxial loading of a MHS material, $\sigma_{\text{MHS}}^*(\epsilon)$, is calculated as

$$\frac{\sigma_{\text{MHS}}^*(\epsilon)}{\sigma_0} = \frac{1}{2\sigma_0} [\sigma_{\text{FCC}}^*(\epsilon) + \sigma_{\text{HCP}}^*(\epsilon)] = \frac{1}{\sigma_0} \left[f_1(\epsilon) \left(\frac{\rho^*}{\rho_s} \right)^2 + f_2(\epsilon) \times \left(\frac{\rho^*}{\rho_s} \right) \right] \quad (25)$$

where

$$f_1 = 0.5(0.071\epsilon^{-0.6295} + 0.0519\epsilon^{-0.5958}), \quad (26a)$$

$$f_2 = 0.5(0.2674\epsilon^{0.1608} + 0.4652\epsilon^{0.4318}). \quad (26b)$$

The material stress–strain curve, $\sigma_{\text{MHS}}^*(\epsilon)$, according to Eqs. (25) and (26), is plotted in Fig. 8(a). An experimental stress–strain curve for uniaxial compression of the MHS material made of smaller hollow spheres, namely $R=0.9$ mm and $t=0.049$ mm ($\sigma_0=344$ MPa) [7], is presented in Fig. 8(b). The actual relative

density of this material is 0.045 , so that ratio $t/R=0.021$ is used in the theoretical model. The comparison between the predicted material stress–strain dependence and the measured one shows a reasonable overall agreement, although the structure of this material is less appropriate for an analysis by the present model since the measured relative density deviates more from the idealized hexagonal packing of the spheres.

The results discussed in this section suggest that the present structural approach for an analysis of the stress–strain response of MHS materials gives a reliable estimate for the material strength in relatively large strains. Note that Eqs. (25) and (26) can be used with sufficient accuracy only for MHS materials having small bonding necks between the spheres, or can be used as lower bound estimate for MHS materials with larger bonding angles. This conclusion is also evident from Fig. 7, where the yield strength for FCC spheres’ packing with 10 deg connecting angle is shown as obtained by Sanders and Gibson [9].

It is observed from the experimental results (Figs. 8(a) and 8(b)) that a notable essentially linear strain hardening occurs for $\epsilon_Y < \epsilon < \epsilon_D$. This material characteristic can be also expressed as a function of the relative density when using Eqs. (25) and (26) and anticipating that a linear hardening modulus, E_T , is an appropriate approximation

$$\frac{E_T^{\text{MHS}}}{\sigma_0} = \frac{1}{\sigma_0} \frac{\sigma_{\text{MHS}}^*(\rho^*/\rho_s, \epsilon_D) - \sigma_{\text{MHS}}^*(\rho^*/\rho_s, \epsilon_Y)}{(\epsilon_D - \epsilon_Y)} \quad (27a)$$

where ϵ_Y and ϵ_D are the strains at yield and densification, respectively. For the two materials examined the densification occurs at $\epsilon_D \approx 0.6$. The yield strength is taken at $\epsilon_Y=0.05$ and the strain hardening modulus is obtained as

$$\frac{E_T^{\text{MHS}}}{\sigma_0} = 0.159 \left(\frac{\rho^*}{\rho_s} \right)^{0.8724}. \quad (27b)$$

The hardening modulus characterizing the stress–strain relationship for each loading directions is

$$\frac{E_T^{\text{FCC}}}{\sigma_0} = 0.0314 \left(\frac{\rho^*}{\rho_s} \right)^{0.602}$$

and

$$\frac{E_T^{\text{HCP}}}{\sigma_0} = 0.332 \left(\frac{\rho^*}{\rho_s} \right)^{0.936} \quad (28)$$

for the Z and z direction, respectively.

It could be of interest for engineering applications related to the energy absorption estimate to obtain a constant stress in MHS material, which is averaged over the entire “plateau” strain interval up to the strain associated with the material densification. According to the above analysis, a characteristic constant stress averaged within the strain interval $\epsilon \in [0.05, 0.6]$ is obtained as

$$\frac{(\sigma^*)_{\text{MHS}}^{\text{av}}}{\sigma_0} = \frac{2}{\sigma_0} \int_{0.05}^{0.6} \sigma_{\text{MHS}}^* d\epsilon = 0.307 \left(\frac{\rho^*}{\rho_s} \right)^{1.051} \quad (29)$$

from which the constant stress values for the two actual MHS materials are shown in Figs. 8(a) and 8(b).

Equations (26), (27), and (29) are used to construct Fig. 11, which shows some predicted characteristics of MHS materials depending on the relative density. The results from the two tested MHS materials [7] are also marked showing good agreement with the results from the analytical model. It should be emphasized that although based on particular spheres’ arrangements, the above-mentioned equations provide averaged characteristics of MHS materials under uniaxial compression. This indicates that the relative density is a more important factor for characterization of the mechanical properties rather than the packing pattern and the variation of the t/R ratio, particularly for materials with low relative density. Furthermore, the relative density is an easy measur-

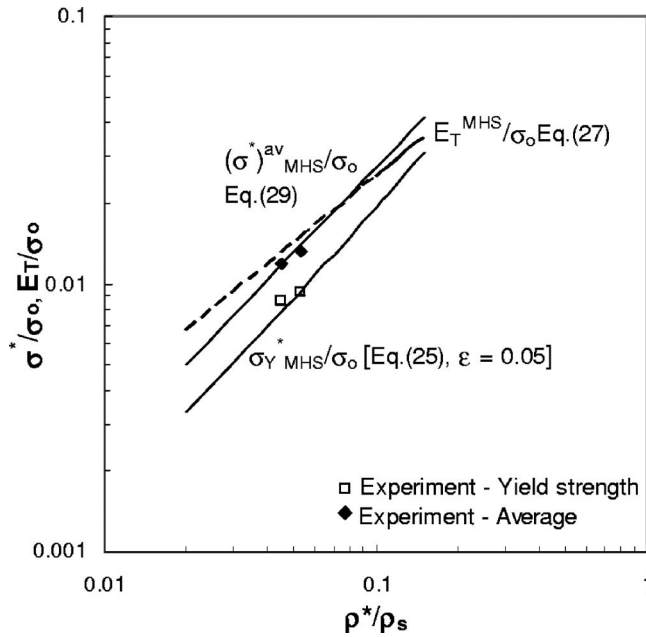


Fig. 11 Comparisons of the analytical model predictions for the yield strength, $\sigma_Y^{MHS}(\epsilon=0.05)$, and average stress, $(\sigma^*)_{MHS}^{av}$, with the available test results [7]

able parameter for MHS material characterization, which implies that Eqs. (25)–(29) can be easily and directly employed in engineering applications.

4 Conclusions

Relevant to important engineering applications, basic mechanical characteristics of MHS materials under uniaxial compression are obtained. Based on the analytical stress–strain dependencies under uniaxial compression, a relationship between the stress–strain curve and material density for MHS material is suggested for the first time. Moreover, a physical explanation is given why the expression for the stress–strain dependence contains both linear and quadratic terms. Further, a noncompact packing is analyzed by assuming regular gaps between the spheres.

The analytical and numerical predictions are compared with some experimental results for quasi-static compression of MHS materials made of mild steel. Reasonable agreement among the analytical predictions, numerical results, and the available experimental data is observed, but it is necessary to further study a wider range of the MHS material specifications in order to verify a more general validity of the proposed material model.

Acknowledgment

This study is part of a project supported by National Natural Science Foundation of China under a grant from Key Research Project No. 10532020 and the Project No. TH-1518/05 supported by the Bulgarian National Research Fund. The Fraunhofer Institute IFAM supplied the MHS materials. All this support is gratefully acknowledged.

Nomenclature

- D, D_m, D_b = total plastic energy, membrane energy, and bending energy, respectively
- $E_T^{FCC}, E_T^{HCP}, E_T^{MHS}$ = hardening moduli for the characteristic loading directions, Eqs. (27) and (28)
- h_C, h_M, h_S = displacement of the top point of a deformable sphere

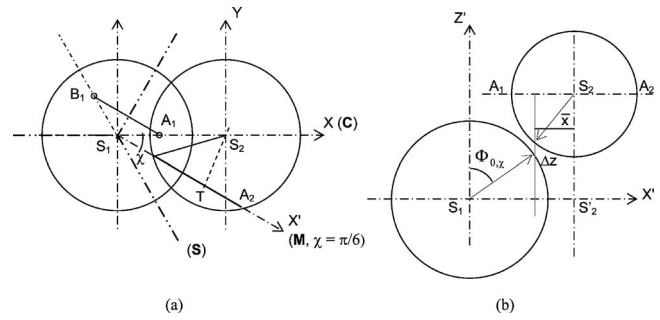


Fig. 12 The definitions of the characteristic angles for the C and M lines

k_{rC}, k_{rM}, k_{rS} = proportionality coefficient between two radii of rolling hinges along C, M, and S lines

M_0 = fully plastic bending moment of the wall, $\sigma_0 t^2/4$

R, t = radius and thickness of a sphere

R_1 = radius of the bottom sphere, $R_1=R$, Fig. 3

R_T = radius of the contact circle of the top sphere used to model the characteristic lines, Fig. 3(b)

W = external work

α = deformation at the stationary plastic hinge for S line, Fig. 3(a)

ϵ, ϵ_D = engineering strain and strain at densification

δ = shortening of a representative block

Φ, Φ_T = current and contact angle for FCC and HCP patterns

$\Phi_{0C}, \Phi_{0M}, \Phi_{0S}$ = initial angles corresponding to contact between the spheres at points on C, M, and S lines, respectively

Φ_C, Φ_M, Φ_S = current angles describing the deformation of C, M, and S lines, respectively

Γ = gap between the spheres, Fig. 9(b)

ψ = characteristic angle determining a contact point between the spheres during deformation of C and M lines, Fig. 3(b)

ρ^*/ρ_s = relative density of the MHS material

σ_0 = yield stress of the base material

σ_Y^* = yield strength of a MHS material

$\sigma_{MHS}^*(\epsilon)$ = predicted engineering stress for an actual MHS material, Eq. (25)

$(\sigma^*)_{MHS}^{av}$ = averaged engineering stress for an actual MHS material, Eq. (29)

C, M, S = subscripts associated with characteristic lines

Appendix: Contact Between the Spheres in the Direction of Loading

The initial contact between the spheres along the characteristic lines is determined by the angles Φ_{0M} and Φ_{0C} as shown in Fig. 3(b) while the variation of the displacement in the direction of loading is described by the variation of the angles Φ_S , Φ_M , and Φ_C . The definition of angles Φ_{0M} and Φ_{0C} are shown in Fig. 12 for compression of the MHS material in the Z direction. Φ_{0M} and Φ_{0C} depend on the angle $\chi \in [0, \pi/3]$, which is $\chi=0$ for the C line and $\chi=\pi/6$ for M line.

In Fig. 12, $S_1 S_2 = 2 \sin(35.2^\circ) + \Gamma$, $\Gamma=0$ for closely packed spheres; $S_1 S_2' = S_1 S_2 \cos \chi$ and the shortest distance between the contacting spheres is

$$\Delta z = 2R \cos \Phi_0 - \sqrt{R_T^2 - \bar{x}^2} - \sqrt{R^2 - (S_1 S_2' - \bar{x})^2} \quad (\text{A1})$$

where $R_T = R\sqrt{1 - 2 \sin^2 \chi}$.

The condition

$$\frac{d}{d\bar{x}}(\Delta z) = \frac{S_1 S_2' - \bar{x}}{\sqrt{R^2 - (S_1 S_2' - \bar{x})^2}} - \frac{\bar{x}}{\sqrt{R_T^2 - \bar{x}^2}} = 0 \quad (\text{A2})$$

gives the value of $\bar{x}$ at the contact point on the surface of the deformable sphere and the initial angle of contact is $\Phi_{0\chi} = \sin^{-1}[(S_1 S_2' - \bar{x})/R]$.

There is no contact between the spheres during deformation but the undeformed part of the S line is defined by angle Φ_{0S} , which is determined by the intersection between the sphere and a plane, which passes points A_1 , B_1 (contact points between the spheres at the central line) and is parallel to the direction of loading (Fig. 12).

In the analysis of the compression the initial angles of contact (for $\Gamma=0$) are taken as follows: $\Phi_{0C}=\pi/4$, $\Phi_{0M}=45.2$ deg, $\Phi_{0S}=\pi/6$ for compression in the z direction (hcp) and $\Phi_{0C}=35.26$ deg, $\Phi_{0M}=39.3$ deg for the compression in the Z direction with $\Phi_{0S}=16.78$ deg (fcc).

References

- [1] Gibson, L. J., and Ashby, M. F., 1997, *Cellular Solids: Structure and Properties*, 2nd ed., Cambridge University Press, Cambridge, UK.
- [2] Lu, G., and Yu, T. X., 2003, *Energy Absorption of Structures and Materials*, Woodhead Publishing Ltd., Cambridge, UK.
- [3] Hurysz, K. M., Clark, J. L., Nagle, A. R., Hardwicke, C. U., Lee, K. J., Cochran, J. K., and Sanders, T. H., 1998, "Steel and Titanium Hollow Sphere Foams," *Porous and Cellular Materials for Structural Applications, MRS Proceedings*, A. Evans, D. Schwartz, D. Shih, and H. Wadley, eds., Materials Research Society, Pittsburgh, PA, Vol. 521, pp. 191–203.
- [4] Clark, J. L., Hurysz, K. M., Lee, K. J., Cochran, J. K., and Sanders, T. H., 1999, "Stainless Steel Hollow Sphere Foams-Fabrication, Carburization, and Properties," *Metal Foams and Porous Metal Structures*, J. Banhart, M. F. Ashby, and N. A. Fleck, eds., MIT Publishers, Bremen, Germany, pp. 171–178.
- [5] Lim, T.-J., Smith, B., and McDowell, D. L., 2002, "Behavior of a Random Hollow Sphere Metal Foam," *Acta Mater.*, **50**, pp. 2867–2879.
- [6] Zhao, H., Elnasri, I., and Abdennadher, S., 2005, "An Experimental Study on the Behavior Under Impact Loading of Metallic Cellular Materials," *Int. J. Mech. Sci.*, **47**, pp. 757–774.
- [7] Gao, Z. Y., Yu, T. X., and Zhao, H., 2007, "Mechanical Behavior of Metallic Hollow Sphere (MHS) Materials: An Experimental Study," *J. Aerosp. Eng.*, submitted.
- [8] Sanders, W. S., and Gibson, L. J., 2003, "Mechanics of Hollow Sphere Foams," *Mater. Sci. Eng., A*, **347**, pp. 70–85.
- [9] Sanders, W. S., and Gibson, L. J., 2003, "Mechanics of BCC and FCC Hollow-Sphere Foams," *Mater. Sci. Eng., A*, **352**, pp. 150–161.
- [10] Gasser, S., Paun, F., Cayzele, A., and Bréchet, Y., 2003, "Uniaxial Tensile Elastic Properties of a Regular Stacking of Braze Hollow Spheres," *Scr. Mater.*, **48**, pp. 1617–1623.
- [11] Gasser, S., Paun, F., and Bréchet, Y., 2004, "Finite Elements Computation for the Elastic Properties of a Regular Stacking of Hollow Spheres," *Mater. Sci. Eng., A*, **379**, pp. 240–244.
- [12] Ruan, H. H., Gao, Z. Y., and Yu, T. X., 2006, "Crushing of Thin-Walled Spheres and Sphere Arrays," *Int. J. Mech. Sci.*, **48**, pp. 117–133.
- [13] Lekie, F. A., and Penny, P. K., 1968, "Plastic Instability of a Spherical Shell," *Engineering Plasticity*, J. Heyman, and F. A. Leckie, eds., Cambridge University Press, Cambridge, UK, pp. 401–411.
- [14] Morris, A., and Calladine, C. R., 1969, "The Local Strength of Thin Spherical Shell Loaded Radially Through a Rigid Boss," *Proceedings 1st International Conference on Pressure Vessel Technology, Delft, Netherlands*, I. Berman, ed., Pergamon, Oxford, UK, pp. 35–44.
- [15] Updike, D. P., 1972, "On the Large Deformation of a Rigid Plastic Spherical Shell Compressed by a Rigid Plate," *J. Eng. Ind.*, **94**, pp. 949–955.
- [16] De Oliveira, J. G., and Wierzbicki, T., 1982, "Crushing Analysis of Rotationally Symmetrical Plastic Shells," *J. Strain Anal. Eng. Des.*, **17**, pp. 229–236.
- [17] Reid, S. R., and Reddy, T. Y., 1983, "Experimental Investigation of Inertia Effects in One-Dimensional Metal Rings Systems Subjected to Impact - I. Fixed Ended Systems," *Int. J. Impact Eng.*, **1**, pp. 85–106.
- [18] Shim, V. P. W., and Stronge, W. J., 1986, "Lateral Crushing of Tightly Packed Arrays of Thin-Walled Metal Tubes," *Int. J. Mech. Sci.*, **28**, pp. 698–617.
- [19] Karagiozova, D., Yu, T. X., and Gao, Z. Y., 2006, "Modeling of MHS Cellular Solid in Large Strains," *Int. J. Mech. Sci.*, **48**, pp. 1273–1286.
- [20] Gao, Z. Y., Yu, T. X., and Karagiozova, D., 2007, "Finite Element Simulation on the Mechanical Properties of MHS Materials," *Acta Mech. Sinica*, **23**, pp. 65–75.

On the Normal Component of Centralized Frictionless Collision Sequences

Pieter J. Mosterman

The MathWorks, Inc.,
3 Apple Hill Dr.,
Natick, MA 01760

A typical assumption for rigid body collisions with multiple impact points is that all collisions occur simultaneously and are synchronized in their compression/expansion behavior, a useful assumption given the microscopic time scale at which collisions occur. In the case in which collisions are dependent upon one another, however, there is interaction between and within compression and expansion phases. Instead of treating the collisions as separate consecutive impacts or by activating all constraints at the same time, a rule is presented that orders the collisions as a sequence of interacting events at a point in time to handle the normal component of the collisions.

[DOI: 10.1115/1.2712237]

Keywords: hybrid modeling, systematic model abstractions, parameter abstraction, time scale abstraction, collision behavior

1 Introduction

This paper discusses a collision rule for the normal component of sequences of frictionless collisions based on the canonical case of centralized impact. Collision behavior of rigid bodies is an extensively studied problem in the field of mechanics, see, e.g., [1–5], and the overview [6], and can be divided into studies on single and multiple collisions. In the case of multiple collisions, the underlying assumption is that there are several contact points in the mechanical systems that exhibit collision behavior at the same point in time. Moreover, the decomposition of the collision effect into a compression and expansion phase is assumed to be synchronized, i.e., all collisions are simultaneously first in their compression and next in their expansion phase.

This assumption has proven accurate for collisions that are initiated by evaluating macroscopic dynamic behavior. However, if collisions become dependent such that one collision triggers another, the separate collision phases do not occur simultaneously even though in a macroscopic view they are still at the same point in time.

For example, consider the three colliding bodies in Fig. 1. It is assumed there is centralized frictionless impact, so only the behavior along the normal component needs to be evaluated. There are two contact points, one between m_1 and m_2 (contact 12) and one between m_2 and m_3 (contact 23). If upon collision of m_1 with m_2 the collision between m_2 and m_3 is evaluated simultaneously for certain restitution coefficients (ϵ), results may be generated that contradict a first-order linearized (spring-damper) approximation [7].

To make this intuitive, consider the bodies have equal mass m , and collisions are perfectly elastic; i.e., $\epsilon^{12} = \epsilon^{23} = 1$. At the time of collision, m_2 and m_3 are at rest, so there is no restitution of relative velocity. Therefore, if executed simultaneously, the collision between m_1 and m_2 appears to be between a body with mass m and a body with mass $2m$, and, consequently, m_1 will obtain a velocity with opposite sign. However, from experiments it is known that m_1 should (almost) come to rest, a result that is derived from evaluating the collisions in sequence. Because of the dependency between the collisions, for $\epsilon^{12} = \epsilon^{23} = 1$ it is critical to evaluate the collisions in sequence.

For other restitution coefficients, quite the opposite is true. Consider again a perfectly elastic collision between m_1 and m_2 ($\epsilon^{12} = 1$), but combined with a perfectly non-elastic collision between m_2 and m_3 ($\epsilon^{23} = 0$). In this case, a linear first-order approximation of the collision behavior between m_1 and m_2 does approximate that of a collision between one body with mass m and another with mass $2m$, and m_1 obtains a velocity with opposite sign. However, if the idealized collisions are evaluated in sequence, first m_1 transfers its momentum to m_2 , and next this is distributed over m_2 and m_3 to obtain equal velocities of half the initial magnitude.

These observations are discussed in detail in previous work [6,7] that explicates the necessity for either a pairwise evaluation or simultaneous evaluation, depending on the coefficients of restitution. This is a rather unsatisfying phenomenon, especially since it leaves unclear for which ϵ to switch between the different evaluations. This paper seeks to overcome this by introducing a pre-collision stage preceding dependent collisions, computed based on Newton's collision rule. This presents a rule for modeling sequences of dependent collisions that converges uniformly between the limit values.

2 Collisions Reviewed

On a macroscopic scale, collision behavior can be modeled by (i) the equations of motion that describe relative velocities and acceleration and forces at the contact points and (ii) the contact conditions, i.e., the conditions that determine whether the collision equations of motion are active. In addition, it is required to find contact points, i.e., perform collision detection, which is beyond the scope of this paper.

The distance between two bodies at a contact point i is given by g_N^i , the relative velocities at a contact point by $\dot{g}_N^i$ and the constraint force in the normal direction by λ_N^i . Methods for computing these values immediately before collision are described in detail elsewhere; e.g., in [3,5]. In the following, because of the assumptions that the collision is without friction, at center, and only occurs in the normal direction, the subscript N can be removed, as all variable values are given in the normal direction.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received December 5, 2000; final manuscript received January 4, 2001. Review conducted by V. Kinra.

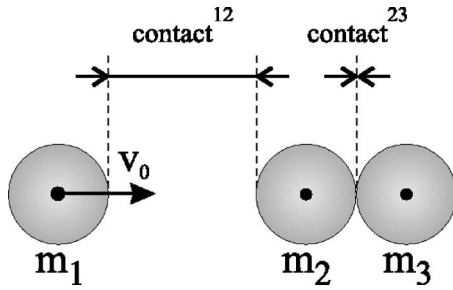


Fig. 1 A sequence of dependent collisions

2.1 A Linear Approximation. To evaluate the accuracy of the collision rules, at least in a gross sense, a linear first-order spring-damper (RC) collision approximation for the system in Fig. 1 is used as a reference. The velocity values after the collision are presented in Table 1 for several parameter combinations (simulations are performed with the hybrid bond graph simulator HY-BRsim [8,9]). Here, the damper parameters R^{12} and R^{23} capture the dissipation of the collisions between m_1 and m_2 , and m_2 and m_3 , respectively. Similarly, the linear approximation of the elasticity is captured by C^{12} and C^{23} . The initial state is $v_1=1$, $v_2=0$, and $v_3=0$, where v_j , $j \in \{1, 2, 3\}$, is the velocity of mass m_j .

Two mass distributions are evaluated; one where $m_1=m_2=m_3=1$ and the other where $m_1=m_2=1$ and $m_3=1000$. The latter corresponds to an example in other work [5] where a ball is bounced off another ball that is at rest on a floor (see Fig. 2). The dissipation is varied between large and small values to serve as a reference for $\epsilon \approx 0$ and $\epsilon \approx 1$, respectively, in the idealized case. Note

that the linear model is a rough approximation of the real collision phenomenon, and only valid as a rather qualitative indication.

2.2 Collision Rules. The equations of motion can be described by a Newton or Poisson type collision rule. Newton's rule specifies the collision at a contact point in terms of the relative velocity before ($\dot{g}_A$) and after the collision ($\dot{g}_E$)¹

$$\dot{g}_E = -\epsilon \dot{g}_A \quad (1)$$

where ϵ ($0 \leq \epsilon \leq 1$) is the *coefficient of restitution* that determines the energy that is dissipated during the collision. For $\epsilon=0$, the collision is *perfectly non-elastic* and there is maximum dissipation, for $\epsilon=1$, the collision is *perfectly elastic* and there is no dissipation.

Poisson's collision rule is formulated on the force level and consists of a compression and expansion stage. During the compression stage, a force builds up until the relative velocity between the colliding bodies has vanished. The velocities of the individual bodies at this point are indexed by C , e.g., v_C . This force is then applied during the expansion stage. In the idealized collision model, these compression and expansion phases occur instantaneously at the same point in time, but are ordered so that the expansion phase begins immediately after the compression phase is completed. Thus the compression and expansion forces become impulsive: Λ_C and Λ_E , respectively. To account for dissipation, the return impulse is modified by a coefficient ϵ

$$\Lambda_E = \epsilon \Lambda_C \quad (2)$$

Note that computationally this coefficient ϵ is not the same as the coefficient of restitution ϵ used in Newton's collision rule. However, conceptually they are similar in that they capture the degree of elasticity of the collision. Therefore, and in order not to produce excessive, and to an extent superfluous, notation, the same variable ϵ is used in the proceedings.

Given the equations of motion, the contact conditions still need to be specified to completely capture collision behavior. In the case of the unilateral constraints of collision phenomena, the complementarity formulation can be used. This formulation states that (i) when the bodies are touching, $g^i=0$, there can be a positive force between them ($\lambda^i > 0$) and (ii) when the bodies are disconnected ($g^i > 0$) there is no force between them ($\lambda^i = 0$). This can be formulated in a compact set of constraints

$$g^i \geq 0 \quad \lambda^i \geq 0 \quad g^i \lambda^i = 0 \quad \forall i \in I^G \quad (3)$$

where I^G is the set of all possible contact points. Typically [3,5,10], these complementarity constraints are formulated on the acceleration level, in which case the g^i variable is replaced by $\ddot{g}^i$

$$\ddot{g}^i \geq 0 \quad \lambda^i \geq 0 \quad \ddot{g}^i \lambda^i = 0 \quad \forall i \in I^G \quad (4)$$

The use of the complementarity principle in physical system modeling is discussed in more detail elsewhere [11]. Other work [7] has used a state transition formulation to determine whether contact exists and whether a collision occurs.

In [3], the collision conditions for applying Newton's collision rule are

$$I^N = \{i \in I^G | g^i = 0; \dot{g}^i \leq 0\} \quad (5)$$

Thus, the relative distance has vanished and the relative velocity is toward each other. The condition for applying Poisson's collision rule, though, is

$$I^P = \{i \in I^G | g^i = 0\} \quad (6)$$

The reason for this is that to ensure the bodies do not penetrate in the case of multiple contacts, Newton's collision rule requires stricter assumptions. Because Poisson's collision rule breaks down in two phases, it can be evaluated as to whether penetration will

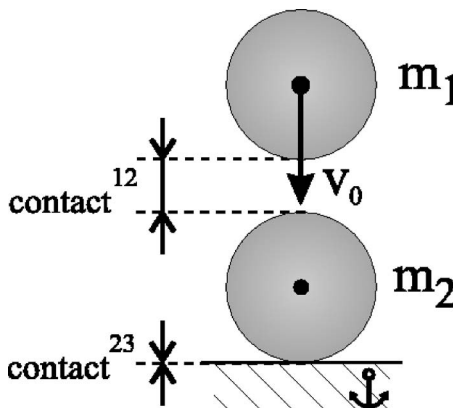


Fig. 2 Collision between two bodies and a floor

¹Notation is derived from [3]. Contact points will be superscripted, whereas reference to bodies around contact points and their variables are subscripted.

occur given the compression impulses in an intermediate evaluation, and at this point it can be corrected by allowing arbitrarily large impulses. Note that this implies that the collisions are evaluated simultaneously.

2.3 Example. In accordance with Table 1, the differences in behavior are illustrated by two mass distributions, i.e., $m_1=m_2=m_3=1$ and $m_1=m_2=1, m_3=1000$, for the system in Fig. 1 with varying degrees of elasticity for the respective collisions.

2.3.1 Newton. For the system in Fig. 1, upon vanishing of g^{12} , it is evaluated whether $g^{12} \leq 0$, which is true. In addition $g^{23}=0$ and $\dot{g}^{23} \leq 0$, so the collision rule generates

$$(v_{2,E} - v_{1,E}) = -\epsilon^{12}(v_{2,A} - v_{1,A}) \quad (7)$$

$$(v_{3,E} - v_{2,E}) = -\epsilon^{23}(v_{3,A} - v_{2,A}) \quad (8)$$

In addition, the forces have to be balanced, i.e., $F_1 + F_2 + F_3 = 0$, which has the instantaneous equivalent (conservation of momentum)

$$m_1(v_{1,E} - v_{1,A}) + m_2(v_{2,E} - v_{2,A}) + m_3(v_{3,E} - v_{3,A}) = 0 \quad (9)$$

This combines into a set of three equations with three unknowns, and it can be solved for $v_{j,E}$, $j \in \{1, 2, 3\}$.

Figure 3 shows the velocities computed from evaluating these constraints for varying ϵ^{12} and ϵ^{23} and fixed $\epsilon^{23}=1$ and $\epsilon^{12}=1$, respectively. Note that, because $(v_{3,A} - v_{2,A})=0$, $v_{3,E}=v_{2,E}$. For $m_1=m_2=m_3=m$ and $\epsilon^{12}=1$, the results equal a collision between a body with mass m and one with mass $2m$, which represents anomalous behavior because momentum should be completely transferred from m_1 to m_3 .

In case m_3 becomes very large, Fig. 3(b), it remains almost at rest and because $(v_{3,A} - v_{2,A})=0$ the same holds for m_2 . In terms of Fig. 2, using Newton's collision rule, m_2 never leaves the floor, which contradicts the results for a linear approximation in Table 1. This conforms the analysis in [5] and is circumvented by Poisson's collision rule.

2.3.2 Poisson. For the Poisson collision rule, upon vanishing of g^i , first the impulsive forces because of the compression phase are computed. This results in $(v_{i,C})$ are the velocities at the end of the compression phase)

$$\Lambda_C^{12} = -m_1(v_{1,C} - v_{1,A}) \quad (10)$$

$$\Lambda_C^{23} = m_3(v_{3,C} - v_{3,A}) \quad (11)$$

where the minus sign is determined by the chosen reference point and is present because the forces between the bodies work in opposite directions. This introduces two additional variables, which requires two more equations that arrive from the constraint that at the end of the compression phase the relative velocity is 0, so

$$v_{1,C} = v_{2,C} \quad (12)$$

$$v_{2,C} = v_{3,C} \quad (13)$$

With the equations

$$\Lambda_E^{12} = \epsilon^{12} \Lambda_C^{12} \quad (14)$$

$$\Lambda_E^{23} = \epsilon^{23} \Lambda_C^{23} \quad (15)$$

and conservation of momentum

$$m_1(v_{1,C} - v_{1,A}) + m_2(v_{2,C} - v_{2,A}) + m_3(v_{3,C} - v_{3,A}) = 0, \quad (16)$$

it can be solved for $v_{1,C}$, $v_{2,C}$, $v_{3,C}$, Λ_C^{12} , Λ_C^{23} , Λ_E^{12} , and Λ_E^{23} . Instead

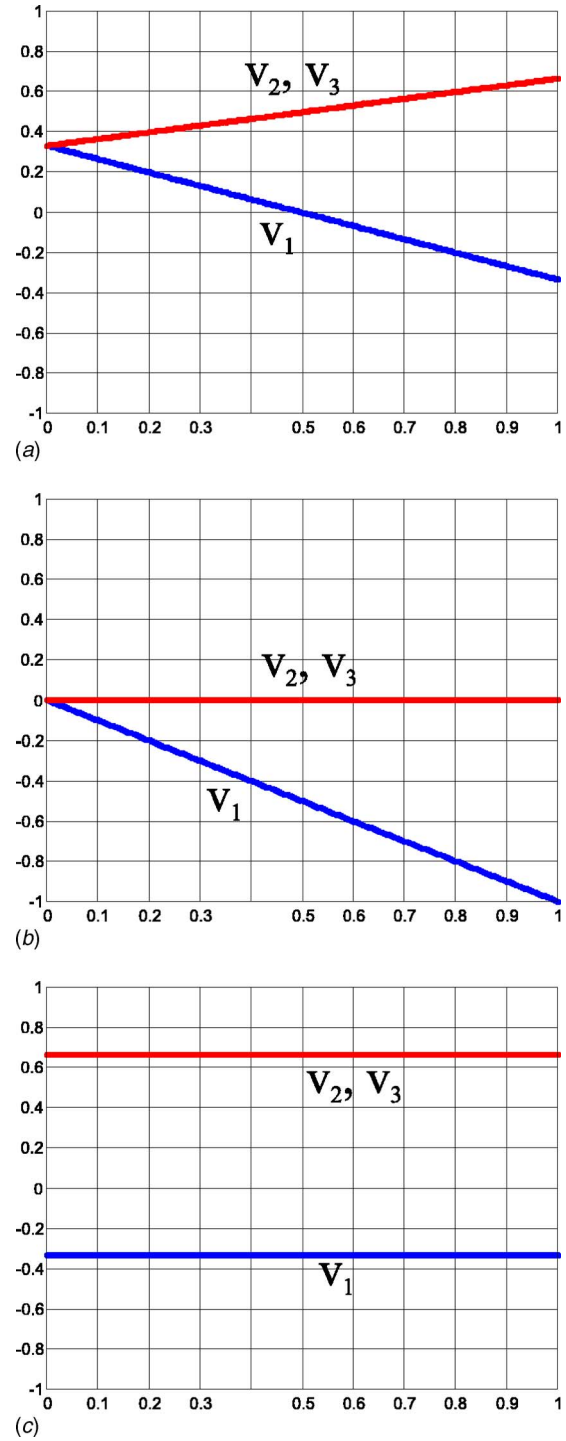


Fig. 3 Resulting velocities for Newton's collision rule: (a) As a function of ϵ^{12} ; (b) as a function of ϵ^{12} with m_3 large; (c) as a function of ϵ^{23}

of Eqs. (14) and (15), the following complementarity problems are solved

$$(\Lambda_E^i - \epsilon^i \Lambda_C^i) \dot{g}_E^i = 0 \quad (17)$$

to ensure either Poisson's rule is applied (the left factor) or the relative velocity vanishes (the right factor). Thus, the impulse Λ_E^i is computed by Poisson's collision rule, or by ensuring there is no penetration.

In the example in Fig. 1, for equal masses and $\epsilon^{23}=1$, the following sequence of values is obtained for $\epsilon^{12} < 0.25$

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ \Lambda_1 \\ \Lambda_2 \end{bmatrix} = \begin{bmatrix} v_0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} \frac{1}{3}v_0 \\ \frac{1}{3}v_0 \\ \frac{1}{3}v_0 \\ \frac{2}{3}v_0 \\ \frac{1}{3}v_0 \end{bmatrix}, \quad \begin{pmatrix} \frac{1}{3}(1-2\epsilon^{12})v_0 \\ \frac{2}{3}\epsilon^{12}v_0 \\ \frac{2}{3}v_0 \\ \frac{2}{3}\epsilon^{12}v_0 \\ \frac{1}{3}v_0 \end{pmatrix}, \quad \begin{bmatrix} \frac{1}{6}v_0 \\ \frac{1}{6}v_0 \\ \frac{2}{3}v_0 \\ \frac{1}{6}v_0 \\ \frac{1}{3}v_0 \end{bmatrix} \quad (18)$$

where the vector in parentheses contains the values that would have been obtained if instead of the complementarity constraints in Eq. (17) to prevent penetration of m_1 and m_2 , Eq. (14) and Eq. (15) were used. To prevent penetration, an arbitrarily large impulse is computed from the $\dot{g}^{12}=0$ constraint; in this case, $\Lambda_E^{12} = \frac{1}{6}v_0$. If $\epsilon^{12} \geq 0.25$, there is no penetration and the vector values in parentheses give the final collision values.

Figure 4 shows the resulting velocities for varying ϵ^{12} and ϵ^{23} . Clearly, in case $\epsilon^{23}=1$, the collision rule does not return values conform the linear approximation in Table 1. Also, for fixed $\epsilon^{12}=1$, behavior in Fig. 4(c) shows no dependence on ϵ^{23} , although this is expected from the linear approximation (compare the $\{R^{12}, R^{23}\}$ entry $\{0.05, 0.05\}$ in Table 1 with entry $\{0.05, 50\}$).

For a large mass m_3 , approximating the setup in Fig. 2, this collision rule does predict m_2 leaving the floor, in accordance with the linear first-order approximation in Table 1 (entry $R^{12}=50$ and $R^{23}=0.05$), but for the other limit case $\epsilon^{12}=1$ and $\epsilon^{23}=1$, its prediction deviates from the linear first-order approximation (entry $R^{12}=0.05$ and $R^{23}=0.05$).

3 A Uniformly Converging Formula

As illustrated in Sec. 2, the collision effects based on Newton and Poisson rules do not provide satisfactory results for centralized frictionless dependent collision sequences for all coefficients of restitution; i.e., $0 \leq \epsilon^i \leq 1$. To devise a formula that does, the behavior of a first-order linear approximation of the collision sequence in Fig. 1 is first analyzed (see also [7]). For two different damper parameters of the collision between m_2 and m_3 ($R^{23}=0.05$ and $R^{23}=5$), the collision behavior is shown in Fig. 5. From this it can be seen that the larger the dissipation between m_2 and m_3 ; i.e., the less elastic the collision, the less the difference in relative velocity that is allowed.

This observation leads to the conjecture that the velocity of bodies in contact, not immediately partaking in a collision, follow each other more closely or not, depending on the elasticity between them. The velocity of one body can then be described as being a fraction of the velocity of the other; e.g., $v_3 = \eta^{23}v_2$ ($0 \leq \eta^{23} \leq 1$) before the collision effect takes place. Here, η^{23} is assumed to depend on the elasticity, or coefficient of restitution, as $\eta^{23} = 1 - \epsilon^{23}$. In general, this leads to the following instantaneous representation at a contact point i

$$v_2^i = \eta^i v_1^i \quad (19)$$

with $\eta^i = 1 - \epsilon^i$ and v_1^i and v_2^i such that $v_2^i > v_1^i$.

To design a uniformly converging collision rule, the second collision in the sequence (i.e., the one between m_2 and m_3 in Fig. 1) has to take the change of velocities during the preceding, interacting collision into account. Including this interaction in an idealized collision rule is difficult because the collisions are active only at one moment in time and partial interaction is difficult to model. For Poisson's collision rule, limited interaction is possible by allowing interaction between compression and expansions phases, but this is insufficient to handle the illustrated complex interaction.

To design a more flexible collision rule, an additional phase in the collision process is introduced that accounts for the change in

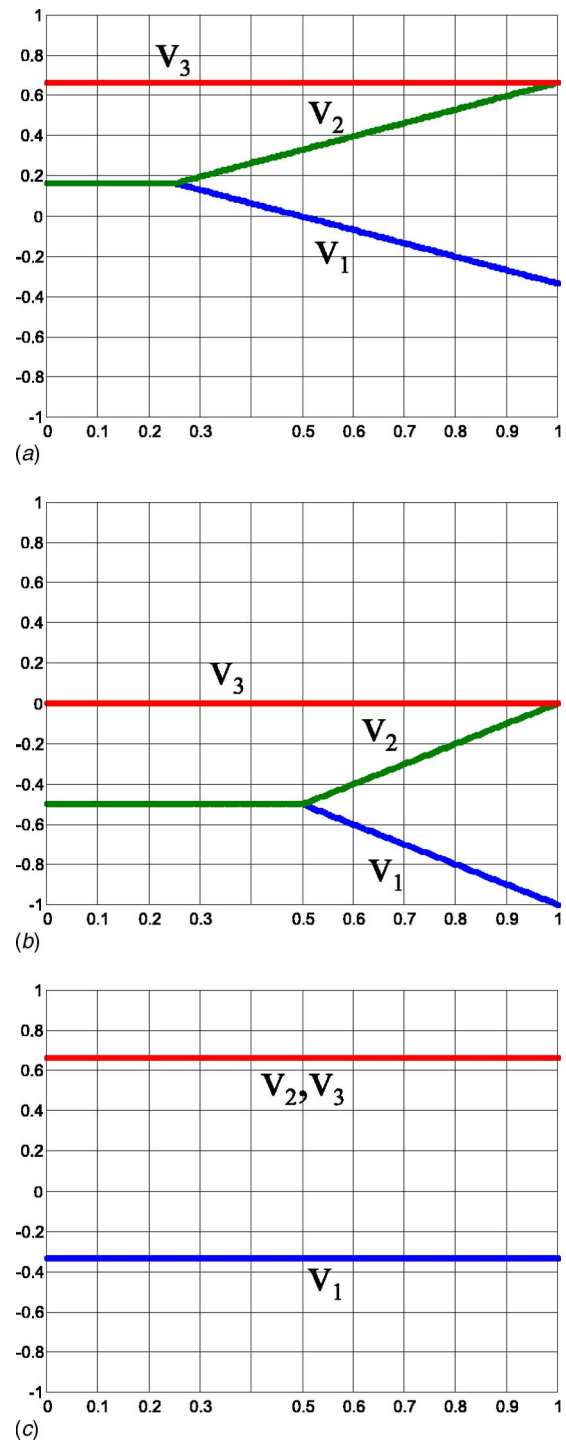


Fig. 4 Resulting velocities for Poisson's collision rule: (a) As a function of ϵ^{12} ; (b) as a function of ϵ^{12} with m_3 large; (c) as a function of ϵ^{23}

relative velocity as given in Eq. (19). This constraint becomes active whenever there is a dependent contact point where a collision takes place that affects the velocity, i.e., for a contact point i , such that $g^i = v_2^i - v_1^i$

$$g^i = 0 \wedge g_E^i < 0 \wedge (v_{1,E}^i \neq v_{1,A}^i \vee v_{2,E}^i \neq v_{2,A}^i) \quad (20)$$

Here the constraints on relative position and velocity determine that a collision occurs and the constraints on the individual veloci-

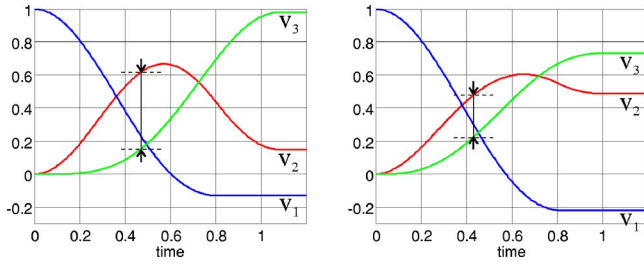


Fig. 5 Different elasticity leads to differences in relative velocity

ties determine that dependent collisions are active that affect the velocities.²

To allow collisions to occur at nonzero velocity, the constraint in Eq. (19) has to be extended. For this case, $|v_{2,E} - v_{2,A}| > |v_{3,E} - v_{3,A}|$, and v_3 reaches a fraction of v_2 while taking an offset $v_{3,A}$ into account

$$v_{3,E} = (1 - \epsilon^{23}) * (v_{2,E} - v_{3,A}) + v_{3,A} \quad (21)$$

or, in general, for a contact point i

$$v_{2,E}^i = (1 - \epsilon^i) * (v_{1,E}^i - v_{2,A}^i) + v_{2,A}^i \quad (22)$$

However, in this formulation, for $\epsilon^i = 1$, $v_{2,E}^i = v_{2,A}^i$, and, therefore, $v_{2,E}^i$ is fixed. Thus, other simultaneous collisions cannot affect $v_{2,E}^i$, an anomalous situation. Instead, the formulation should be such that for this value of ϵ^i the velocity $v_{2,E}^i$ is not affected by the dependent collision; i.e., there is no force acting. This requires a formulation at the force (impulse) level, instead of in terms of velocities. Accounting for the substitute mass, $m_s^i = m_1^i m_2^i / (m_1^i + m_2^i)$, where m_1^i and m_2^i are the two masses involved in the collision at contact i , this leads to the formula

$$\epsilon^i \Lambda^i + (1 - \epsilon^i) m_s^i g_E^i = 0, \quad 0 \leq \epsilon^i \leq 1 \quad (23)$$

Thus, for $\epsilon^i = 0$, the velocities become equal and for $\epsilon^i = 1$ there is no interaction because $\Lambda^i = 0$.

For one contact point between two masses, the impulsive force acting on m_2 can be derived from Eq. (23) by writing

$$\epsilon^i m_2^i (v_{2,E}^i - v_{2,A}^i) + (1 - \epsilon^i) m_s^i (v_{2,E}^i - v_{1,E}^i) = 0 \quad (24)$$

From this, in the case $m_1^i = m_2^i = 1$, Eq. (22) follows.

Putting the pieces together, the collision rule extends the standard Newton's collision rule as follows. If $g^i = 0 \wedge \dot{g}_E^i < 0$ contact behavior is evaluated based on two cases:

1. A pre-collision phase because of dependency between collisions (i.e., $v_{2,E}^i \neq v_{2,A}^i \vee v_{1,E}^i \neq v_{1,A}^i$). Behavior is computed based on an initial velocity change and by balancing forces, i.e.

$$\begin{aligned} \epsilon^i \Lambda^i + (1 - \epsilon^i) m_s^i g_E^i &= 0 \\ F_1^i &= -\Lambda^i \\ F_2^i &= \Lambda^i \end{aligned} \quad (25)$$

with m_s^i the substitute mass.

2. Execution of the collision effect, computed based on Newton's collision rule and by balancing forces, i.e.

$$v_{2,E}^i - v_{1,E}^i = -\epsilon^i * (v_{2,A}^i - v_{1,A}^i)$$

²During continuous behavior, $v_{1,E}^i = v_{1,A}^i$ and $v_{2,E}^i = v_{2,A}^i$, as the left and right limit values of a point on a continuous (C^0) curve have to be equal.

$$F_1^i = -F_2^i \quad (26)$$

Thus, behavior at the contact points is then governed by either the velocity restitution based on Newton's collision rule or by the pre-collision effect (in addition to conservation of momentum, no longer explicitly mentioned in the proceedings). The particular contact points are elements of the sets I_k^N and I_k^P , respectively. Furthermore, the pre-collision effect is always first completed by executing Newton's collision rule after the pre-collision at that contact point, provided $\dot{g}_E^i < 0$ still holds, before new independent collisions are inferred. The forces are summed in case multiple contact points exert a force on the same mass.

In summary, the independent collision set that may trigger a sequence of collision and pre-collision effects can be specified by

$$I_k^{N1} = \{i \in I_G | (g^i = 0 \wedge \dot{g}_E^i \leq 0) \wedge (v_{2,E}^i = v_{2,A}^i \wedge v_{1,E}^i = v_{1,A}^i)\} \quad (27)$$

and applies Newton's collision rule.³ Because of the second clause, the sequence of pre-collision and collision effects is first completed before a possible new collision sequence is triggered.⁴ Each pre-collision in a set I_{k-1}^P may be followed by Newton's collision rule as well, gathered in the set I_k^{N2}

$$I_k^{N2} = \{i \in I_G | i \in I_{k-1}^P \wedge (g^i = 0 \wedge \dot{g}_E^i \leq 0)\} \quad (28)$$

and $I_k^N = I_k^{N1} \cup I_k^{N2}$. Given these sets of contact points, I_k^{N1} and I_k^{N2} , the "intermediate" velocities $\dot{g}_p^i$ are computed using Eq. (26) with $v_{j,E}^i$ substituted for the input $v_{j,A}^i$. These are intermediate velocities as their values may cause pre-collision effects I_k^P

$$I_k^P = \{i \in I_G | \dot{g}_p^i \leq 0\} \quad (29)$$

When all sets, i.e., I_k^{N1} , I_k^{N2} , and I_k^P , are determined, the velocities are updated, i.e., $v_{j,A}^i = v_{j,E}^i$, and the new $v_{j,E}^i$ are computed using Eqs. (25) and (26) for the contact points in I_k^N and I_k^P , respectively.

During this sequence of computations, $v_{j,E}^i \neq v_{j,A}^i$, and, therefore, $I_{k+1}^{N1} = \emptyset$. When the sequence of pre-collisions and collisions has terminated, $I_{k+m}^{N2} = \emptyset$ and $\dot{g}_p^i = g^i$, and, consequently, $I_{k+m}^P = \emptyset$ as well, where $0 < l < m$. Therefore, $v_{j,E}^i = v_{j,A}^i$ and I_{k+m}^{N1} may be populated again and further collisions may occur. The first collision is indexed $k=1$, $I_0^P = I_0^{N1} = I_0^{N2} = \emptyset$, and k is reset to 0 when no further collisions occur (either dependent or independent).

The results of this new collision rule for three colliding bodies, as shown in Fig. 1, can now be compared against those for Newton and Poisson collision rules presented in Sec. 2. Both the Newton and Poisson collision rules did not properly transfer momentum from m_1 to m_3 in the case $\epsilon^{12} = \epsilon^{23} = 1$, as can be seen in Figs. 3 and 4. The new collision rule does properly exhibit this behavior, as illustrated by Fig. 6(a). In addition, Figs. 3(c) and 4(c) show the Newton and Poisson collision rules to have no ϵ^{23} dependence when $\epsilon^{12} = 1$. In Fig. 6(c), the ϵ^{23} dependence of the new collision rule is shown to achieve correct values for the limit values, $\epsilon^{23} = 0$ and $\epsilon^{23} = 1$.

In case a large mass m_3 is used, mimicking a collision between two bodies, one of which is at rest on a floor (see Fig. 2), Newton's collision rule would not properly compute m_2 leaving the floor, as shown in Fig. 3(b). The new collision rule does yield this behavior, as shown in Fig. 6(b), and it does not produce an abrupt change in behavior for varying ϵ^{12} , as illustrated for Poisson's collision rule in Fig. 4(b).

Note that for large m_3 and $\epsilon^{12} = \epsilon^{23} = 1$, similar to Poisson's collision rule, the new collision rule computes velocities that differ from the first-order approximation in Table 1, entry $m_3 = 1000$, $R^{12} = 0.05$, $R^{23} = 0.05$. This approximation, however, is a crude one, and real-life experiments should be performed to establish

³Note that g^i is not indexed by either E or A as it is a C^0 variable.

⁴Alternatively, to achieve the same effect, the clause can be replaced by $I_{k-1}^P = \emptyset$.

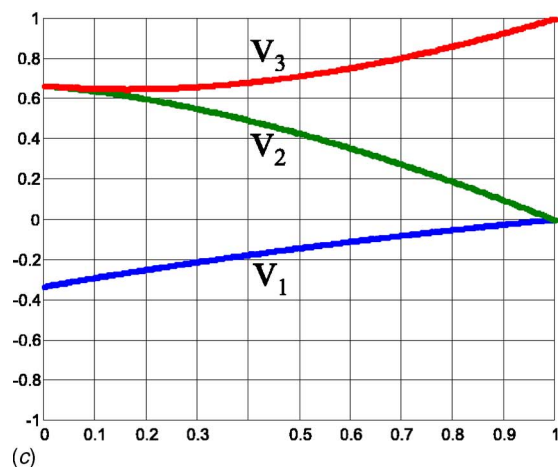
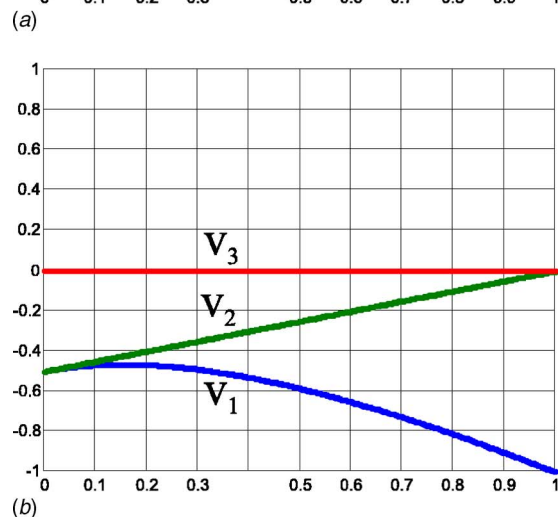
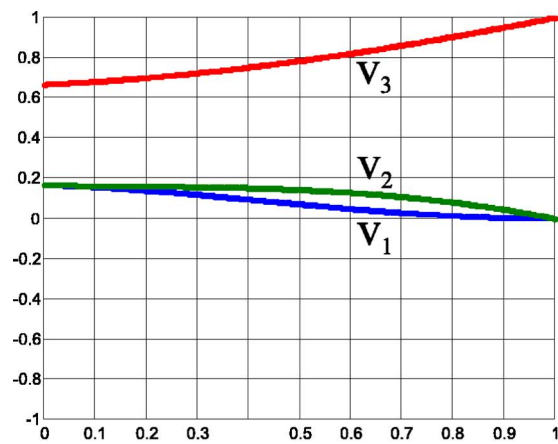


Fig. 6 Resulting velocities for the new collision rule: (a) As a function of ϵ^{12} ; (b) as a function of ϵ^{12} with m_3 large; (c) as a function of ϵ^{23}

realistic behavior. Intuitively, for $\epsilon^{12} = \epsilon^{23} = 1$, a sequence of isolated collisions would occur, transferring momentum completely from m_1 to m_2 , which then collides with the large mass m_3 , causing m_2 to reverse its velocity and to then transfer its momentum completely back to m_1 . Therefore, m_2 would remain at rest.

4 Increasing Complexity

Because three bodies only allow two collisions, and, therefore, analysis of only one pair of dependent collisions, generality is

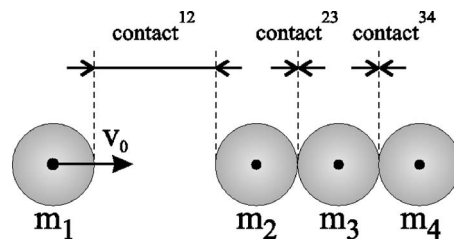


Fig. 7 A sequence of four colliding bodies

limited. To investigate a more general situation, sequences of collisions of four bodies, as shown in Fig. 7, are analyzed. The results for varying ϵ^i are presented in Fig. 8 for $m_1 = m_2 = m_3 = m_4 = m$. Collision behavior for the different sets of active rules is derived and computed by MATLAB[®] [12] and its Symbolic Math Toolbox [13].⁵ Note that, again, collision behavior is smooth and converges between the two limit values that can be verified to match collision behavior of more detailed models.

To illustrate the execution of the collision law, consider the case where $\epsilon^{23} = 0.5$ in Fig. 8(b). The sequence of collision effects for this parameter value is given in Table 2. The velocities achieved during the collision are

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix}^k = \begin{bmatrix} 0 \\ v_0 \\ 0 \\ 0 \end{bmatrix}, \left(\begin{bmatrix} 1 \\ 0 \\ v_0 \\ 0 \end{bmatrix} \right), \begin{bmatrix} -\frac{1}{7}v_0 \\ \frac{6}{7}v_0 \\ \frac{2}{7}v_0 \\ 0 \end{bmatrix}, \left(\begin{bmatrix} -\frac{1}{7}v_0 \\ \frac{3}{7}v_0 \\ \frac{5}{7}v_0 \\ 0 \end{bmatrix} \right), \begin{bmatrix} -\frac{1}{7}v_0 \\ \frac{3}{7}v_0 \\ \frac{5}{7}v_0 \\ 0 \end{bmatrix}, \left(\begin{bmatrix} \frac{3}{7}v_0 \\ -\frac{1}{7}v_0 \\ \frac{3}{7}v_0 \\ 0 \end{bmatrix} \right), \begin{bmatrix} -\frac{1}{7}v_0 \\ \frac{15}{49}v_0 \\ \frac{3}{49}v_0 \\ \frac{38}{49}v_0 \end{bmatrix}, \left(\begin{bmatrix} -\frac{1}{7}v_0 \\ \frac{6}{49}v_0 \\ \frac{12}{49}v_0 \\ \frac{38}{49}v_0 \end{bmatrix} \right) \quad (30)$$

where the values in parentheses are the results after computing g_p^i , these are replaced because analysis shows that further pre-collision effects occur. Note that pre-collision effects do not trigger further pre-collision effects; otherwise, at step 1, contact 34 would be included as a pre-collision because $v_3 > v_4$.

A detailed physical explanation for this is difficult to provide as the propagation of the collision shockwaves is a complicated and still not a well-understood phenomenon. Given the accuracy of the presented model, it appears as though there is a degree of coupling between the “secondary” contact points (those that partake in the pre-collision) that allow shockwaves of the primary collision to travel across them, the degree of which is related to the coefficient of restitution. Given this, the secondary contact points represent a first-order effect, and these higher order effects of this coupling can be abstracted away from the collision model.

This observation is supported by experiments reported in [14] where an “impulse transmission ratio” is introduced similar to the coupling coefficient η in Sec. 3. The impulse transmission ratio is applied to triplets of adjacent bodies, or two adjacent contact

⁵The MATLAB code of the collision models is available upon request.

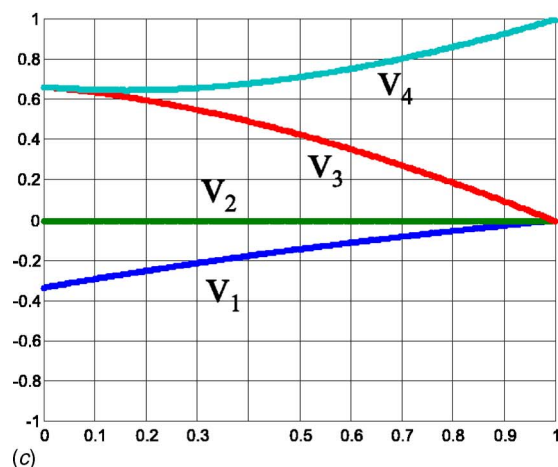
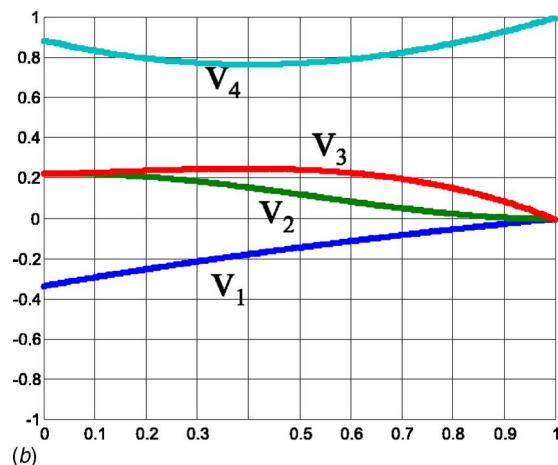
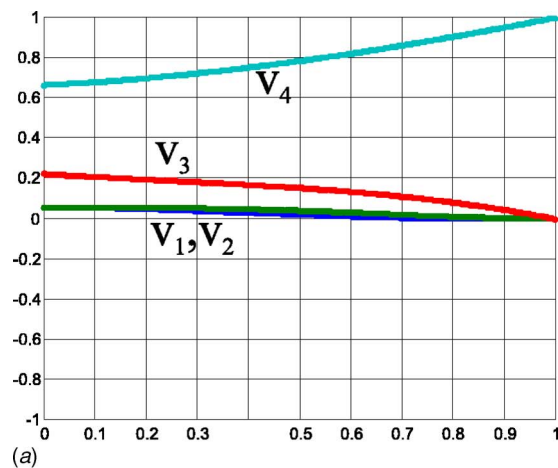


Fig. 8 Resulting velocities for the new collision rule: (a) As a function of ϵ^{12} ; (b) as a function of ϵ^{23} ; (c) as a function of ϵ^{34}

points, as well. Note that, in general, η can be chosen an independent parameter, as opposed to substituting $\eta = 1 - \epsilon$.

Table 2 Collision sets

k	I_k^{N1}	I_k^{N2}	I_k^P
1	{1}	{}	{2}
2	{}	{2}	{3}
3	{}	{3}	{2}
4	{}	{2}	{}

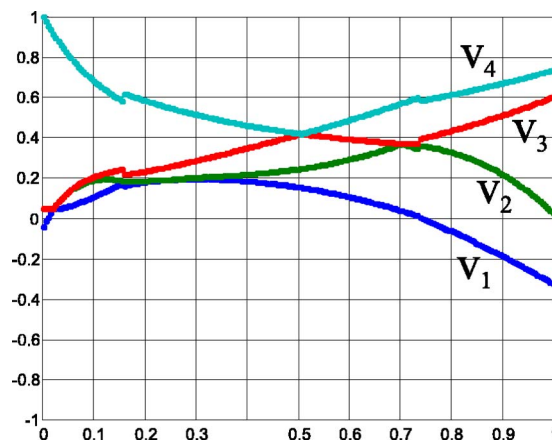


Fig. 9 Final velocities for $m_2=0.1$ and $0 \leq \epsilon^{23} \leq 1$

This system also includes behavior where a pre-collision is triggered for a contact but not followed by Newton's collision rule, i.e., it does not appear in I_1^{N2} . If $\epsilon^{23}=0$, the first collision is $I_1^{N1}=\{1\}$, $I_1^{N2}=\{\}$, and $I_1^P=\{2\}$, and the post-velocities $v_1=-\frac{1}{3}v_0$, $v_2=v_3=\frac{2}{3}v_0$, and $v_4=0$. The next iteration yields $I_1^{N1}=\{\}$, $I_1^{N2}=\{\}$, and $I_1^P=\{\}$ because $v_{j,E}^i \neq v_{j,A}^i$, $i \in \{1,2,3\}$, $j \in \{1,2\}$, and, therefore, contact 34 is not included in I_1^{N1} . After the $v_{j,A}^i$ are updated, the collision computation is resumed with $I_2^{N1}=\{3\}$, $I_2^{N2}=\{\}$, and $I_2^P=\{2\}$.

The presented extended collision rule is still sensitive to initial value perturbations. Consider the mass distribution $m_1=1$, $m_2=0.1$, $m_3=1$, and $m_4=1$. For $0 \leq \epsilon^{23} \leq 1$, the velocities after collision are shown in Fig. 9. For certain values of ϵ^{23} , discontinuous changes in final velocities occur. This is caused by a change in the sequence of active contact points when it includes simultaneous collisions; i.e., I_k^{N2} has a size larger than 1 (see the results in Table 3). Even when such a collision is triggered with very small relative velocity, it causes a significant change in the post-collision stage as well when collisions with low relative velocity take place. The effect of these pre-collisions can gradually decrease until for larger relative velocity a true collision takes place; i.e., $\eta^j=f(\dot{g}^i)$. Future work will concentrate on further exploring this phenomenon.

5 Conclusions

Often, in collision models the normal component is handled by either a Newton or Poisson type collision law. In the case of multiple impacts, the collisions are assumed to occur at the same point in time and to be synchronized in terms of their compression and expansion behavior.

This assumption is violated when sequences of dependent collisions occur and compression and expansion phases overlap. In the limit cases, i.e., restitution coefficients of value 0 and 1, either all constraints need to be activated simultaneously or sequentially. This is discussed in detail in other work [7]. Because of the two

Table 3 Collision sets, $m_2=0.1$, $\epsilon^{23}=0.6$

k	I_k^{N1}	I_k^{N2}	I_k^P
1	{1}	{}	{2}
2	{}	{2}	{1,3}
3	{}	{1,3}	{2}
4	{}	{2}	{1}
5	{}	{1}	{2}
6	{}	{2}	{1,3}
7	{}	{1,3}	{}

different methods of handling the limit cases, there is no uniform convergence between them based on one type of activation constraint only.

This paper presents a method wherein a pre-collision phase precedes the application of Newton's collision rule in the case of multiple dependent sequential collisions. It is shown how the resulting characteristic converges uniformly between the limit cases. Furthermore, it is shown that this extended collision rule conforms better with a linear first-order approximation.

Because of the superior characteristics of Poisson's collision rule for multiple contacts [5], future work will concentrate on extending this collision rule similarly. Furthermore, sensitivity to simultaneous independent collisions will be further investigated.

Acknowledgment

During this work, Pieter J. Mosterman was supported by a grant from the DFG Schwerpunktprogramm KONDISK and employed by the German Aerospace Center (DLR) Oberpfaffenhofen. Mehmet Yunt provided some valuable comments on an earlier draft of this paper that are much appreciated.

References

- [1] Brach, R. M., 1991, *Mechanical Impact Dynamics*, John Wiley and Sons, New York.
- [2] Chatterjee, A., and Ruina, A., 1998, "A New Algebraic Rigid Body Collision Law Based on Impulse Space Considerations," *ASME Trans. J. Appl. Mech.*, **65**(4), pp. 939–951.
- [3] Glocker, C., 1995, "Dynamik von starrkörpersystemen mit reibung und stößen." Ph.D. dissertation, Technical University of Munich, Germany (in German).
- [4] Ivanov, A. P., 1995, "On Multiple Impact," *J. Appl. Math. Mech.*, **59**(6), pp. 887–902.
- [5] Pfeiffer, F., and Glocker, C., 1996, *Multibody Dynamics With Unilateral Contacts*, John Wiley & Sons, Inc., New York.
- [6] Brogliato, B., 1999, *Nonsmooth Mechanics*, Springer-Verlag, London, ISBN 1-85233-143-7.
- [7] Mosterman, P. J., and Breedveld, P., 1999, "Some Guidelines for Stiff Model Implementation With the Use of Discontinuities," *International Conference on Bond Graph Modeling and Simulations (ICBGM '99)*, San Francisco, January 17–20, pp. 175–180.
- [8] Mosterman, P. J., and Biswas, G., 1999, "A Java Implementation of an Environment for Hybrid Modeling and Simulation of Physical Systems," *International Conference on Bond Graph Modeling and Simulations (ICBGM '99)*, San Francisco, January 17–20, pp. 157–162.
- [9] Mosterman, P. J., 2002, "HyBrSim – A Modeling and Simulation Environment for Hybrid Bond Graphs," *J. Syst. Control Eng.*, **216**, pp. 35–46.
- [10] Lötstedt, P., 1981, "Coulomb Friction in Two-Dimensional Rigid Body Systems," *ZAMM*, **61**, pp. 605–615.
- [11] Chatterjee, A., 1999, "On the Realism of Complementarity Conditions in Rigid Body Collisions," *Nonlinear Dyn.*, **20**, pp. 159–168.
- [12] MATLAB[®], 1997, *The Language of Technical Computing*, The MathWorks, Natick, MA.
- [13] MathWorks, 2003, *Symbolic Math Toolbox User's Guide*, The MathWorks, Natick, MA.
- [14] Ceanga, V., and Hurmuzlu, Y., 2001, "A New Look at an Old Problem: Newton's Cradle," *ASME J. Appl. Mech.*, **68**(4), pp. 575–583.

Xue-qian Fang¹
Department of Aerospace Engineering &
Mechanics,
Harbin Institute of Technology,
Harbin 150001, China
e-mail: fangxueqian@163.com

Chao Hu
Department of Aerospace Engineering &
Mechanics,
Harbin Institute of Technology,
Harbin 150001, China and
School of Aerospace Engineering and
Mechanics,
Tongji University,
Shanghai, 200092, China

Shan-yi Du
Department of Aerospace Engineering &
Mechanics,
Harbin Institute of Technology,
Harbin 150001, China

Dynamic Stress of a Circular Cavity Buried in a Semi-Infinite Functionally Graded Material Subjected to Shear Waves

The multiple scattering of shear waves and dynamic stress in a semi-infinite functionally graded material with a circular cavity is investigated, and the analytical solution of this problem is derived. The analytical solutions of wave fields are expressed by employing the wave function expansion method, and the expanded mode coefficients are determined by satisfying the boundary condition of the cavity. The image method is used to satisfy the traction-free boundary condition of the material structure. As an example, the numerical solution of the dynamic stress concentration factors around the cavity is also presented. The effects of the buried depth of the cavity, the incident wave number, and the nonhomogeneity parameter of materials on the dynamic stress concentration factors are analyzed. Analyses show that when the nonhomogeneity parameter of materials is <0 , it has less influence on the maximum dynamic stress around the cavity; however, it has greater influence on the distribution of dynamic stress around the cavity. When the nonhomogeneity parameter of materials is >0 , it has greater influence on both the maximum dynamic stress and the distribution of dynamic stress around the cavity, especially in the case that the buried depth is comparatively small. [DOI: 10.1115/1.2712238]

Keywords: exponentially graded material, multiple scattering of elastic waves, dynamic stress concentration factor, circular cavity

1 Introduction

Functionally graded materials (FGMs) are the new generation of composites and an important area of materials science research. The volume fraction of the materials changes gradually; the nonhomogeneous microstructures in the materials produce continuous graded macroscopic properties, such as the heat conductivity, specific heat, and mass density. All the properties have many potential applications, e.g., thermal barrier coating, thermal protection of the reentry capsule, furnace liners, personal body armor, the materials of elevated temperature, and heat resistance for the electromagnetic sensors, graded refractive index materials for optical applications. In an ideal FGM, the material properties may vary smoothly along one direction. As an example, having a smooth transition region between a pure metal and pure ceramic would result in a new type of materials, which combines the desirable high-temperature properties and thermal resistance of a ceramic with the fracture toughness of a metal [1,2].

In engineering, composite plate structures have found extensive use. During tailoring, connection, and serving of these structures, it is inevitable to make cavities, and some failures (such as holes and cracks) may also occur inside the structures. Under dynamic loads, the stress around and near the discontinuities may increase sharply, which causes the decrease of the strength of structures and the fatigue and fracture of structures. With the advent of FGMs, significant efforts have been made in the study of the dynamic stress around the cavities and cracks in the materials under dynamic loading. These include analytical, numerical, and experimental investigations.

At present, the numerical method is employed extensively to solve the dynamic stress around and near the discontinuities in

composite materials. Using the boundary element method, Rice and Sadd [3] investigated the propagation and scattering of SH-waves in semi-infinite homogeneous materials containing subsurface cavities, and the numerical solution of the dynamic stress around the cavity was obtained. By making use of Laplace and Fourier integral transforms and a numerical Laplace inversion technique, Li and Weng [4] presented the dynamic stress intensity factor of a cylindrical crack located in a functionally graded material interlayer between two coaxial elastic dissimilar homogeneous cylinders and subjected to a torsional impact loading, and the effect of parameters on dynamic stress intensity factor was also analyzed. Assuming an exponential spatial variation of the elastic properties, Ueda [5] adopted the Fourier transform technique to compute the dynamic stress intensity factor of the surface crack in a layered plate with a functionally graded nonhomogeneous interface and analyzed the effect of the geometric and material parameters on the variations of dynamic stress intensity factors. Finite element method was also applied to analyze the effect of different elastic gradient profiles on the fracture behavior of dynamically loaded functionally graded materials having cracks parallel to the elastic gradient [6]. Based on the integral equation for the crack in a nonhomogeneous medium with a continuously differentiable shear modulus, Chan et al. [7] studied the dynamic stress of the crack under shear waves in FGMs. Based on a meshless local boundary integral equation method, Sladek et al. [8] solved the dynamic stress of the crack numerically in functionally graded materials under the dynamic anti-plane shear loading.

Although these numerical methods are very useful tools for these problems, it is also very important to determine the physical behavior of the problems with analytical method. Pao and Chao [9] studied elastic wave scattering and dynamic stress concentrations in thin plate with cutouts, and the analytical and numerical solutions of the problem were presented. Kung [10] also studied dynamic stress concentrations in thin plates and gave the expressions of bending moment and shearing force versus frequency.

¹Corresponding author.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received May 18, 2006; final manuscript received August 16, 2006. Review conducted by Sridhar Krishnaswamy.

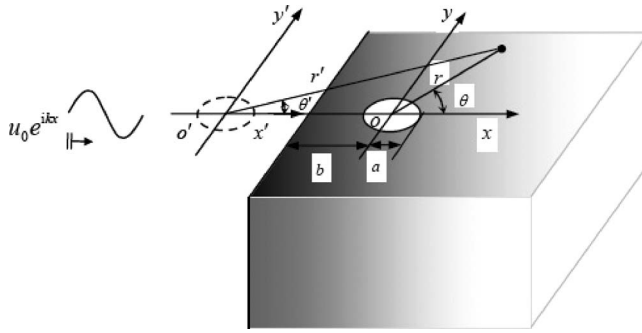


Fig. 1 Schematic of the buried cavity and the incident elastic waves in a semi-infinite graded material

Klyukin et al. [11] studied the scattering of flexural waves produced by a row of circular inclusions in plates, and the analytical solution was presented. The image method was applied to analyze scattering and dynamic stress concentrations of elastic waves in plates having a circular cavity subjected to a plane harmonic SH wave [12].

It is well known that the model of semi-infinite structures is one of the most common types of structural elements encountered in many practical engineering structures. However, because of the effects of the boundaries of the investigated areas, complex problems (e.g., multiple scattering of elastic waves) may arise. In this area, very little literature was reported in the past. In this paper, based on the theory of multiple scattering, employing image method, and wave functions expansion method, the multiple scattering and dynamic stress in a semi-infinite functionally graded material with a circular cavity is investigated and the analytical solution of this problem is presented. The addition theorem for Bessel function is used to accomplish the translation between different coordinate systems. The numerical solutions are graphically illustrated. The effects of the geometric and material parameters on the dynamic stress concentration factors around the cavity are also analyzed. The results can provide great help for the design of semi-infinite functionally graded materials.

2 Wave Motion Equation and Its Solution

Consider a semi-infinite plate structure of exponentially graded materials, as depicted in Fig. 1. For simplicity, we assume that the shear modulus and density of materials vary continuously, and the same nonhomogeneity parameter is used to describe the variation of them, i.e.,

$$\mu(x) = \mu_0 \exp(2\beta x), \quad \rho(x) = \rho_0 \exp(2\beta x) \quad (1)$$

where μ_0 and ρ_0 are the shear modulus and density of materials at the position of $x=0$, respectively, β is a nonhomogeneity parameter that denotes the exponent of spatial variation of the shear modulus and density of materials [2]. A circular cavity is buried at a depth of b beneath the surface of the semi-infinite material. An anti-plane shear wave propagates along the positive x direction in materials.

The anti-plane governing equation in materials is described as

$$\frac{\partial \tau_{xz}}{\partial x} + \frac{\partial \tau_{yz}}{\partial y} = \rho(x) \frac{\partial^2 u}{\partial t^2} \quad (2)$$

where τ_{xz} and τ_{yz} are the anti-plane shear stresses, and u is the displacement field.

The constitutive relations of anti-plane shear displacement are

$$\tau_{xz} = \mu(x) \frac{\partial u}{\partial x}, \quad \tau_{yz} = \mu(x) \frac{\partial u}{\partial y} \quad (3)$$

Substituting Eq. (3) into Eq. (2), the following equation can be obtained:

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + 2\beta \frac{\partial u}{\partial x} = \frac{1}{c_s^2} \frac{\partial^2 u}{\partial t^2} \quad (4)$$

where $c_s = \sqrt{\mu_0/\rho_0}$ is the wave speed of shear waves.

The steady solution of the problem is investigated, let $u = U \exp(-i\omega t)$, Eq. (4) can be changed into

$$\nabla^2 U + 2\beta \frac{\partial U}{\partial x} + k^2 U = 0 \quad (5)$$

where ω is the frequency of the incident waves, $k = \omega/c_s$ is the wave number of elastic waves.

The form of the solution of Eq. (5) can be proposed as

$$U = \exp(-\beta x) w(x, y) \quad (6)$$

where $w(x, y)$ is the function introduced for derivation.

Substituting Eq. (6) into Eq. (5), one can see that the function $w(x, y)$ should satisfy the following equation:

$$\nabla^2 w + \kappa^2 w = 0 \quad (7)$$

Here, $\kappa = (k^2 - \beta^2)^{1/2}$.

According to Eqs. (5)–(7), one can see that there exist elastic waves whose form is $U e^{-i\omega t} = u_0 \exp(-\beta x) e^{i(\kappa x - \omega t)}$, which denotes the propagating wave with its amplitude of vibration attenuating in the x direction.

According to Eq. (5), the general solution of the scattered field resulting from the cavity in FGMs can be described as

$$u^s = \exp(-\beta r \cos \theta) \sum_{n=-\infty}^{\infty} A_n H_n^{(1)}(\kappa r) e^{in\theta} \quad (8)$$

where $r = (x, y)$, $H_n^{(1)}(\cdot)$ is the n th Hankel function of the first kind, A_n are the mode coefficients of the scattered waves, and determined by satisfying the boundary condition. The solution of the scattered-reflected waves is the same as that of the scattered waves [12].

3 Excitation of Elastic Waves and the Total Wave Field

Consider an antiplane shear wave propagating along the positive x direction. In the local coordinate system (r, θ) of the real cavity, the incident waves can be described as

$$\begin{aligned} u_1^{(i)} &= u_0 \exp(-\beta x) \exp[i(\kappa x - \omega t)] \\ &= u_0 \exp(-\beta r \cos \theta) \sum_{n=-\infty}^{\infty} i^n J_n(\kappa r) \exp(in\theta) \exp(-i\omega t) \end{aligned} \quad (9)$$

where u_0 is the amplitude of the incident waves, κ is the wave number of the propagating waves, $J_n(\cdot)$ is the n th Bessel function of the first kind.

For the image cavity, the incident waves propagate in the negative x' direction and are expressed as

$$\begin{aligned} u_2^{(i)} &= u_0 \exp(\beta x') \exp[-i(\kappa x' + \omega t)] \\ &= u_0 \exp(\beta r' \cos \theta') \sum_{n=-\infty}^{\infty} i^{-n} J_n(\kappa r') \exp(in\theta') \exp(-i\omega t) \end{aligned} \quad (10)$$

The reflected waves at the edge of the semi-infinite FGM are described by the scattered waves resulting from the virtual image cavity. When the multiple scattering between the real and image cavities is considered, in the local coordinate system (r, θ) of the real cavity, the scattered field resulting from the real cavity can be described as

$$u_1^{(s)} = \exp(-\beta r \cos \theta) \sum_{l=1}^{\infty} \sum_{n=-\infty}^{\infty} A_n^l H_n^{(1)}(\kappa r) \exp(in\theta) \exp(-i\omega t) \quad (11)$$

where l denotes the time of the scattering between the real and image cavities, A_n^l ($l=1, 2, \dots, \infty$) determined by satisfying the boundary condition are the mode coefficients of the l th scattering of the real cavity.

Likewise, in the local coordinate system (r', θ') , the scattered field resulting from the image cavity can be described as

$$\begin{aligned} u_2^{(s)} &= \exp(\beta r' \cos \theta') \sum_{l=1}^{\infty} \sum_{n=-\infty}^{\infty} B_n^l H_n^{(1)}(\kappa r') \exp(in\theta') \exp(-i\omega t) \\ &= \exp(\beta r' \cos \theta') \sum_{l=1}^{\infty} \sum_{n=-\infty}^{\infty} B_n^l H_n^{(1)}(\kappa r') \exp(in\theta') \exp(-i\omega t) \end{aligned} \quad (12)$$

where B_n^l ($l=1, 2, \dots, \infty$) determined by satisfying the boundary condition is the mode coefficient of the l th scattering of the image cavity.

Thus, the total field of elastic waves in materials is taken to be the superposition of the incident field, the scattered field, and the reflected field at the edge of materials, namely,

$$u = u_1^{(i)} + u_1^{(s)} + u_2^{(s)} \quad (13)$$

To make computation tractable, the expression of elastic wave in the local coordinate system (r', θ') can be translated into another local coordinate system (r, θ) . According to the addition theorem for Bessel function [13], the following relation can be derived:

$$H_n^{(1)}(\kappa r') e^{in\theta'} = \sum_{m=-\infty}^{\infty} (-1)^{m-n} H_{m-n}^{(1)}(2\kappa b) J_m(\kappa r) e^{im\theta} \quad (14)$$

Similarly,

$$H_n^{(1)}(\kappa r) e^{in\theta} = \sum_{m=-\infty}^{\infty} H_{m-n}^{(1)}(2\kappa b) J_m(\kappa r') e^{im\theta'} \quad (15)$$

Thus, the following translation of coordinate systems can be obtained:

$$\begin{aligned} &\exp(\beta r' \cos \theta') \sum_{n=-\infty}^{\infty} H_n^{(1)}(\kappa r') \exp(in\theta') \exp(-i\omega t) \\ &= \exp[\beta(2b + r \cos \theta)] \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} (-1)^{m-n} H_{m-n}^{(1)}(2\kappa b) \\ &\quad \times J_m(\kappa r) \exp(im\theta) \exp(-i\omega t) \end{aligned} \quad (16)$$

where $r' = \sqrt{r^2 + 4b^2 + 4rb \cos \theta}$, $\cos \theta' = (r'^2 + 4b^2 - r^2)/4br$.

$$\begin{aligned} &\exp(-\beta r \cos \theta) \sum_{n=-\infty}^{\infty} H_n^{(1)}(\kappa r) \exp(in\theta) \exp(-i\omega t) \\ &= \exp[\beta(2b - r' \cos \theta')] \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} H_{m-n}^{(1)}(2\kappa b) J_m(\kappa r') \\ &\quad \times \exp(im\theta') \exp(-i\omega t) \end{aligned} \quad (17)$$

where $r = \sqrt{r'^2 + 4b^2 - 4r'b \cos \theta'}$, $\cos \theta = -(r'^2 + 4b^2 - r^2)/4br$.

Without loss of generality, the case that the cavity is free of traction is investigated. The boundary condition is that the radial shear stress is equal to zero, namely,

$$\tau_{rz}|_{r=a} = \mu \frac{\partial u}{\partial r} \Big|_{r=a} = 0 \quad (18)$$

4 Determination of Mode Coefficients and Dynamic Stress Concentration

Multiple scattering of elastic waves takes place between the real and image cavities. By satisfying the boundary conditions around the cavities, mode coefficients of elastic waves are determined. When $l=1$ and the time factor is omitted, the relations among every mode coefficient of scattered waves are written as

$$\begin{aligned} A_n^1 &\left\{ \beta \cos \theta H_n^{(1)}(\kappa a) - \frac{1}{a} [n H_n^{(1)}(\kappa a) - \kappa a H_{n+1}^{(1)}(\kappa a)] \right\} \\ &= -i^n \left\{ \beta \cos \theta J_n(\kappa a) - \frac{1}{a} [n J_n(\kappa a) - \kappa a J_{n+1}(\kappa a)] \right\} \end{aligned} \quad (19)$$

$$\begin{aligned} B_n^1 &\left\{ \beta \cos \theta' H_n^{(1)}(\kappa a) + \frac{1}{a} [n H_n^{(1)}(\kappa a) - \kappa a H_{n+1}^{(1)}(\kappa a)] \right\} \\ &= -i^{-n} \left\{ \beta \cos \theta' J_n(\kappa a) + \frac{1}{a} [n J_n(\kappa a) - \kappa a J_{n+1}(\kappa a)] \right\} \end{aligned} \quad (20)$$

when $l=2, 3, \dots, \infty$ and the time factor is omitted, the relations among every mode coefficient of scattered waves are written as

$$\begin{aligned} A_n^l \exp(-\beta a \cos \theta) &\left\{ -\beta \cos \theta H_n^{(1)}(\kappa a) \right. \\ &\quad \left. + \frac{1}{a} [n H_n^{(1)}(\kappa a) - \kappa a H_{n+1}^{(1)}(\kappa a)] \right\} \\ &= -B_n^{l-1} \exp[\beta(2b + a \cos \theta)] \left\{ \beta \cos \theta \sum_{m=-\infty}^{\infty} (-1)^{m-n} \right. \\ &\quad \times H_{m-n}^{(1)}(2\kappa b) J_m(\kappa r) + \sum_{m=-\infty}^{\infty} (-1)^{m-n} H_{m-n}^{(1)} \\ &\quad \times (2\kappa b) \frac{1}{a} [m J_m(\kappa a) - \kappa a J_{m+1}(\kappa a)] \left. \right\} \exp[i(m-n)\theta] \end{aligned} \quad (21)$$

$$\begin{aligned} B_n^l \exp(\beta a \cos \theta') &\left\{ \beta \cos \theta' H_n^{(1)}(\kappa a) \right. \\ &\quad \left. + \frac{1}{a} [n H_n^{(1)}(\kappa a) - \kappa a H_{n+1}^{(1)}(\kappa a)] \right\} \\ &= -A_n^{l-1} \exp[\beta(2b - a \cos \theta')] \\ &\quad \times \left\{ -\beta \cos \theta' \sum_{m=-\infty}^{\infty} H_{m-n}^{(1)}(2\kappa b) J_m(\kappa a) + \sum_{m=-\infty}^{\infty} H_{m-n}^{(1)} \right. \\ &\quad \times (2\kappa b) \frac{1}{a} [m J_m(\kappa a) - \kappa a J_{m+1}(\kappa a)] \left. \right\} \\ &\quad \times \exp[i(m-n)\theta'], \quad (l=2, 3, \dots, \infty) \end{aligned} \quad (22)$$

Equations (19)–(22) are the algebra equations determining the mode coefficients A_n^l and B_n^l of the scattered waves.

In the following analysis, it is convenient to make the variables dimensionless. To accomplish this step, we may introduce a characteristic length a , where a is the radius of the cavity. The following dimensionless variables and quantities have been chosen for computation: the incident wave number is $ka=0.01-3.0$, the rela-

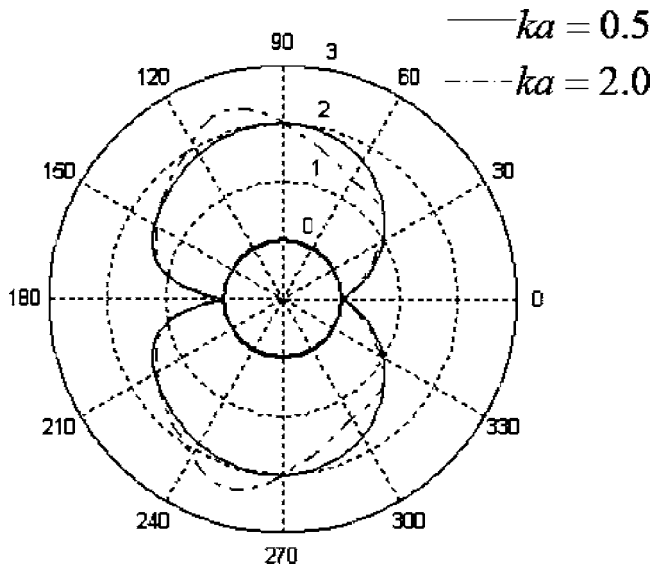


Fig. 2 Distribution of dynamic stress concentration factor around the cavity ($\beta=0, b/a=1.1$)

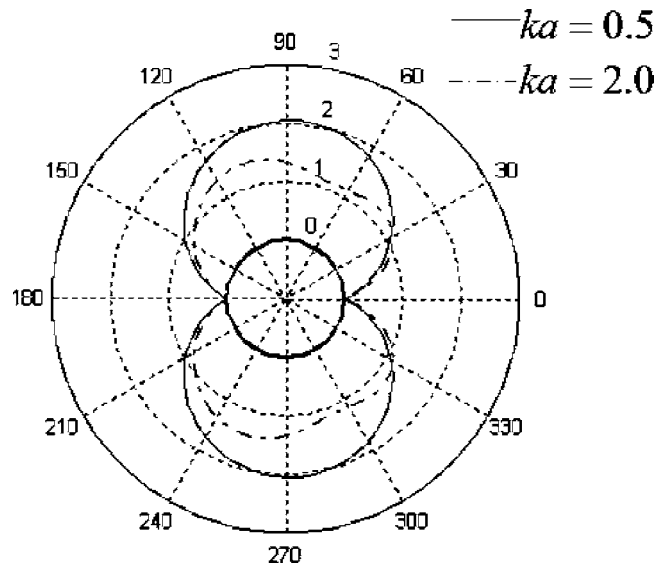


Fig. 3 Distribution of dynamic stress concentration factor around the cavity ($\beta=0, b/a=5.0$)

tive buried depth of the cavity beneath the surface of FGMs is $b/a=1.1-10$, the nonhomogeneity parameter is $\beta a=-0.50-0.50$.

According to the definition, the dynamic stress concentration factor (DSCF) is the ratio of the hoop shear stress around the cavity and the maximum stress [14]. Thus, the DSCF around the circular cavity in FGMs is expressed as

$$\text{DSCF} = \tau_{\theta z}^* = \left| \frac{\tau_{\theta z}}{\tau_0} \right| \quad (23)$$

$$\begin{aligned} \tau_{\theta z}^* = \frac{1}{\kappa} \exp(-\beta a \cos \theta) & \left\{ \beta a \sin \theta \sum_{n=-\infty}^{\infty} i^n J_n(\kappa a) \right. \\ & + \sum_{n=-\infty}^{\infty} i^{n+1} n J_n(\kappa a) \left. \right\} \exp(in\theta) + \frac{1}{\kappa} \exp(-\beta a \cos \theta) \\ & \times \left\{ \beta a \sin \theta \sum_{l=1}^{\infty} \sum_{n=-\infty}^{\infty} A_n^l H_n^{(1)}(\kappa a) + \sum_{l=1}^{\infty} \sum_{n=-\infty}^{\infty} i n A_n^l H_n^{(1)}(\kappa a) \right. \\ & \times (\kappa a) \left. \right\} \exp(in\theta) + \frac{1}{\kappa} \exp[\beta(2b + a \cos \theta)] \\ & \times \left\{ -\beta a \sin \theta \sum_{l=1}^{\infty} \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} (-1)^{m-n} B_n^l H_{m-n}^{(1)}(2\kappa b) J_m(\kappa a) \right. \\ & + \sum_{l=1}^{\infty} \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} (-1)^{m-n} i m B_n^l H_{m-n}^{(1)}(2\kappa b) J_m(\kappa a) \left. \right\} \exp(im\theta) \end{aligned} \quad (24)$$

5 Numerical Examples

Fatigue failures often occur in regions with high stress concentration; thus, an understanding of the distribution of the dynamic stress is very useful in structural design. According to the expression of DSCF, the DSCFs around the circular cavity are computed. It is found that the truncations after $l=10$ and $n=m=12$ gives practically adequate results at any desired frequency.

Figures 2 and 3 illustrate the angular distribution of the DSCFs around the circular cavity with $\beta=0$ for the cases of $b/a=1.1$ and $b/a=5$, respectively. It is clear that when the distance is b/a

$=1.1$ because of the multiple scattering between the cavity and the edge, the DSCFs at the positions near the edge are greater than that at the symmetrical positions about the y-axis, whereas at the positions of $\theta=0, \pi$, the DSCFs are minimum. When the variables are $ka=0.5, b/a=5.0$, the maximum DSCFs are at the position of $\theta=\pi/2, 3\pi/2$, and the angular distributions of DSCFs are symmetric about both axes. The above results are consistent with those in literature [3].

Figures 4–7 display the angular distribution of the DSCFs around the circular cavity when the nonhomogeneous parameter β and the value of b/a are different. It can be seen that when the incident frequent is the same, the greater the nonhomogeneity parameter β , the greater the maximum dynamic stress around the cavity is.

In contrast to $\beta=0$, if the nonhomogeneity parameter is $\beta<0$, the maximum dynamic stress increases very little; however, the position of it has a trend of shifting toward the shadow side of the cavity. Because of the effect of boundary, the trend of shifting is

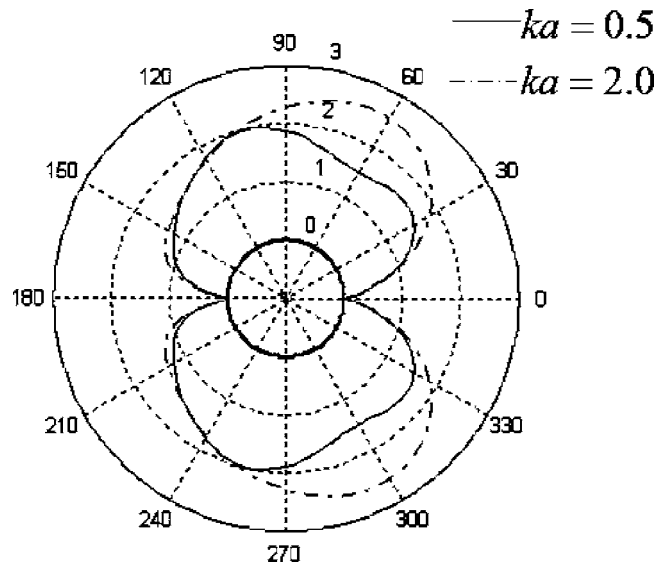


Fig. 4 Distribution of dynamic stress concentration factor around the cavity ($\beta a=-0.2, b/a=1.1$)

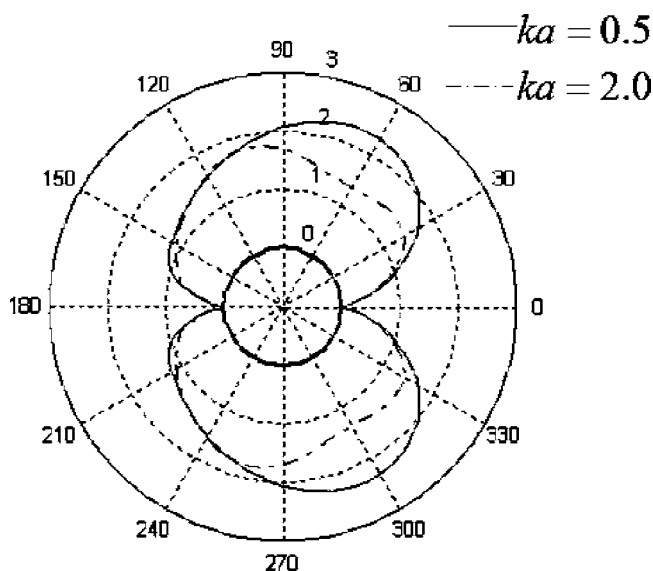


Fig. 5 Distribution of dynamic stress concentration factor around the cavity ($\beta a = -0.2, b/a = 5.0$)

more evident when the distance b/a is comparatively great. In the case of $ka = 0.5, b/a = 5.0$, the maximum dynamic stress is near the position of $\theta = \pi/3, 5\pi/3$.

If the nonhomogeneity parameter is $\beta > 0$, then the maximum dynamic stress increases greatly, and the position of it has a trend of shifting toward the illuminated side of the cavity. The trend of shifting is more evident when the distance b/a is comparatively little. In the case of $ka = 0.5, b/a = 5.0$, the maximum dynamic stress is near the position of $\theta = 2\pi/3, 4\pi/3$.

Figures 8 and 9 show the effect of nonhomogeneity parameter βa on DSCFs at the position of $\theta = \pi/2$ as a function of the dimensionless wave number ka for the cases of $b/a = 1.1$ and $b/a = 5.0$. From Figs. 8 and 9, it can be seen that at the position of $\theta = \pi/2$, if the nonhomogeneity parameter is $\beta < 0$, then the fluctuation of the DSCFs is little as the dimensionless wave number varies, and the distance b/a has less effect on the dynamic stress at $\theta = \pi/2$. If the nonhomogeneity parameter is $\beta > 0$, the fluctuation of the DSCFs is great as the dimensionless wave number

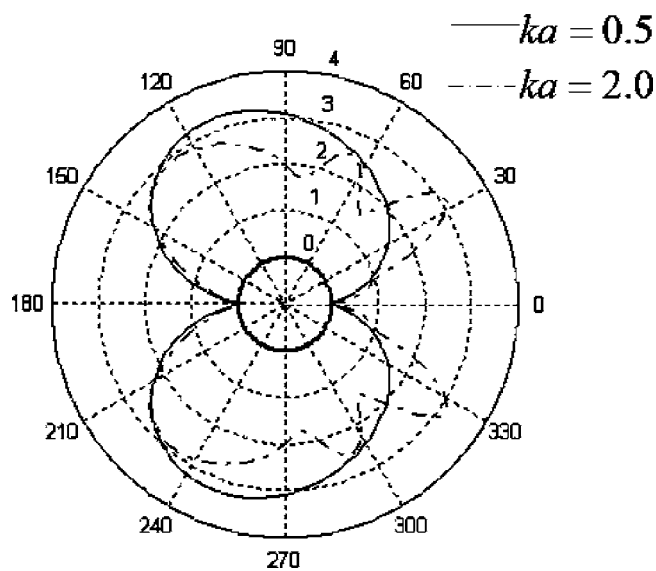


Fig. 7 Distribution of dynamic stress concentration factor around the cavity ($\beta a = 0.2, b/a = 5.0$)

varies, the greater the nonhomogeneity parameter, the more evident the fluctuation is. The effect of the distance on the dynamic stress is also great when $\beta > 0$. It is interesting to note that in the case of $\beta > 0$ and the distance b/a is comparatively little, the magnitude of the fluctuation becomes bigger and bigger as the dimensionless wave number (frequency of loading) increases.

In Figs. 10 and 11, we respectively plot the effect of dimensionless wave number ka on DSCFs at the position of $\theta = \pi/2$ as a function of the nonhomogeneity parameter βa for the cases of $b/a = 1.1$ and $b/a = 5.0$. It can be seen that when the nonhomogeneity parameter is less than a certain number β^* , the DSCFs decrease with the increase of dimensionless wave number (frequency of loading), and the magnitude of the variation is comparatively little. The DSCFs also show very little variation with the increase of nonhomogeneity parameter, especially when the values of b/a and wave number are great, the DSCFs show no variation. However, when the nonhomogeneity parameter is $\beta > \beta^*$, the variation of DSCFs is great as the wave number (frequency of loading) changes, and the wave number corresponding

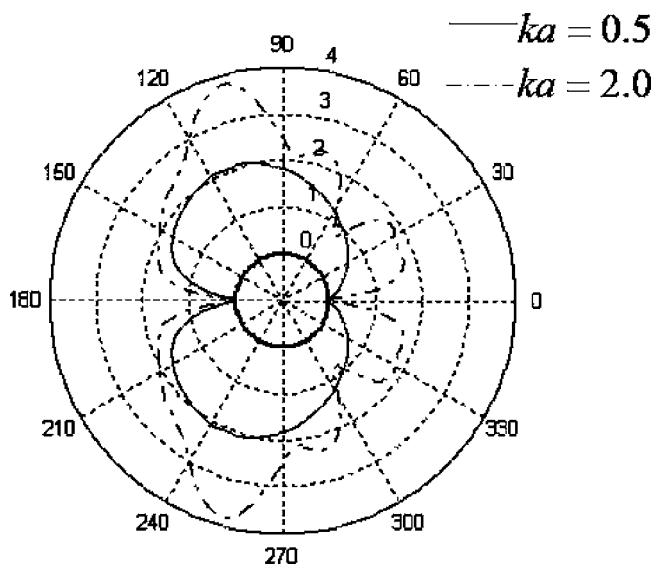


Fig. 6 Distribution of dynamic stress concentration factor around the cavity ($\beta a = 0.2, b/a = 1.1$)

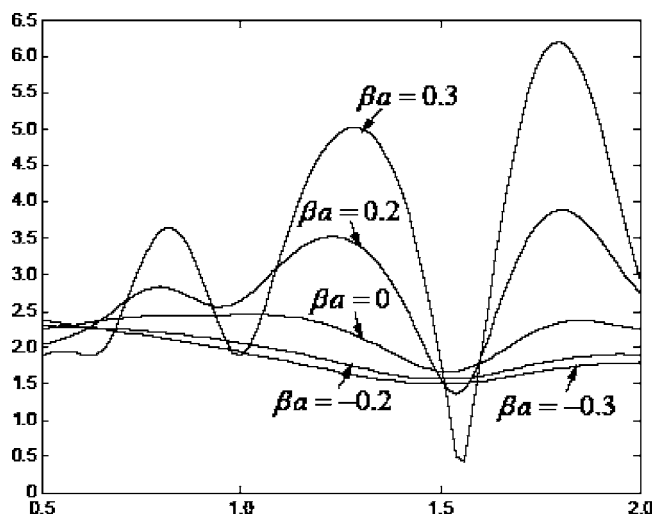


Fig. 8 Effect of nonhomogeneity parameter on dynamic stress concentration factor with $\theta = \pi/2, b/a = 1.1$

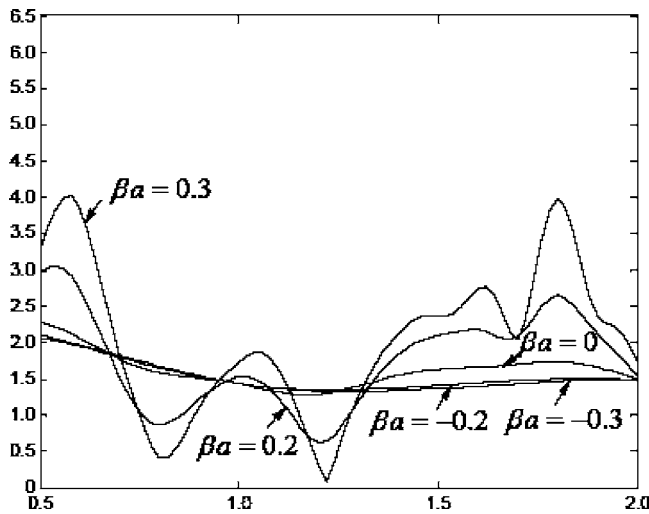


Fig. 9 Effect of nonhomogeneity parameter on dynamic stress concentration factor with $\theta = \pi/2, b/a = 5.0$

to the maximum dynamic stress has no disciplinary. It can be found that the greater the distance b/a , the greater the corresponding value of β^* is.

It is interesting to note that if the nonhomogeneity parameter satisfies $\beta < \beta^*$ and the wave number is greater than a certain number, the DSCF at the position of $\theta = \pi/2$ approaches a constant value. The constant value changes as the value of b/a varies. When $b/a = 1.1$, it is ~ 1.75 , and when $b/a = 5.0$, it is about ~ 1.45 .

6 Conclusions

The elastodynamic problem of a circular cavity in a semi-infinite functionally graded material under anti-plane impact load is analyzed by employing image method and wave functions expansion method. The case that the cavity is free of traction is investigated. The analytical solution and numerical solution of this problem are presented. For the homogeneous materials, our results are in good agreement with the solutions in previous literatures.

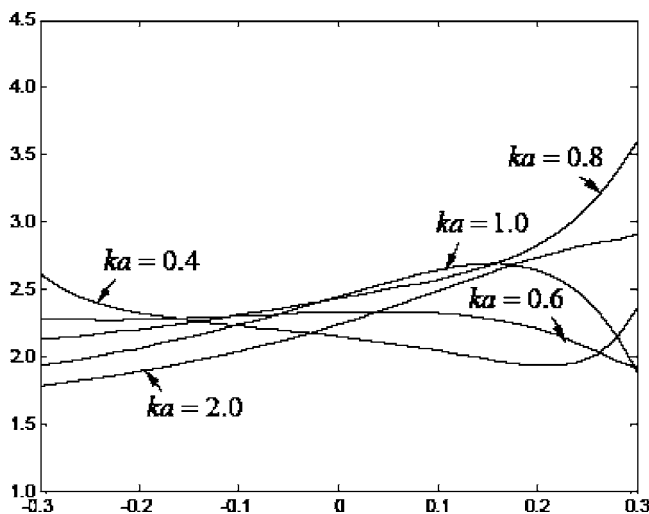


Fig. 10 Effect of incident wave number parameter on dynamic stress concentration factor with $\theta = \pi/2, b/a = 1.1$

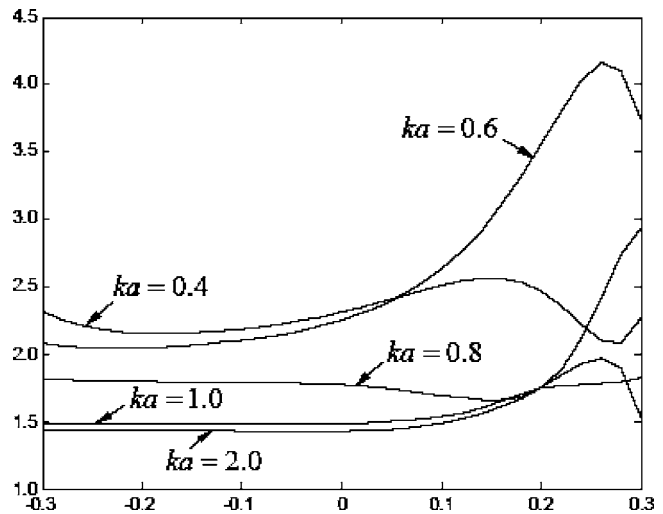


Fig. 11 Effect of incident wave number parameter on dynamic stress concentration factor with $\theta = \pi/2, b/a = 5.0$

The nonhomogeneity parameter of materials has great influence on the value and distribution of the dynamic stress concentration factors around the cavity. When the nonhomogeneity parameter is < 0 , the effect of it on the maximum dynamic stress is little. When the nonhomogeneity parameter is > 0 , the effect is greater, and the maximum dynamic stress increases with the increase of nonhomogeneity parameter. With the increase of the values of b/a , the effect of nonhomogeneity parameter and wave number on the dynamic stress decreases. Thus, to reduce the dynamic stress and avoid fatigue failures of structures, it is proposed that the nonhomogeneity parameter should be < 0 in the x direction in Fig. 1, namely, the shear modulus and density of semi-infinite functionally graded materials decrease in the x direction. The smaller the value of b/a , the less the nonhomogeneity parameter should be. When the value of b/a is smaller, the maximum dynamic stress around the cavity increases greatly with a small increase of the frequency of the impact load. Thus, we should choose a greater value of b/a , when designing the semi-infinite functionally graded materials under higher frequency.

Acknowledgment

The work is supported by the National Natural Science Foundation of China (Grant No. 10572045) and the Outstanding Youth Foundation of Hei Longjiang Province of China (Grant No. JC-9).

References

- [1] Gray, L. J., Kaplan, T., and Richardson, J. D., 2003, "Green's Functions and Boundary Integral Analysis for Exponentially Graded Materials: Heat Conduction," *ASME J. Appl. Mech.*, **70**, pp. 543–549.
- [2] Kuo, H.-Y., and Chen, T., 2005, "Steady and Transient Green's Functions for Anisotropic Conduction in an Exponentially Graded Solid," *Int. J. Solids Struct.*, **42**, pp. 1111–1128.
- [3] Rice, J. M., and Sadd, M. H., 1984, "Propagation and Scattering of SH-Waves in Semi-Infinite Domains Using a Time-Dependent Boundary Element Method," *ASME J. Appl. Mech.*, **51**, pp. 641–645.
- [4] Li, C., and Weng, G. J., 2000, "Dynamic Stress Intensity Factor of a Cylindrical Interface Crack With a Functionally Graded Interlayer," *Mech. Mater.*, **33**, pp. 325–333.
- [5] Ueda, S., 2001, "The Surface Crack Problem for a Layered Plate With a Functionally Graded Nonhomogeneous Interface," *Int. J. Fract.*, **110**, pp. 189–204.
- [6] Rousseau, C.-E., and Tippur, H. V., 2001, "Influence of Elastic Gradient Profiles on Dynamically Loaded Functionally Graded Materials: Cracks Along the Gradient," *Int. J. Solids Struct.*, **38**, pp. 7839–7856.

- [7] Chan, Y.-S., Paulino, G. H., and Fannjiang, A. C., 2001, "The Crack Problem for Nonhomogeneous Materials Under Antiplane Shear Loading—A Displacement Based Formulation," *Int. J. Solids Struct.*, **38**, pp. 2989–3005.
- [8] Sladek, J., Sladek, V., and Zhang, C., 2005, "A Meshless Local Boundary Integral Equation Method for Dynamic Anti-Plane Shear Crack Problem in Functionally Graded Materials," *Eng. Anal. Boundary Elem.*, **29**, pp. 334–342.
- [9] Pao, Y. H., and Chao, C. C., 1964, "Diffractions of Flexural Waves by a Cavity in an Elastic Plate," *AIAA J.*, **2**(11), pp. 2004–2010.
- [10] Kung, G. C. S., 1964, "Dynamical Stress Concentration in an Elastic Plate," M.S. thesis, Cornell University, Ithaca.
- [11] Klyukin, I. I., Klyukin, I. I., Satedkim, B. H., and Tirbln, D. C., 1964, "Scattering of Flexural Waves by Antivibrators on a Plate," *Sov. Phys. Acoust.*, **10**, pp. 49–53.
- [12] Hayir, A., and Bakirtas, I., 2004, "A Note on Plate Having a Circular Cavity Excited by Plane Harmonic SH Waves," *J. Sound Vib.*, **271**, pp. 241–255.
- [13] Stratton, J. A., 1941, *Electromagnetic Theory*, McGraw-Hill, New York, pp. 372–374.
- [14] Pao, Y.-H., and Mow, C. C., 1973, *Diffraction of Elastic Waves and Dynamic Stress Concentrations*, Crane, Russak, New York, Chap. 4.

F. D. Fischer
 Institute of Mechanics,
 Montanuniversität Leoben,
 Franz-Josef-Straße 18,
 A-8700 Leoben, Austria,
 Erich Schmid Institute for Materials Science,
 Austrian Academy of Sciences,
 Jahnstrasse 12,
 A-8700 Leoben, Austria
 e-mail: mechanik@unileoben.ac.at

J. Svoboda
 Institute of Physics of Materials,
 Academy of Sciences of the Czech Republic,
 Žitkova 22, CZ-616 62,
 Brno, Czech Republic
 e-mail: svobj@ipm.cz

A Note on the Principle of Maximum Dissipation Rate

The Principle of Maximum Dissipation Rate (PMD) can be exploited to derive homogeneous kinetic rate laws for the internal variables. A “normality structure” expressing the rates of the internal variables as normal to convex functions (entropy production rate, dissipation function as flow potentials) in the space of the conjugate thermodynamic forces is a direct consequence of the PMD. This paper can be considered as a note to Yang et al., 2005, ASME J. Appl. Mech., 72, pp. 322–329. [DOI: 10.1115/1.2722304]

Keywords: principle of maximum dissipation rate, normality structure, flow potentials, plasticity

1 Introduction

The Principle of Maximum Dissipation Rate (PMD) goes back to Onsager [1] in 1931 who introduced it for the case of transient heat conduction. Ziegler [2–5] demonstrated, that Onsager’s original formulation of the PMD can be generalized to arbitrary processes obeying the laws of linear nonequilibrium thermodynamics. Classical relations of applied physics, e.g., Ohm’s law of electric resistance, Hagen/Poiseuille’s law for laminar flow, and Fick’s first law of diffusion, can be derived in a straightforward way. During the last 15 years the PMD has been applied to the development of models in materials science [6–15]. The authors have demonstrated that the PMD seems to be a handy tool for the solution of practical problems of the thermodynamics of irreversible processes, which are often unsolvable or solvable only with great complications in the conventional way. However, apparently the PMD, its generalization, and its application to modeling have still not penetrated sufficiently into the material mechanics community. Most recently a series of papers by Yang et al. [16–19] has appeared which looks on homogeneous kinetic rate laws derived from flow potentials yielding a normality structure and relates these laws to the PMD in the sense of a constraint. These papers have motivated us to show that homogeneous kinetic rate laws with a normality structure can directly be derived from the PMD. This concept is valid at any length scale (e.g., also on the microlevel) and can be considered as a further justification of introducing a normality structure to the evolution equations as pioneered by Rice [20,21].

2 Problem Description

2.1 Definitions. We concentrate on a thermomechanical system whose internal structure is described by a set of internal variables ξ_i , $i=1, \dots, n$ and their rates or fluxes $\dot{\xi}_i$ being conjugate to the thermodynamic forces f_i defined per unit volume. Classical continuum thermodynamics teaches that the dissipation rate follows as

$$d = \sum_{i=1}^n \dot{\xi}_i f_i \quad (1)$$

and the entropy production rate as

$$\eta_p = d/\vartheta \quad (2)$$

with ϑ being the temperature. For systems without diffusive fluxes the f_i can be replaced by $-\partial\varphi/\partial\xi_i$ with φ being the specific Helmholtz free energy. In the case of diffusive fluxes the reader is referred to [22–24]. In the case of an elastic-plastic material the ξ_i are the components of the plastic strain state ϵ^p , the equivalent plastic strain ϵ_{pv} , and strain-type quantities κ^p corresponding to the back stress. The conjugate forces to the fluxes $\dot{\xi}_i$ are the components of the stress state σ , the change in the yield stress (isotropic hardening) $\sigma_s - \sigma_s^0$ with σ_s^0 being the initial value of the yield stress, and the backstress $\mathbf{a}$. The strain and stress components are supposed to be collected in vectors ϵ^p , κ^p and σ , $\mathbf{a}$, resp., see the Voigt notation.

The total entropy production function p in a system is a functional of the fluxes $\dot{\xi}_i$. Since for a system out of thermodynamic equilibrium the specific entropy production rate p must be a positive function for arbitrary fluxes (i.e., second Law of Thermodynamics), the Taylor expansion of p , denominated as p_2 with respect to the $\dot{\xi}_i$, truncated after three terms, must yield

$$p_2 = \sum_{i=1}^n \sum_{j=1}^n \dot{\xi}_i R_{ij} \dot{\xi}_j > 0 \quad (3)$$

The matrix R_{ij} is symmetric since $R_{ij} = 1/2(\partial^2 p / \partial \dot{\xi}_i \partial \dot{\xi}_j)$ and, due to $p > 0$ for any $\dot{\xi}_i \neq 0$, positive definite. The components R_{ij} can be functions of the thermodynamic forces f_ℓ , $\ell=1, \dots, n$ and the temperature ϑ . Note that p_2 approaches p near the thermodynamic equilibrium since terms with powers higher than 2 are neglected by using p_2 instead of p . For the sake of simplicity we assume a temperature field constant in space and time, at least within each incremental step. This means we exclude heat conduction.

An orthogonal point-transformation to a coordinate system formed by the eigenvectors of R_{ij} allows us to rewrite p_2 with $\dot{\xi}_i'$ as the transformed variables as

$$p_2 = \sum_{i=1}^n \lambda_i \dot{\xi}_i'^2 \quad (4)$$

The quantities λ_i , $i=1, \dots, n$, are the positive eigenvalues of p_2 . A further transformation $\dot{\xi}_i = \sqrt{\lambda_i} \dot{\xi}_i'$ yields

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received July 5, 2005; final manuscript received December 28, 2006. Review conducted by Zhigang Suo.

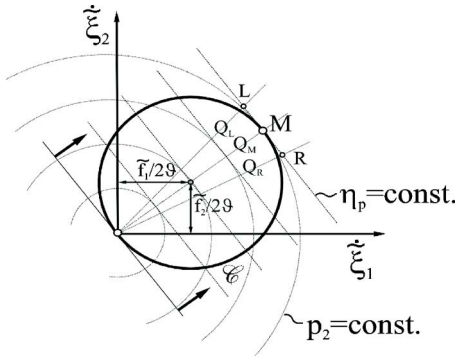


Fig. 1 Identifying the position of the maximum Q_M of Q in the $\dot{\xi}_1, \dot{\xi}_2$ plane; the lines $\eta_p = \text{const.}$ and $p_2 = \text{const.}$ are contour lines of the function to be maximized along $\mathcal{C}$

$$p_2 = \sum_{i=1}^n \dot{\xi}_i^2 \quad (5)$$

This operation allows us to transform p_2 in “diagonal form.” A generalized form of p_2 is dealt with in Sec. 2.3.

A comparison of (2)–(5) furnishes the necessary equivalence or energy conservation condition

$$p_2(\dot{\xi}_i, \dot{\xi}_j) = \eta_p \quad (6)$$

if no further dissipation process is active in addition to that controlled by the $\dot{\xi}_i, i = 1, \dots, n$.

2.2 Realization of the PMD. We maximize now the entropy production rate η_p (2) with respect to the fluxes $\dot{\xi}_i$, constrained by the side condition (6) yielding

$$\delta(\eta_p + \alpha(p_2 - \eta_p)) = 0 \quad (7)$$

As outlined in [7,11] and in detail in [25], the Lagrange multiplier α is found to be -1 resulting finally in a free extremum of $\varphi_p = 2\eta_p - p_2$ yielding

$$\sum_{j=1}^n R_{ij} \dot{\xi}_j = \frac{f_i}{\vartheta} \quad (8)$$

or by inversion

$$\dot{\xi}_i = \sum_{j=1}^n R_{ij}^{(-1)} \left(\frac{f_j}{\vartheta} \right) \quad (9)$$

The evolution Eq. (9) is homogeneous of order 1 in the forces $f_i, i = 1, \dots, n$, if the coefficients $R_{ij}^{(-1)}$ are independent of f_i . Inserting relation (9) into (3) allows us to express the dissipation function by the forces f_i as

$$p_2 = \sum_{i=1}^n \sum_{j=1}^n \left(\frac{f_i}{\vartheta} \right) R_{ij}^{(-1)} \left(\frac{f_j}{\vartheta} \right) \quad (10)$$

The components $R_{ij}^{(-1)}$ correspond to the matrix $\mathbf{R}^{-1}$ which is the inverse matrix of $\mathbf{R}$ corresponding to R_{ij} . Since $\mathbf{R}$ is positive definite, also $\mathbf{R}^{-1}$ is positive definite. Of course, the same result must be obtained if p_2 is maximized with the side condition (6).

The realization of the PMD can easily be shown for a system with two fluxes $\dot{\xi}_1, \dot{\xi}_2$ with p_2 in the diagonal form (5) as

$$p_2 = \dot{\xi}_1^2 + \dot{\xi}_2^2 \quad \eta_p = (\tilde{f}_1/\vartheta)\dot{\xi}_1 + (\tilde{f}_2/\vartheta)\dot{\xi}_2$$

The function η_p represents a plane with contour lines parallel to the direction $\dot{\xi}_2 = -\tilde{f}_1/\tilde{f}_2 \cdot \dot{\xi}_1$. The equivalence condition (6) yields a

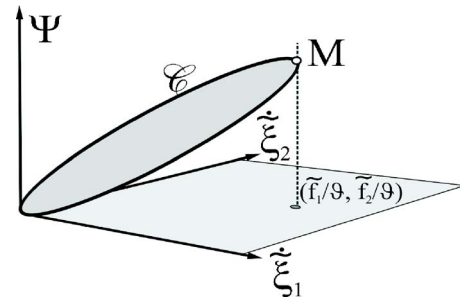


Fig. 2 Inclined projection of the function of the entropy production rate η_p with its maximum along $\mathcal{C}$

circular cylinder $\mathcal{C}$ with the basis circle according to

$$(\dot{\xi}_1 - \tilde{f}_1/2\vartheta)^2 + (\dot{\xi}_2 - \tilde{f}_2/2\vartheta)^2 = (\tilde{f}_1/2\vartheta)^2 + (\tilde{f}_2/2\vartheta)^2$$

As illustrated in Figs. 1 and 2 the maximum value of the entropy production is positioned at the point M with the coordinates $\dot{\xi}_1 = \tilde{f}_1/\vartheta, \dot{\xi}_2 = \tilde{f}_2/\vartheta$. If p_2 is maximized instead of η_p , the contour lines are circles which touch the equivalence condition exactly in M .

It is interesting to observe that a constant value of η_p yields for any variation of $\dot{\xi}_1, \dot{\xi}_2$ an increase of the dissipation function, marked by the points L and R in Fig. 1. This finally means that on the one hand a maximum dissipation rate is sought which is on the other hand a minimum for all admissible fluxes yielding the same entropy production rate.

2.3 Generalization of the PMD. To obtain homogeneous evolution laws of the order q as discussed by Yang et al. [16–19] the total entropy production p and p_2 , resp., is replaced by p_q as

$$p_q = \sum_{i=1}^n \lambda_i \dot{\xi}_i^{(q+1)/q}, \quad (11)$$

note that the primes to $\dot{\xi}_i$ are omitted. The quantities λ_i are assumed to be positive, and q is selected so that $p_q > 0$ is guaranteed. Here one should keep in mind that $\dot{\xi}_i$ may be positive or negative. Looking for the extremum of $\eta_p + \alpha(p_q - \eta_p)$ yields $\alpha = -q$ and $\varphi_q = ((q+1)/q)\eta_p - p_q$ to be freely extremized.

The first derivatives of φ_q with respect to $\dot{\xi}_i$, equated to zero, yield

$$\frac{f_i}{\vartheta} = \lambda_i \dot{\xi}_i^{1/q} \quad \dot{\xi}_i = \left(\frac{f_i}{\vartheta \lambda_i} \right)^q \quad (12)$$

The evolution Eq. (12) is now homogeneous of the order q in the forces $f_i, i = 1, \dots, n$. Inserting (12) into (11) results in

$$p_q = \sum_{i=1}^n \lambda_i \left(\frac{f_i}{\vartheta \lambda_i} \right)^{q+1} \quad (13)$$

The second derivatives of φ_q with respect to $\dot{\xi}_i$ yield

$$\partial^2 \varphi_q / \partial \dot{\xi}_i^2 = -((q+1)/q^2) \lambda_i \dot{\xi}_i^{(q+1)/q-2}$$

The sign of these derivatives is negative for proper q values, pointing to a maximum.

3 The Normality Structure

3.1 The Flow Potential. The derivative of p_2 , Eq. (10) with respect to f_i/ϑ is

$$\frac{\partial p_2}{\partial(f_i/\vartheta)} = 2 \sum_{j=1}^n R_{ij}^{(-1)} \left(\frac{f_j}{\vartheta} \right) \quad (14)$$

A comparison with (9) yields

$$\dot{\xi}_i = \frac{1}{2} \frac{\partial p_2}{\partial(f_i/\vartheta)} \quad (15)$$

with the consequence that the function $p_2/2$ can be used as a “flow potential.” Since the vector formed by the n components $\dot{\xi}_i$ in an n -dimensional vector space is parallel to the normal to the surface $p_2 = \text{const}$ described by the vector components f_i/ϑ as coordinates, the application of PMD automatically yields a normality structure. This is also practically valid, if the $R_{ij}^{(-1)}$ are held constant for the next time increment, when Eq. (15) is applied to calculate the $\dot{\xi}_i$, and afterwards updated by the current values f_i/ϑ .

If we use p_q , Eq. (13), instead of (15) immediately

$$\dot{\xi}_i = \frac{1}{(q+1)} \frac{\partial p_q}{\partial(f_i/\vartheta)} \quad (16)$$

follows. This relation agrees exactly with the findings by Yang et al. [16–19] with respect to the normality structure. Q is their flow potential, ϕ is their dissipation function, $Q = \phi/(1+q)$, q is the order of homogeneity. Trivially, the use of p_2 yields in comparison with p_q that $q=1$. It should be noted that the current concept is also similar to that proposed by Edelen [26] whose formulation works with an inequality instead of the equivalence condition (6).

3.2 Application to Plasticity. Concentrating only on plasticity with isotropic hardening the dissipation rate follows as

$$d^p = \boldsymbol{\sigma}^T \cdot \dot{\boldsymbol{\epsilon}}^p = \mathbf{s}^T \cdot \dot{\boldsymbol{\epsilon}}^p \quad (17)$$

$\mathbf{s}$ is the deviator to $\boldsymbol{\sigma}$. The six thermodynamic forces are the six components of $\mathbf{s}$; note that only five independent components do exist for $\dot{\boldsymbol{\epsilon}}^p$, $\mathbf{s}$ due to deviatoric properties. During a plastification process the yield condition must be fulfilled enforcing

$$\mathbf{s}^T \cdot \mathbf{s} = 2/3 \cdot \sigma_s^2 \quad (18)$$

The value of the actual yield stress is σ_s which is assumed to be a given constant quantity, at least during a certain time interval (or a certain load increment). As entropy production rate (3) we introduce

$$p_2 = \frac{F(\sigma_s, \vartheta; t)}{\vartheta} (\dot{\boldsymbol{\epsilon}}^p)^T \cdot \dot{\boldsymbol{\epsilon}}^p$$

The function $F(\sigma_s, \vartheta; t)$ is positive and depends on σ_s , ϑ and may also depend explicitly on the time t , to describe e.g., primary creep. The matrix $\underline{\mathbf{R}}$ is simply $\mathbf{F}/\vartheta \cdot \underline{\mathbf{I}}$ with $\underline{\mathbf{I}} = \delta_{ij}$ being the unity matrix. Equation (9) yields for $\dot{\boldsymbol{\epsilon}}^p$ with $\underline{\mathbf{R}}^{-1} = \vartheta/\mathbf{F} \cdot \underline{\mathbf{I}}$

$$\dot{\boldsymbol{\epsilon}}^p = \frac{\vartheta}{F(\sigma_s, \vartheta; t)} \frac{\mathbf{s}}{\vartheta} = \frac{\mathbf{s}}{F(\sigma_s, \vartheta; t)}$$

If this relation is inserted in p_2 according to (10), then p_2 follows as

$$p_2 = \frac{1}{\vartheta F(\sigma_s, \vartheta; t)} \mathbf{s}^T \cdot \mathbf{s}$$

Differentiation with respect to $\mathbf{s}/\vartheta$ yields

$$\dot{\boldsymbol{\epsilon}}_p = \frac{1}{2} \frac{\partial p_2}{\partial(\mathbf{s}/\vartheta)} = \frac{\mathbf{s}}{F(\sigma_s, \vartheta; t)} \quad (19)$$

which identifies p_2 as a flow potential with $\dot{\boldsymbol{\epsilon}}_p$ being normal to p_2 . In the case of viscoplasticity $F(\sigma_s, \vartheta; t)$ is an a priori given function. In the case of time-independent flow plasticity $F(\sigma_s, \vartheta; t)$ can be written as the product $\sigma_s \tilde{F}(\vartheta; t)$. The positive function $\tilde{F}(\vartheta; t)$

has to be calibrated at each time instant to meet the yield condition (18). It turns out that $\tilde{F}(\vartheta; t)$ is equivalent to $2/(3\dot{\epsilon}_{pv})$.

If we assume the element $R_{ij}^{(-1)}$ in (10) to be a function of the f_i , e.g., in the form of $\sigma_s(\mathbf{s})\tilde{F}(\vartheta; t)$ not being constant, then the derivative of p_2 with respect to f_i/ϑ (being $\mathbf{s}/\vartheta$) follows with (10) and (18)

$$\begin{aligned} \frac{\partial p_2}{\partial(\mathbf{s}/\vartheta)} &= \frac{\partial}{\partial(\mathbf{s}/\vartheta)} \left(\frac{1}{\sigma_s(\mathbf{s})\tilde{F}(\vartheta; t)} \left(\frac{\mathbf{s}}{\vartheta} \right)^T \cdot \frac{\mathbf{s}}{\vartheta} \right) = \frac{\partial}{\partial(\mathbf{s}/\vartheta)} \left(\frac{2}{3\tilde{F}(\vartheta; t)} \frac{\sigma_s}{\vartheta^2} \right) \\ &= \frac{1}{\sigma_s(\mathbf{s})\tilde{F}(\vartheta; t)} \mathbf{s} \end{aligned} \quad (20)$$

Note that $\partial\sigma_s/\partial\mathbf{s} = 3/2 \cdot \mathbf{s}/\sigma_s$.

This result, differing in the derivation above by a factor 2, is not surprising, since p_2 is now homogeneous, not of order 2 in the forces, but a positive function being homogeneous of order 1 in the forces, now reducing q in Yang et al. [16] from 1 to 0. The normality structure itself is, however, conserved, if one compares (20) and (19).

An equivalent derivation can be performed if a deviatoric back-stress $\mathbf{a}$ does exist, too. Then $\mathbf{s}$ is to be replaced by $\mathbf{s} - \mathbf{a}$.

3.3 A Final Remark. The reader may be aware of the “Principle of Maximum Plastic Resistance” by Bishop and Hill, sometimes denoted as a “Principle of Maximum Plastic Dissipation,” outlined e.g., in the Handbook [27]. This principle works at the level of d^p , Eq. (16), by keeping the plastic strain rates fixed and looking via the principle of virtual power for statically admissible stress states $\tilde{\boldsymbol{\sigma}}$ with the dissipation $\tilde{d}^p = \tilde{\boldsymbol{\sigma}}^T \cdot \dot{\boldsymbol{\epsilon}}^p$. Then one can show that an admissible stress state $\tilde{\boldsymbol{\sigma}}$ may not obtain the critical resolved shear stress on all the slip systems active to produce $\dot{\boldsymbol{\epsilon}}^p$ according to the actual stress state $\boldsymbol{\sigma}$, yielding $\tilde{d}_p \leq d_p$. Any dissipation function or normality structure is not engaged. Furthermore, the quantity to be varied is the stress state corresponding to the f_i instead to the $\dot{\xi}_i$ corresponding to $\dot{\boldsymbol{\epsilon}}^p$.

4 Conclusion

The Principle of Maximum Dissipation Rate (PMD) can be used as a general principle for the derivation of evolution equations and delivers their corresponding coefficients in accordance with Onsager’s reciprocity law as well as their normality structure with respect to the entropy production function. If this function is a quadratic form of the thermodynamic fluxes, e.g., the plastic strain rates, the application of the PMD is demonstrated at the hand of viscoplasticity and classical flow plasticity coupled with a back stress tensor. For a more general formulation of the entropy production function, the evolution equations introduced by Yang et al. [16–19] with a higher order of homogeneity as well as their normality structure can be derived by the PMD.

References

- [1] Onsager, L., 1931, “Reciprocal Relations in Irreversible Processes I,” *Phys. Rev.*, **37**, pp. 405–426.
- [2] Ziegler, H., 1961, “Zwei Extremalprinzipien der irreversiblen Thermodynamik,” *Ing.-Arch.*, **30**, pp. 410–416.
- [3] Ziegler, H., 1963, *Some Extremum Principles in Irreversible Thermodynamics With Application to Continuum Mechanics*, I. N. Sneddon and R. Hill, eds., North-Holland, Amsterdam, Vol. 4, Chap. II, pp. 91–193.
- [4] Ziegler, H., 1977, *An Introduction to Thermomechanics*, North-Holland, Amsterdam.
- [5] Ziegler, H., and Wehrli, C., 1987, “On a Principle of Maximal Rate of Entropy Production,” *J. Non-Equil. Thermodyn.*, **12**, pp. 229–243.
- [6] Svoboda, J., Turek, I., and Sklenička, V., 1990, “Unified Thermodynamic Treatment of Cavity Nucleation and Growth in High Temperature Creep,” *Acta Metall. Mater.*, **38**, pp. 573–580.
- [7] Svoboda, J., and Turek, I., 1991, “On Diffusion-Controlled Evolution of Closed Solid State Thermodynamic Systems at Constant Temperature and Pressure,” *Philos. Mag. B*, **64**, pp. 749–759.
- [8] Cocks, A. C. F., Gill, S. P. A., and Pan, J. Z., 1999, “Modelling Microstructure Evolution in Engineering Materials,” *Advances in Applied Mechanics*, E. v. d.

- Giessen and Th. Y. Wu, eds., Academic, San Diego, Vol. 36, pp. 81–162.
- [9] Gill, S. P. A., Cornforth, M. G., and Cocks, A. C. F., 2001, “Modelling Microstructure Evolution in Engineering Materials,” *Int. J. Plast.*, **17**, pp. 669–690.
 - [10] Svoboda, J., and Lukáš, P., 2000, “Creep Deformation Modelling of Superalloy Single Crystals,” *Acta Mater.*, **48**, pp. 2519–2528.
 - [11] Svoboda, J., Fischer, F. D., Fratzl, P., and Kroupa, A., 2002, “Diffusion in Multi-Component Systems With No or Dense Sources and Sinks for Vacancies,” *Acta Mater.*, **50**, pp. 1369–1381.
 - [12] Fischer, F. D., Svoboda, J., and Fratzl, P., 2003, “A Thermodynamical Approach to Grain Growth and Coarsening,” *Philos. Mag.*, **83**, pp. 1075–1093.
 - [13] Svoboda, J., Gamsjäger, E., Fischer, F. D., and Fratzl, P., 2004, “Application of the Thermodynamic Extremal Principle to the Diffusional Phase Transformations,” *Acta Mater.*, **52**, pp. 959–967.
 - [14] Svoboda, J., Fischer, F. D., Fratzl, P., and Kozeschnik, E., 2004, “Modelling of Kinetics in Multi-Component Multi-Phase Multi-Particle Systems I.—Theory,” *Mater. Sci. Eng., A*, **385**, pp. 166–174.
 - [15] Kozeschnik, E., Svoboda, J., Fratzl, P., and Fischer, F. D., 2004, “Modelling of Kinetics in Multi-Component Multi-Phase Multi-Particle Systems II.—Numerical Solution and Application,” *Mater. Sci. Eng., A*, **385**, pp. 157–165.
 - [16] Yang, Q., Tham, L. G., and Svoboda, G., 2005, “Normality Structures with Homogeneous Kinetic Rate Laws,” *ASME J. Appl. Mech.*, **72**, pp. 322–329.
 - [17] Yang, Q., Chen, X., and Zhou, W. Y., 2005, “Relationship Between Normality Structure and Orthogonality Condition,” *Mech. Res. Commun.*, **32**, pp. 582–589.
 - [18] Yang, Q., Chen, X., and Zhou, W. Y., 2005, “Microscopic Thermodynamic Basis of Normality Structure of Inelastic Constitutive Relations,” *Mech. Res. Commun.*, **32**, pp. 590–596.
 - [19] Yang, Q., Chen, X., and Zhou, W. Y., 2005, “On Microscopic Thermodynamic Mechanisms of Damage Evolution Laws,” *Int. J. Damage Mech.*, **14**, pp. 261–293.
 - [20] Rice, J. R., 1971, “Inelastic Constitutive Relations for Solids: An Integral Variable Theory and its Application to Metal Plasticity,” *J. Mech. Phys. Solids*, **19**, pp. 433–455.
 - [21] Rice, J. R., 1975, “Continuum Mechanics and Thermodynamics of Plasticity in Relation to Microscale Deformation Mechanisms,” *Constitutive Equations in Plasticity*, A. S. Argon, ed., MIT Press, Cambridge, MA, pp. 23–79.
 - [22] Müller, I., and Ruggeri, T., 1993, *Extended Thermodynamics*, Springer-Verlag, New York.
 - [23] Fried, E., and Gurtin, M. E., 2004, “A Unified Treatment of Evolving Interfaces According to Small Deformations and Atomic Transport With Emphasis on Grain-Boundaries and Epitaxy,” *Advances in Applied Mechanics*, H. Aref, E. v. d. Giessen, eds., Elsevier, Amsterdam, Vol. 40, pp. 1–177.
 - [24] Fischer, F. D., and Simha, N. K., 2004, “Influence of Material Flux on the Jump Relations at a Singular Interface in a Multicomponent Solid,” *Acta Mech.*, **171**, pp. 213–223.
 - [25] Svoboda, J., Turek, I., and Fischer, F. D., 2005, “Application of the Thermodynamic Extremal Principle to Modeling of Thermodynamic Processes in Material Sciences,” *Philos. Mag.*, **85**, pp. 3699–3707.
 - [26] Edelen, D. G. B., 1972, “A Nonlinear Onsager Theory of Irreversibility,” *Int. J. Eng. Sci.*, **10**, pp. 481–490.
 - [27] Aernoudt, E., Van Houtte, P., and Leffers, T., 1993, “Deformation and Textures of Metals of Large Strain,” *Materials Science and Technology*, H. Mughrabi, ed., VCH, Weinheim, Vol. 6, pp. 89–136.

Analysis of One-Dimensional and Two-Dimensional Thin Film “Pull-in” Phenomena Under the Influence of an Electrostatic Potential

Gang Duan
Graduate Student

Kai-Tak Wan
Mechanical Engineering,
University of Missouri-Rolla,
Rolla, MO 65409-0050
e-mail: wankt@umr.edu

A thin one-dimensional rectangular or two-dimensional axisymmetric film is clamped at the perimeter. An electrostatic potential (V_0^*) applied to a pad directly underneath the film leads to a “pull-in” phenomenon. The electromagnetic energy stored in the capacitive film-pad dielectric gap is decoupled from the mechanical deformation of the film using the Dugdale-Barenblatt-Maugis cohesive zone approximation. The ratio of film-pad gap (g) to film thickness (h), or, $\gamma = g/h$, is found to play a crucial role in the electromechanical behavior of the film. Solution spanning a wide range of γ is found such that $V_0^* \propto \gamma^{3/2}$ for $\gamma < 0.5$ and $V_0^* \propto \gamma^{5/2}$ for $\gamma > 5$. The new model leads to new design criteria for MEMS-RF-switches. [DOI: 10.1115/1.2722311]

Keywords: MEMS, RF-switch, electrostatic potential, surface forces, pull-in phenomena

1 Introduction

When a thin film clamped at the perimeter is subjected to an external force (e.g., electrostatic potential, long-range intersurface forces), “pull-in” occurs when a tunable surface force reaches a threshold or when the film is brought into close proximity of a substrate. There are numerous applications of this phenomenon, e.g., microactuators [1,2], micropumps [3–5], and strain gauges [6]. In this paper, we will focus on the operation of a MEMS device and will allude to measurement of the range and magnitude of intersurface forces. In a typical MEMS-RF-switch, a mechanically suspended bridge is pulled by an electrostatic voltage (V_0) applied directly underneath an electrode pad (Fig. 1) [7–11]. When V_0 exceeds a certain pull-in threshold, V_0^* , the bridge makes direct contact with the pad so that either an “on” or “off” signal is induced; and when the voltage is removed, the bridge resumes its original undeformed configuration that induces the complementary signal. Note that the electrode pad falls short of the bridge span in virtually all actual RF switches, but the assumption of the same length is most commonly adopted by the literature. There is also a two-dimensional (2D) version of this switch where a circular bridge is clamped at the perimeter. To understand the device operation and to optimize the design parameters (e.g., dimension of the bridge), it is necessary to construct a rigorous elastic model for the electromechanical interaction.

The rudimentary “lumped model” assumes a rigid rectangular plate with one surface attached to an elastic recoil spring while another surface interacts with a rigid substrate via the attractive electrostatic forces. This simple parallel-plate capacitor model predicts a pull-in event (i.e., a spontaneous collapse of the bridge onto the electrostatic pad) to occur when the midspan of the bridge reaches 1/3 of the bridge-pad gap [12,13]. More sophisticated models became available recently to account for the fringing field as a result of the finite bridge width, residual stress due to

fabrication and thermal expansion, inclusion of air-cushion, etc. [14–16]. One major difficulty in formulating the exact electromechanical behavior is the *nonlinear* governing solid-mechanics equation, which forbids an analytical solution. In the literature, numerical approaches using variational method with series of predetermined orthogonal trial functions and finite element analysis (FEA) are devised to solve for the bridge profile and the associated device behavior [2,12,13,15,17,18]. Several limitations are noted: (i) these solutions do not agree with one another because some models are based on pure plate bending of the bridge, some on pure membrane-stretching, and others on some well-defined mathematical functions; (ii) since the normalized bridge profile takes on a fixed shape that is unable to account for mixed plate bending and membrane stretching; (iii) the numerical routine must be repeated for new design in device geometry and dimension; and (iv) the coupled electromagnetic and mechanical parameters do not lead to well-defined design criteria. A latest approach [19] is to adopt the Galerkin method where the electrostatic potential is expressed in a Taylor series with the terms higher than w^4 ignored. The method is also limited to a specific set of dimension and working parameters and might need to be repeated to cover a range of bridge stiffnesses and thicknesses. A comparison between the Galerkin method and our new model will be discussed.

In this paper, the electromagnetic and mechanical components will be *decoupled* based on an assumption that the electrostatic field in the bridge-pad gap is uniform, resulting in an *exact* analytical solution. Despite the inevitable inaccuracy involved, the new solution is capable of formulating new design criteria as the bridge gets thinner and shorter. The critical operational parameters, such as pull-in voltage and critical bridge-pad gap, will also be derived. Deviation resulting from the uniform field assumption will be assessed. Edge effects due to finite bridge width and anti-clastic deformation are ignored. The one-dimensional (1D) model will be extended to 2D.

2 Theory

Figure 1 shows a rectangular bridge mechanically clamped at the two opposite ends and suspended above an electrode pad that is a distance g below and has identical length as the bridge. The

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received April 27, 2006; final manuscript received September 18, 2006. Review conducted by Matthew R. Begley.

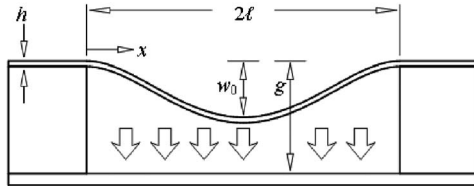
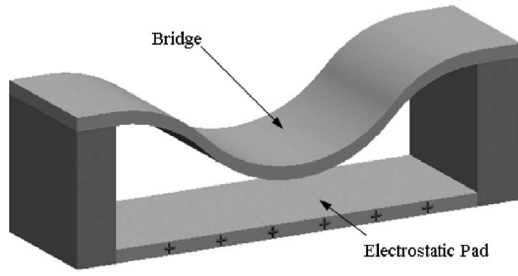


Fig. 1 Sketch of a typical MEMS-RF switch. The suspended bridge deforms in the presence of an electrostatic force induced by the electrode-pad directly underneath.

bridge is assumed to be free of prestress or residual stress and possesses a unit width, length 2ℓ , thickness h , elastic modulus E , Poisson's ratio ν , and flexural rigidity, $\kappa = Eh^3/12(1-\nu^2)$. An electrical potential, V_0 , is applied to the pad to set up a uniform electric field. The bridge is compelled and deformed by bending and a longitudinal membrane stress σ to a profile, $w(x)$, governed by the classical von Kármán equation [20,21],

$$\underbrace{-\kappa \nabla^4 w + (\sigma h) \nabla^2 w}_{\text{Mechanical deformation}} = \underbrace{-\left(\frac{\epsilon_0 V_0^2}{2}\right) \frac{1}{(g-w)^2}}_{\text{Electrostatic force}} \quad (1)$$

where ϵ_0 is the permittivity of free space and ∇^2 is the Laplacian operator in the rectilinear or curvilinear coordinate systems. The right-hand side of (1) denotes the electrostatic force on the bridge, while the left-hand side represents the mechanical response of the bridge in terms of the two deformation modes of plate bending and membrane stretching, respectively. Since $w(x)$ appears on both sides of (1), the coupled electromechanical equation leads to nonlinearity and thus forbids an analytical solution. To decouple the two components, the Dugdale-Barenblatt-Maugis cohesion zone approximation [22] is adopted here. The electrostatic force is replaced by a uniform mechanical pressure p , which is related to the applied voltage by averaging the traction over the bridge span. Equation (1) will first be solved for a 1D rectangular switch, followed by an extension to a 2D axisymmetric film where (1) remains valid for the radial profile, $w(r)$.

2.1 1D Rectangular Switch. For a 1D switch, a rectangular film is clamped at the opposite ends. A set of normalized parameters is defined in Appendix A.² Note that β gages the ratio of membrane stress to film rigidity such that (i) $\beta \approx 0$ corresponds to a plate-bending dominant deformation in a thick and stiff bridge

²Note that Ref. [19] normalizes w and V_0 using g instead of h as in the present work. It will become apparent that normalization using h will lead to a better description of the mixed bending-stretching bridge behavior.

and (ii) $\beta \rightarrow \infty$ refers to a membrane-stretching dominant deformation in a thin and flexible bridge. A few boundary conditions are noted,

(i) At the clamped ends:

$$w_{x=0} = 0 \quad \text{or} \quad \omega_{\xi=0} = 0 \\ (\partial w / \partial x)_{x=0} = 0 \quad \text{or} \quad (\partial \omega / \partial \xi)_{\xi=0} = 0$$

(ii) At the center:

$$(\partial w / \partial x)_{x=\ell} = 0 \quad \text{or} \quad (\partial \omega / \partial \xi)_{\xi=1} = 0$$

Applying the cohesive zone approximation, (1) becomes

$$-\kappa \nabla^4 w + (\sigma h) \nabla^2 w = -p \quad \text{or} \quad \frac{d^4 \omega}{d\xi^4} - \beta^2 \frac{d^2 \omega}{d\xi^2} = \rho \quad (2)$$

with the equivalent pressure

$$p = \frac{\epsilon_0 V_0^2}{4\ell} \int_0^{2\ell} \frac{dx}{[g-w(x)]^2} \quad \text{or} \quad \rho = v_0^2 \int_0^1 \frac{d\xi}{[\gamma - \omega(\xi)]^2} \quad (3)$$

Equation (2) can be reduced to a second-order linear differential equation [23] that leads to an *exact* analytical bridge profile

$$\omega = \rho \left(\frac{1}{\beta^3} \right) \left\{ \frac{1}{\tanh \beta} [\cosh(\beta \xi) - 1] - \sinh(\beta \xi) + \beta \left(\xi - \frac{\xi^2}{2} \right) \right\} \quad (4)$$

with a central deflection, $\omega_0 = \omega(\xi=1)$,

$$\omega_0 = \rho \left(\frac{1}{\beta^3} \right) \left\{ \frac{1}{\tanh \beta} (\cosh \beta - 1) - \sinh \beta + \frac{\beta}{2} \right\} \quad (5)$$

Note that (4) is a rigorous solution of (2), rather than a predetermined profile, as in most variational methods in the literature (cf. Table 1). Figure 2 shows the changing bridge profiles for a range of β . The volume of the reduced dielectric space between the bridge and pad is found by integrating (4)

$$V = \int_0^{2\ell} w dx \quad \text{or} \quad \vartheta = \int_0^1 2\omega d\xi = \rho \left(\frac{2}{\beta^4} \right) \left(1 + \frac{\beta^2}{3} - \frac{\beta}{\tanh \beta} \right) \quad (6)$$

The uniform membrane stress on the bridge can be found by elementary elasticity [23,24]

$$\sigma = \frac{1}{2\ell} \left(\frac{E}{1-\nu^2} \right) \int_0^{2\ell} \frac{1}{2} \left(\frac{dw}{dx} \right)^2 dx \quad \text{or} \quad \beta^2 = 6 \int_0^1 \left(\frac{\partial \omega}{\partial \xi} \right)^2 d\xi \quad (7)$$

Substituting (4) into (7) yields

$$\rho = \frac{\beta^4 \sinh \beta}{[(6 + \beta^2) \cosh(2\beta) - 9\beta \cosh \beta \sinh \beta - 6 - 4\beta^2]^{1/2}} \quad (8)$$

By eliminating β from (5) and (8), the mechanical response, $\rho(\omega_0)$, can be obtained, though it is a mathematically formidable task because of the transcendental functions $\sinh(x)$ and $\cosh(x)$. An alternative to derive the *exact* form of $\rho(\omega_0)$ is to trace a parametric plot of $\rho(\omega_0)$ by taking β as a varying parameter since both ρ and ω_0 are functions of β (Fig. 3). The bending to stretching transition can be expressed in an alternative manner as

$$\rho = k(\beta) \omega_0^{n(\beta)} \quad (9)$$

where both $k(\beta)$ and $n(\beta)$ are well defined functions of β . If $\rho(\omega_0)$ is shown in a log-log plot, $n(\beta)$ is the gradient and is defined as

$$n = \frac{d(\log \rho)}{d(\log \omega_0)} = \frac{\omega_0}{\rho} \left(\frac{d\rho}{d\omega_0} \right) \quad (10)$$

The exact form of $n(\beta)$ can be found by MATHEMATICATM, though it is too lengthy to be given here. Figure 4 shows n as a function

Table 1 Comparison of the pull-in parameter (w_0^*/g)

Methods	Pull-in (w_0^*/g)
1D and 2D lumped model [12,13] Assumption: rigid plates with one attached to an elastic spring and another stationary	1/3=0.3333
1D Variational method [15] Trial function: $w = w_0 \cos^2(\pi \xi)$	1/3=0.3333
1D Variational method [18] Trial functional: $w = w_0 \xi^2(\xi-1)^2$	~0.45
1D Energy method for multi-layered bridge [17] Trial function: $w = (w_0/2)[1 + \cos(2\pi \xi)]$	0.40–0.67
1D Galerkin method [19]	0.55 (zero residual stress) 0.42–0.63 (range of residual stress) 0.4545–0.6791 (force balance) 0.3970–0.6583 (energy balance)
1D Present work	~0.40
2D Variational method [2] Trial function: $w = C_1 J_0(\Omega_m^{1/2} r) + C_2 I_0(\Omega_m^{1/2} r)$	0.5723–0.7500 (force balance) 0.4633–0.7135 (energy balance)
2D Present work	

of β with $1 \leq n \leq 3$.

Deformation of the bridge is bounded by two limiting cases. In case of a thick and stiff bridge, the deformation is small ($w_0 < 0.5$), the membrane stress is negligible ($\sigma \approx 0$ and $\beta \approx 0$), and only plate bending is present. It can be easily shown that (4)–(6) and (10) reduce to $w_{\text{bend}} = (\rho/24)\xi^2(\xi-2)^2$, $\rho = 24w_0$, $\vartheta = (16/15)w_0$, and $n=1$, respectively, which is consistent with the classical Timoshenko's linear solution [24] shown in Figs. 2–4 as asymptotes. In case of a thin and flexible bridge with a zero bending inertia ($\kappa=0$), the deformation is large ($w_0 > 5$), the normalized membrane stress becomes infinite ($\beta \rightarrow \infty$) and only membrane stretching is present. The bridge behavior now becomes $w_{\text{stretch}} = (\rho/\beta^2)(\xi - \xi^2/2)$, $\rho = 16w_0^3$, $\vartheta = (4/3)w_0$, and $n=3$. Note that w_{stretch} is parabolic such that $(\partial w/\partial \xi)_{\xi=0}$ is undefined, which violates boundary condition (i). However, a film with zero flexural rigidity does not require a differentiable profile at the clamped

edges. Figures 2–4 show the membrane-stretching asymptotes. Note that $\rho(w_0)$ for membrane stretching is cubic (Fig. 3) and is consistent with our earlier results for rectangular film deformed by a central line load [23,25]. When the deformation is intermediate ($0.5 < w_0 < 5$), mixed bending stretching must be considered. The transition can be arbitrarily taken as the intersection between the two limiting cases in Fig. 3 and is roughly $w_0 \approx 1.20$. Here, $1 < n < 3$ and $16/15 < (\vartheta/w_0) < 4/3$.

There are two ways to investigate the electromechanical behavior of the RF switch: (i) The first and most common method is to balance the mechanical force due to bridge deformation (cf. (5) and (8)) and the electrostatic attraction due to applied voltage (3), and (ii) a balance of electromagnetic and mechanical energy involved. The stretching limit is chosen in this section to demonstrate the general behavior that is applicable also to mixed bending-stretching films. Figure 5 shows the mechanical and electrostatic forces for a range of applied voltage. When the applied voltage increases from null, there are two distinct intersections between the force curves at A and B as shown. It will become apparent that A corresponds to a stable configuration while B is unstable. As v_0 increases further, $(w_0)_A$ and $(w_0)_B$ move closer

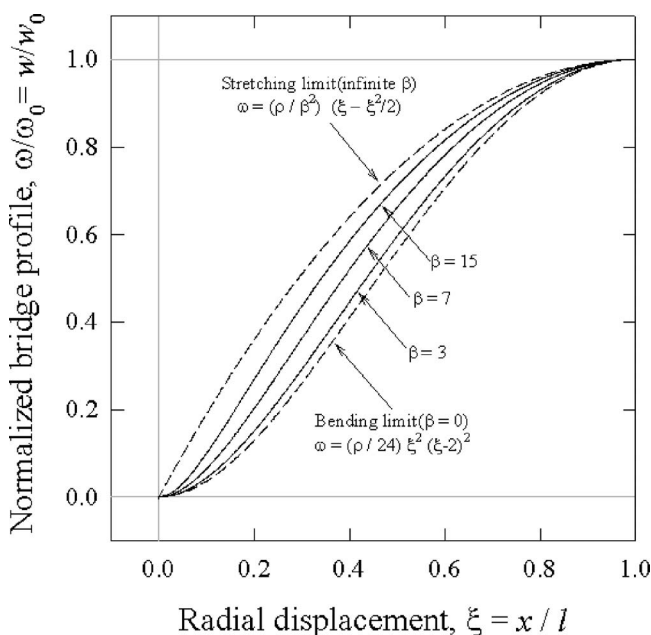


Fig. 2 Normalized bridge deformed profile as a function of membrane stress. The bridge anchors at $\xi=0$ and has its mid-span at $\xi=1$. The dashed curves show the plate-bending and membrane-stretching limits.

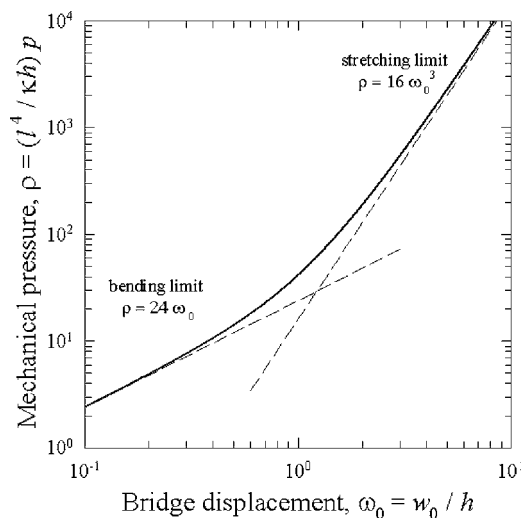


Fig. 3 Mechanical response of the bridge under a uniform pressure across the span. The dashed curves show the plate-bending and membrane-stretching limits.

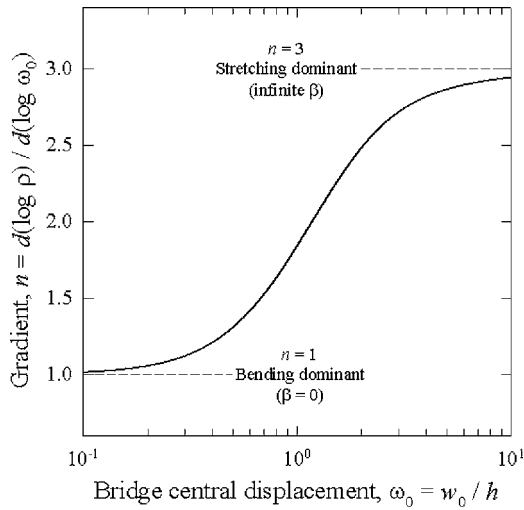


Fig. 4 The gradient $n(\omega_0)$ of the mechanical response $\rho(\omega_0)$

until they converge to C. Further increase in $v_0 > v_0^*$ (with the superscript asterisk denoting pull-in hereafter) leads to pull-in, i.e. spontaneous collapse of the bridge onto the electrode pad. The electromechanical force balance is maintained along path OAC.

The device behavior can be further scrutinized by a simple energy balance. The total energy of the system is given by $U_T = U_C - U_E$, where U_C and U_E are the energies stored in the capacitive dielectric medium at the bridge-pad gap and in the elastic bridge, respectively,

$$U_E = - \int p dV \quad \text{or} \quad \Sigma_E = - \int_0^\beta \rho(\beta) \frac{\partial \beta}{\partial \beta} d\beta \quad (11)$$

$$U_C = - \frac{\epsilon_0 V_0^2}{2} \int_0^{2\ell} \frac{dx}{g - w(x)} \quad \text{or} \quad \Sigma_C = - 2v_0^2 \int_0^1 \frac{d\xi}{\gamma - \omega(\beta, \xi)} \quad (12)$$

Figure 6 shows the energetics of the device. Since the bridge is bounded by the gap, $0 \leq w_0 \leq g$ or $0 \leq (\omega_0 / \gamma) \leq 1$. As v_0 increases

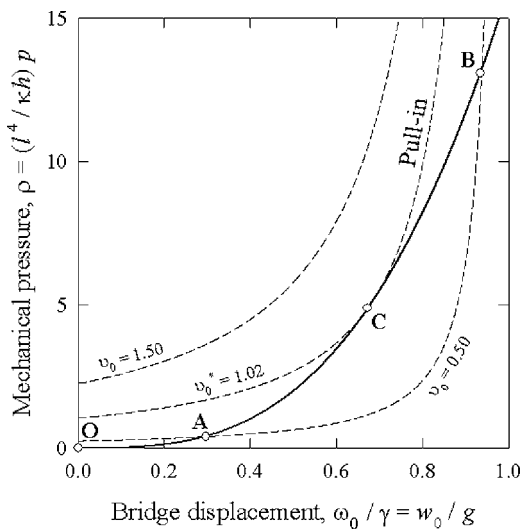


Fig. 5 Forces acting on the bridge in the stretching limit, with the attractive electrostatic force shown as dashed curves for a range of applied voltage, and the cubic mechanical force on the bridge shown as dark curve (OACB). Stable equilibrium is maintained along the path OAC. Pull-off occurs at C.

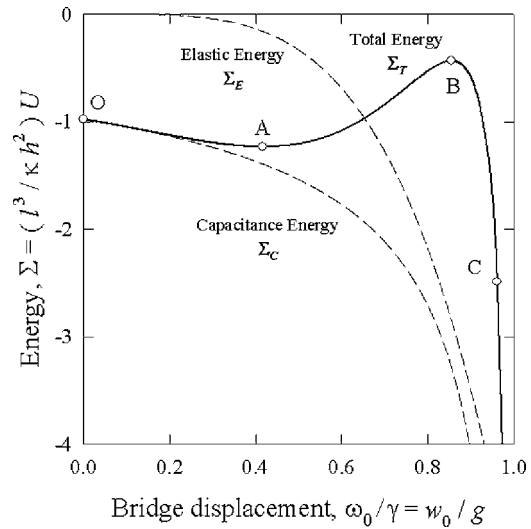


Fig. 6 Energetics of the MEMS-RF switch with $v_0=1.00$ in the stretching limit, showing various energy terms as functions of bridge central displacement

from null, the bridge deforms. Therefore, both Σ_E and Σ_C are monotonic decreasing in (ω_0 / γ) , and $\Sigma_T = \Sigma_C - \Sigma_E$ is shown as OABC. At a nonzero v_0 , the bridge moves to a stable equilibrium at A where Σ_T is minimal. An unstable equilibrium is found at B. Figures 7(a) and 7(b) shows $\Sigma_T(\omega_0)$ for a range of fixed v_0 and $\Sigma_T(v_0, \omega_0)$, respectively. As v_0 increases, $(\omega_0)_A$ and $(\omega_0)_B$ move close to each other and eventually merge at C, corresponding to a neutral equilibrium. Further increase in v_0 leads to pull-in. Energy balance is maintained along path OAA'C. The branch CB'B is obtained mathematically but is physically inaccessible. The stable equilibrium at A can be found by putting $(\partial \Sigma_T / \partial \omega_0) = 0$ and $[\partial^2 \Sigma_T / \partial (\omega_0)^2] < 0$. At pull-in at C, $[\partial^2 \Sigma_T / \partial (\omega_0)^2] = 0$, resulting in a point of inflexion. Exact solution for the pull-in parameters ω_0^* and v_0^* can be derived for the limiting cases of bending and stretching only, while the mixed bending-stretching behavior requires numerical integration.

The switch behavior depends significantly on the bridge-pad gap. Figure 8 shows (ω_0^* / γ) as a function of the bridge-pad gap γ . In theory, the force and energy balance should yield identical results. However, the cohesive zone approximation leads to a small inconsistency as shown in the shaded region, which cannot be resolved by the present model. Pull-in is expected anywhere within this zone. In fact, Fig. 5 shows $v_0^* = 1.02$ (force balance), whereas Fig. 7 shows $v_0^* = 0.92$ (energy balance) for the stretching limit. A bending-stretching transition occurs roughly at $g \approx 1.2h$, i.e., when the bridge-pad gap is slightly larger than the bridge thickness. A thick and stiff bridge combined with a small gap with $g < 0.5h$ leads to a bending dominant mode, while a thin (and flexible) bridge with a large gap with $g > 5h$ leads to a stretching dominant behavior. Bridges with the intermediate thickness and gap ($0.5h < g < 5h$) requires the full bending-stretching solution. Force balance requires $0.455 \leq (\omega_0^* / g) \leq 0.679$ and energy balance requires $0.397 \leq (\omega_0^* / g) \leq 0.658$, with the lower and upper limits referring to the pure bending and pure stretching modes, respectively.

Figure 9 shows the pull-in voltage v_0^* as a function of the bridge-pad gap. It can be easily shown that $v_0^* \propto \gamma^{3/2}$ in the plate-bending limit and $v_0^* \propto \gamma^{5/2}$ in the membrane-stretching limit. A small difference between force and energy balances is found as shown by the narrow shaded strip. In the bending limit, $v_0^* = 2.342 \gamma^{3/2}$ (force balance) and $v_0^* = 2.101 \gamma^{3/2}$ (energy balance). In the stretching limit, $v_0^* = 1.023 \gamma^{5/2}$ (force balance) and v_0^*

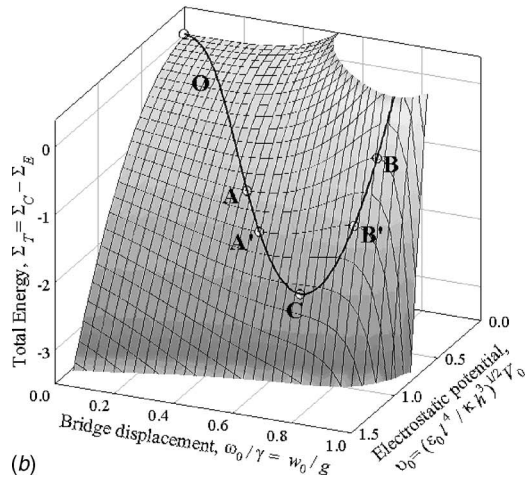
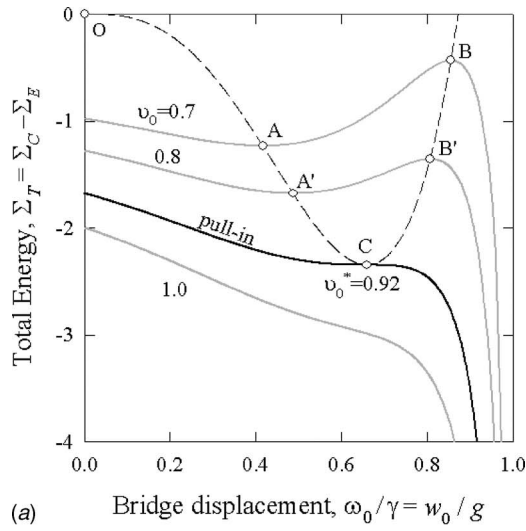


Fig. 7 (a) Total energy as a function of central bridge displacement for a range of applied voltage in the stretching limit. Stable equilibrium is maintained along the path OAA'C. Path CB'B is unstable and physically inaccessible. Pull-in occurs at C. (b) Total energy as a function of both central bridge displacement and applied voltage.

$=0.916 \gamma^{5/2}$ (energy balance). Bending-stretching transition occurs roughly at $g \approx 2.5h$ where the two limiting cases intersect.

2.2 2D Axisymmetric Switch. Figure 10 shows a 2D axisymmetric MEMS-RF-switch clamped at the circular perimeter. The set of normalized parameters is redefined as in Appendix B. The elastic deformation equation (1) remains valid, though the exact solution to the axisymmetric problems requires a nonlinear von Kármán equation in cylindrical coordinates to be solved. To avoid the mathematical complexity, an average stress approximation is adopted (i.e., $\sigma = \sigma_r = \sigma_t$) in association with the cohesive zone approximation. The boundary conditions are given by

(iii) At the clamped circumference:

$$\begin{aligned} w_{r=a} = 0 \quad \text{or} \quad \omega_{\xi=1} = 0 \\ (\partial w / \partial r)_{r=a} = 0 \quad \text{or} \quad (\partial \omega / \partial \xi)_{\xi=1} = 0 \end{aligned}$$

(iv) At the center:

$$(\partial w / \partial r)_{r=0} = 0 \quad \text{or} \quad (\partial \omega / \partial \xi)_{\xi=0} = 0$$

Equation (1) is reduced to the modified Bessel equation [26] with the profile gradient given by

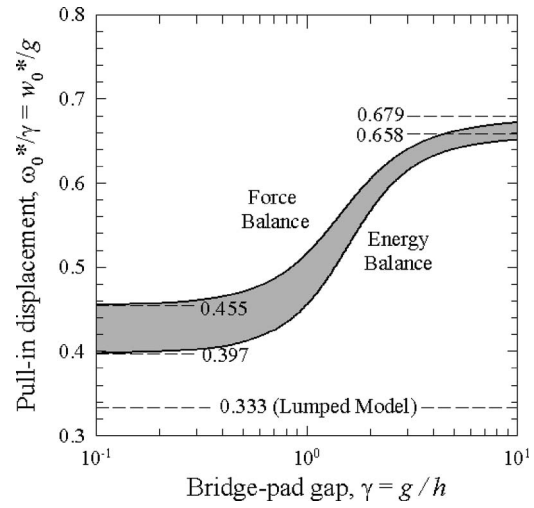


Fig. 8 Pull-in (w_0^*/g) as a function of the bridge-pad gap. Both force and energy balances are shown. Pull-in occurs within the shaded area.

$$\xi^2 \frac{\partial^2 \theta}{\partial \xi^2} + \xi \frac{\partial \theta}{\partial \xi} - (1 + \beta^2 \xi^2) \theta = \rho \xi^3 \quad (13)$$

with the apparent mechanical pressure on the film is given by

$$p = \frac{\epsilon_0 V_0^2}{2\pi a^2} \int_0^a \frac{2\pi r dr}{[g - w(r)]^2} \quad \text{or} \quad \rho = v_0^2 \int_0^1 \frac{\xi d\xi}{[\gamma - \omega(\xi)]^2} \quad (14)$$

Equation (13) can be solved analytically to yield the film profile

$$\omega = \rho \left(\frac{1}{\beta^3 I_1(\beta)} \right) \left\{ \frac{\beta}{2} (1 - \xi^2) I_1(\beta) + I_0(\beta \xi) - I_0(\beta) \right\} \quad (15)$$

with a central deflection

$$\omega_0 = \rho \left(\frac{1}{\beta^3 I_1(\beta)} \right) \left\{ \frac{\beta}{2} I_1(\beta) - I_0(\beta) + 1 \right\} \quad (16)$$

The corresponding average membrane stress is given by

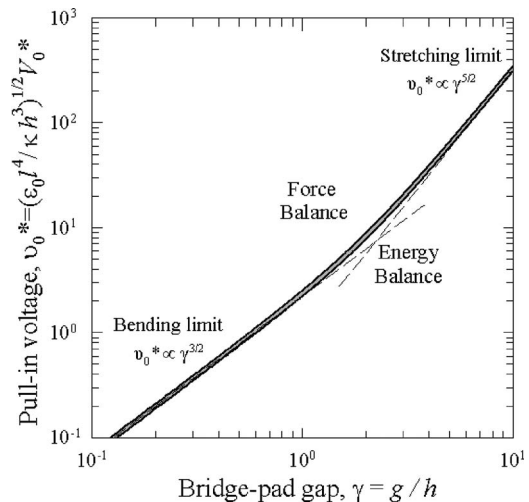


Fig. 9 Pull-in voltage as a function of the bridge-pad gap. Both force and energy balances are shown. Pull-in occurs within the shaded area. The dashed lines show the plate-bending and membrane-stretching limits.

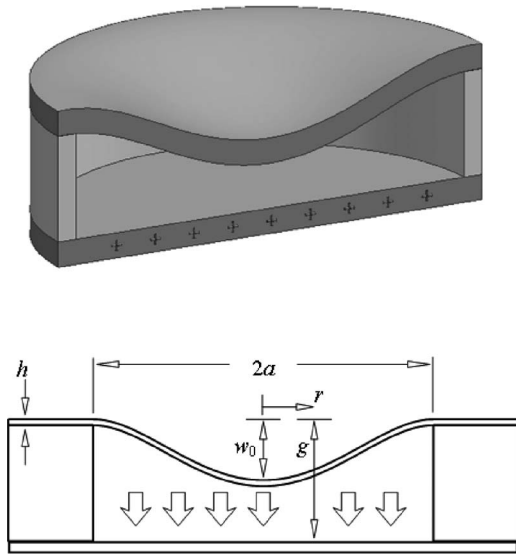


Fig. 10 Sketch of a 2D axisymmetric MEMS-RF switch

$$\sigma = \frac{1}{2a^2} \left(\frac{E}{1-\nu^2} \right) \int_0^a \left(\frac{dw}{dr} \right)^2 r dr \quad \text{or} \quad \beta^2 = 6 \int_0^1 \left(\frac{\partial \omega}{\partial \xi} \right)^2 \xi d\xi \quad (17)$$

which yields a relation between pressure and membrane stress

$$\rho = \frac{\beta^{7/2} I_1(\beta)}{\{(9\beta/2)I_1(\beta)^2 - 3I_2(\beta)[\beta I_0(\beta) + 4I_1(\beta)]\}^{1/2}} \quad (18)$$

The volume of the reduced dielectric space is found to be

$$V = \int_0^a w 2\pi r dr \quad \text{or} \quad \vartheta = \int_0^1 2\omega \xi d\xi = \rho \left(\frac{1}{\beta^4} \right) \left(2 + \frac{\beta^2}{4} - \frac{\beta I_0(\beta)}{I_1(\beta)} \right) \quad (19)$$

The mechanical response, $\rho(\omega_0)$ can be obtained by eliminating β from (16) and (18). The bending to stretching transition $\rho = k(\beta)\omega_0^{n(\beta)}$ is similar to the 1D counterpart with $1 \leq n \leq 3$. The limiting plate-bending solution becomes $\omega_{\text{bend}} = (\rho/32)(1-\xi^2)^2$, $\rho = 32\omega_0$, $n=1$, and $\vartheta = (1/3)\omega_0$. The limiting membrane-stretching solution becomes $\omega_{\text{stretch}} = (\rho/2\beta^2)(1-\xi^2)$, $\rho = 12\omega_0^3$, $n=3$, and $\vartheta = (1/2)\omega_0$.

The energetics and pull-in phenomenon for a 2D film is derived by the similar energy balance method as in the 1D model. Figures 11(a) and 11(b) show $\Sigma_T(\omega_0)$ for a range of fixed v_0 and $\Sigma_T(\omega_0, v_0)$, respectively. The trajectory OAA'C traces the energy balance locus, and pull-in occurs at C. Figure 12 shows (ω_0^*/γ) as a function of γ . A shaded region of uncertainty is found because of the discrepancies due to the average membrane stress approximation and the cohesive zone approximation. Figure 13 shows $v_0^*(\gamma)$. Similar to the 1D model, $v_0^* \propto \gamma^{3/2}$ is expected in the bending limit and $v_0^* \propto \gamma^{5/2}$ in the stretching limit. In summary, $v_0^* = 4.483 \gamma^{3/2}$ (force balance) and $v_0^* = 3.773 \gamma^{3/2}$ (energy balance) in the bending limit ($\gamma < 2$); $v_0^* = 1.591 \gamma^{5/2}$ (force balance) and $v_0^* = 1.338 \gamma^{5/2}$ (energy balance) in the stretching limit ($\gamma > 4$), and the transition at $\gamma \approx 3$.

3 Discussion

A solid mechanics model is derived for the electromechanical deformation of a bridge in a capacitive MEMS-RF switch and the associated pull-in phenomenon for both 1D and 2D. The analyti-

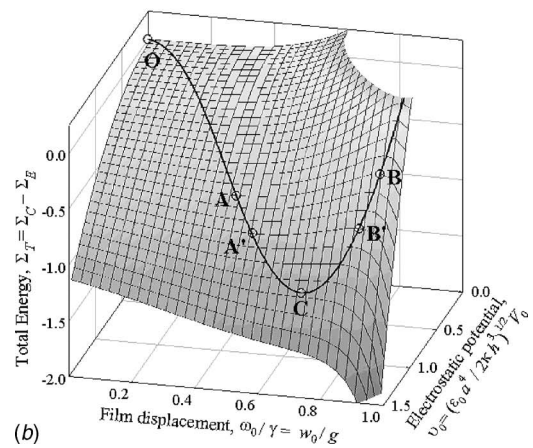
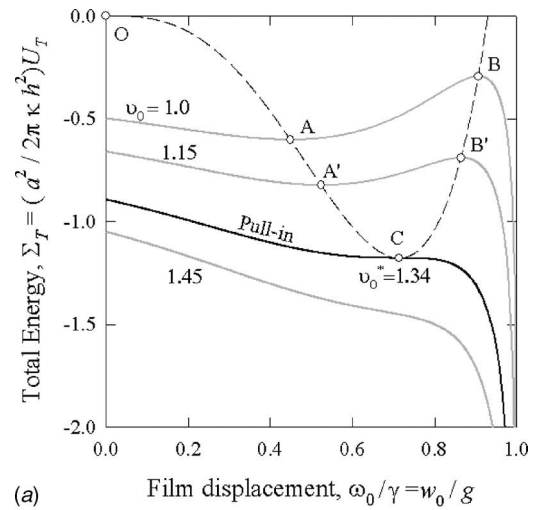


Fig. 11 (a) Total energy $\Sigma_T(\omega_0, v_0)$ for fixed v_0 in the stretching limit. Pull-in occurs at C. (b) Total energy $\Sigma_T(\omega_0, v_0)$.

cal solution has certain advantages over the existing models in formulating the design criteria. First, the combinatorial influences on the device are derived analytically [27] in terms of (i) materials parameters—elastic modulus, Poisson ratio, and flexural rigidity of bridge; (ii) geometrical parameters—bridge-pad gap separation,

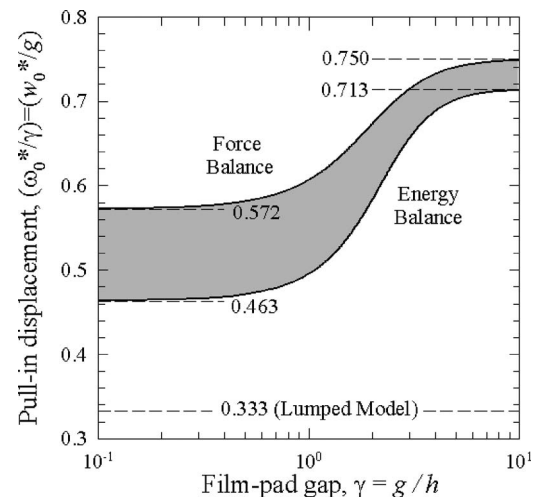


Fig. 12 Pull-in (ω_0^*/g) as a function of the bridge-pad gap

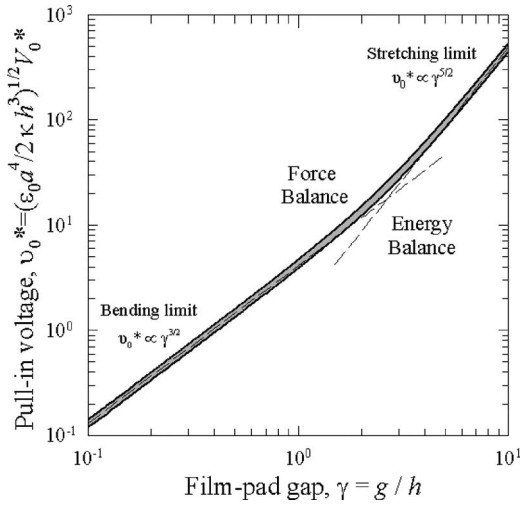


Fig. 13 Pull-in voltage as a function of the bridge-pad gap. Pull-in occurs within the shaded area.

bridge length and thickness; and (iii) structural index—mixed bending-stretching deformation, and the limiting cases of pure bending and pure stretching. Second, the ratio of gap to bridge thickness (g/h) is found to play a critical role in determining the pull-in voltage. The relations for a bending bridge ($v_0^* \propto \gamma^{3/2}$) and a stretching bridge ($v_0^* \propto \gamma^{5/2}$) are crucial in designing the device and assessing the performance, especially when the device dimensions shrink from micro- (MEMS) to nanoscale (NEMS). Note that the actual (g/h) ratio in most actual devices falls in the range of 0.5–5. The pull-in voltage in the plate-bending limit is consistent with literature [17,28], but the bending-stretching transition and the stretching limit are virtually unavailable in current literature. Table 1 compares the present work to various existing models. It is remarkable that the celebrated lumped model predicts the smallest (w_0^*/g)=1/3 and underestimates the critical pull-in voltage. Our new model essentially covers the entire range of literature values (besides the lumped model) and shows the bending-stretching transition to be the main cause of inconsistencies in the literature values. Most existing models do not allow bridge profile change (w/w_0) as the gap widens and are therefore incapable of predicting the bending-stretching in the electromechanical behavior. Third, when an AC voltage is applied to the electrode pad, the resonance frequency of the bridge is determined by the governing constitutive relation, $\rho \propto (\omega_0)^n$, with $n=1$ for thick and stiff bridge and $n=3$ for thin and flexible bridge. In the linear bending region, resonance can be investigated using the simple harmonic motion equation, but deviation is expected as the gap widens. For instance, the nonlinear van der Pol equation will be needed to solve for $n=3$ [29]. Failure to realize the bending-stretching transition in design will undermine the device performance.

The present model can be extended to include other important parameters not covered above. For instance, residual stress (σ_0) due to thermal mismatch is inevitable during device fabrication and operation. To accommodate its effect, the total membrane stress in Eq. (1) is rewritten as $\sigma = \sigma_0 + \sigma_m$, where σ_m is the concomitant stress due to change in bridge profile. Similarly, σ in Eq. (7) will be replaced by σ_m . The new constitutive relation and the subsequent pull-in parameters will yield useful information for switch design and can be checked against literature (e.g., [18,19]), though it is beyond the present scope. Another interesting extension is that of pull-off. When the electrostatic potential is turned off, the bridge adhered to the pad is expected to detach from the substrate and resume its undeformed geometry reversibly. However, in the presence of undesirable intersurface forces (e.g., capillary at high relative humidity, stray charges on surfaces), the

bridge must overcome the energy barrier in order to delaminate from the substrate. The thin-film delamination mechanics can be obtained using the present model. In fact, we have investigated the delamination mechanics of a clamped circular film earlier for an ideal zero-range surface force, and derived the critical mechanical force, bridge-pad gap, and radius at “pull-off” [30]. The model can be modified to allow transformation from 2D to 1D.

Another related area alluded to in the Introduction is the measurement of long-range intrinsic surface forces, such as van der Waals potential or stray charges left at the interface, etc. Such interactions can be incorporated into the present model by assigning an extra term on the right-hand side of (1) according to the Dugdale-Barenblatt-Maugis cohesive zone theory. In an earlier paper, we reported how a small graphite cylinder compelled a clamped silicone film into adhesive contact by means of a long-range surface force [31]. A solid mechanics model was also constructed to account for the subsequent delamination and pull-off. The present theoretical model here presents a thorough analysis for the pull-in event *prior* to the adhesion contact between the two adherends and is capable of analyzing the magnitude and range of surface forces involved. Since the intrinsic intersurface force potential is not tunable as in the MEMS switch but a fixed function, depending on the materials nature and the dielectric gap, one necessary modification to the present model is to allow the film-substrate gap to vary. Detailed analysis is beyond the scope of the present work.

4 Conclusion

Understanding the performance of a MEMS-RF switch in terms of the device geometry, materials, and structural index is crucial in the design criteria. In this study, a rigorous analytical elastic model is derived to account for the bridge deformed geometry and its effects on the pull-in voltage and other pull-in parameters. The ratio of bridge-pad gap to bridge thickness (g/h) is found to play a significant role in the device behavior.

Acknowledgment

This work is funded by National Science Foundation (Grant No. CMS-0527912) and University of Missouri Research Board (Grant No. 2428).

Appendix A: Normalized Parameters for the 1D Model

Coordinates and profile

$$\xi = \left(\frac{1}{\ell}\right)x, \quad \omega = \left(\frac{1}{h}\right)w, \quad \omega_0 = \left(\frac{1}{h}\right)w_0, \quad \vartheta = \left(\frac{1}{h\ell}\right)V$$

Device geometry (bridge-pad gap)

$$\gamma = \left(\frac{1}{h}\right)g$$

Electrostatic potential (applied voltage)

$$v_0 = \left(\frac{\epsilon_0 \ell^4}{2\kappa h^3}\right)^{1/2} V_0$$

Membrane stress

$$\beta = \left(\frac{\ell^2 h}{\kappa}\right)^{1/2} \sigma^{1/2}$$

Equivalent pressure

$$\rho = \left(\frac{\ell^4}{\kappa h}\right)p$$

Energies involved

$$\Sigma_E = \left(\frac{\ell^3}{\kappa h^2}\right)U_E, \quad \Sigma_C = \left(\frac{\ell^3}{\kappa h^2}\right)U_C, \quad \Sigma_T = \left(\frac{\ell^3}{\kappa h^2}\right)U_T$$

Appendix B: Normalized Parameters for the 2D Model

Coordinates and profile

$$\xi = \left(\frac{1}{a}\right)r, \quad \omega = \left(\frac{1}{h}\right)w, \quad \omega_0 = \left(\frac{1}{h}\right)w_0, \quad \vartheta = \left(\frac{1}{\pi a^2 h}\right)V$$

Device geometry (film-pad gap)

$$\gamma = \left(\frac{1}{h}\right)g$$

Electrostatic potential (applied voltage)

$$v_0 = \left(\frac{\epsilon_0 a^4}{2\kappa h^3}\right)^{1/2} V_0$$

Membrane stress

$$\beta = \left(\frac{a^2 h}{\kappa}\right)^{1/2} \sigma^{1/2}$$

Equivalent pressure

$$\rho = \left(\frac{a^4}{2\kappa h}\right)p$$

Energies involved

$$\Sigma_E = \left(\frac{a^2}{2\pi\kappa h^2}\right)U_E, \quad \Sigma_C = \left(\frac{a^2}{2\pi\kappa h^2}\right)U_C, \quad \Sigma_T = \left(\frac{a^2}{2\pi\kappa h^2}\right)U_T$$

References

- [1] Najjar, F., Choura, S., El-Borgi, S., Adbdel-Rahman, E., and Nayfeh, A., 2005, "Modeling and Design of Variable-Geometry Electrostatic Microactuators," *J. Micromech. Microeng.*, **15**, pp. 419–429.
- [2] Vogl, G., and Nayfeh, A., 2005, "A Reduced-Order Model for Electrically Actuated Clamped Circular Plates," *J. Micromech. Microeng.*, **15**, pp. 684–690.
- [3] Xu, T.-B., and Su, J., 2005, "Development, Characterization, and Theoretical Evaluation of Electroactive Polymer-Based Micropump Diaphragm," *Sens. Actuators, A*, **121**, pp. 267–274.
- [4] Prochaska, A. M., Nemirowsky, Y., and Dinnar, U., 2005, "A Membrane Micropump Electrostatically Actuated Across the Working Fluid," *J. Micromech. Microeng.*, **15**, pp. 2309–2316.
- [5] Ullmann, A., Fono, I., and Taitel, Y., 2001, "A Piezoelectric Valve-Less Pump—Dynamic Model," *ASME J. Fluids Eng.*, **123**, pp. 92–98.
- [6] Schomburg, W. K., Rummeler, Z., Shao, P., Walff, K., and Xie, L., 2004, "The Design of Metal Strain Gauges on Diaphragms," *J. Micromech. Microeng.*, **14**, pp. 1101–1108.
- [7] Elders, J., Spiering, V., and Walsh, S., 2001, "Microsystems Technology (MST) and MEMS Applications: An Overview," *Micromech. Microeng.*, **26**, pp. 312–317.
- [8] Rebeiz, G. M., and Muldavin, J. B., 2001, "RF MEMS Switches and Switch Circuits," *IEEE Microw. Mag.*, **2**, pp. 59–71.
- [9] Osterberg, P., and Senturia, S., 1997, "M-Test: A Test Chip for MEMS Material Property Measurement Using Electrostatically Actuated Test Structures," *J. Microelectromech. Syst.*, **6**, pp. 107–118.
- [10] Nemirowsky, Y., 2001, "A Methodology and Model for the Pull-in Parameters of Electrostatic Actuators," *J. Microelectromech. Syst.*, **10**, pp. 601–615.
- [11] Chowdhury, S., Ahmadi, M., and Miller, W., 2005, "A Closed-Form Model for the Pull-in Voltage of Electrostatically Actuated Cantilever Beams," *J. Micromech. Microeng.*, **15**, pp. 756–763.
- [12] Rocha, L. A., Cretu, E., and Wolffenbuttel, R. F., 2004, "Behavioural Analysis of the Pull-in Dynamic Transition," *J. Micromech. Microeng.*, **14**, pp. S37–S42.
- [13] Rottenberg, X., Brebels, S., De Raedt, W., Nauwelaers, B., and Tilmans, H. A. C., 2004, "RF-Power: Driver for Electrostatic RF-MEMS Devices," *J. Micromech. Microeng.*, **14**, pp. S43–S48.
- [14] Grettillat, M., Grettillat, F., and Rooij, N., 1999, "Micromechanical Relay With Electrostatic Actuation and Metallic Contacts," *J. Micromech. Microeng.*, **9**, pp. 324–331.
- [15] O'Mahony, C., Hill, M., Duane, R., and Mathewson, A., 2003, "Analysis of Electromechanical Boundary Effects on the Pull-in of Micromachined Fixed-Fixed Beams," *J. Micromech. Microeng.*, **13**, pp. S75–S80.
- [16] Lee, H., Coutu, Jr. R. A., Mall, S., and Kladitis, P. E., 2005, "Nanoindentation Technique for Characterizing Cantilever Beam Style RF Microelectromechanical Systems (MEMS) Switches," *J. Micromech. Microeng.*, **15**, pp. 1230–1235.
- [17] Rong, H., Huang, Q.-A., Nie, M., and Li, W., 2004, "An Analytical Model for Pull-in Voltage of Clamped-Clamped Multilayer Beams," *Sens. Actuators, A*, **116**, pp. 15–21.
- [18] Schauwecker, B., Mehner, J., Strohm, K. M., Haspeklo, H., and Luy, J.-F., 2004, "Investigations of RF Shunt Airbridge Switches Among Different Environmental Conditions," *Sens. Actuators, A*, **114**, pp. 49–58.
- [19] Zhang, Y., and Zhao, Y.-p., 2006, "Numerical and Analytical Study on the Pull-in Instability of Micro-Structure Under Electrostatic Loading," *Sens. Actuators, A*, **127**, pp. 366–380.
- [20] Gould, P. L., 1988, *Analysis of Shells and Plates*, Springer, New York.
- [21] Gupta, R., 1997, "Electrostatic Pull-in Test Structure Design for In Situ Mechanical Property Measurements of Micro-Electro-Mechanical Systems (MEMS)," Ph.D. thesis, MIT, Cambridge, MA.
- [22] Maugis, D., 2000, *Contact, Adhesion and Rupture of Elastic Solids*, Springer, New York.
- [23] Wan, K.-T., and Duan, J., 2002, "Adherence of a Rectangular Flat Punch Onto a Clamped Plate—Transition From a Rigid Plate to a Flexible Membrane," *ASME J. Appl. Mech.*, **69**, pp. 104–109.
- [24] Timoshenko, S. P., and Woinowsky-Krieger, S., 1959, *Theory of Plates and Shells*, 2nd ed., McGraw-Hill, New York.
- [25] Wan, K.-T., 1999, "Fracture Mechanics of a V-Peel Adhesion Test—Transition From a Bending Plate to a Stretching Membrane," *J. Adhes.*, **70**, pp. 197–207.
- [26] Wan, K.-T., 2002, "Adherence of an Axisymmetric Flat Punch Onto a Clamped Circular Plate—Transition From a Rigid Plate to a Flexible Membrane," *ASME J. Appl. Mech.*, **69**, pp. 110–116.
- [27] Ashby, M. F., 2005, *Materials Selection in Mechanical Design*, 3rd ed., Butterworth-Heinemann, Woburn, MA.
- [28] Begley, M. R., 2005, "The Impact of Materials Selection and Geometry on Multi-Functional Bilayer Micro-Sensors and Actuators," *J. Micromech. Microeng.*, **15**, pp. 2379–2388.
- [29] Stoker, J., 1992, *Nonlinear Vibrations in Mechanical and Electrical Systems*, Wiley, New York.
- [30] Wan, K.-T., and Kogut, L., 2005, "The Coupling Effect of Interfacial Adhesion and Tensile Residual Stress on a Thin Membrane Adhered to a Flat Punch," *J. Micromech. Microeng.*, **15**, pp. 778–784.
- [31] Ju, B. F., Wan, K.-T., and Liu, K. K., 2004, "Indentation of a Square Elastomeric Thin Film by a Flat-Ended Cylindrical Punch in the Presence of Long-Range Intersurface Forces," *J. Appl. Phys.*, **96**, pp. 6159–6163.

Parametric Formulation of the Finite-Volume Theory for Functionally Graded Materials—Part I: Analysis

Marcio A. A. Cavalcante

Severino P. C. Marques

Center of Technology,
Federal University of Alagoas,
Maceio, Alagoas, Brazil

Marek-Jerzy Pindera

Civil Engineering Department,
University of Virginia,
Charlottesville, VA 22903

The recently reconstructed higher-order theory for functionally graded materials is further enhanced by incorporating arbitrary quadrilateral subcell analysis capability through a parametric formulation. This capability significantly improves the efficiency of modeling continuous inclusions with arbitrarily-shaped cross sections of a graded material's microstructure previously approximated using discretization based on rectangular subcells, as well as modeling of structural components with curved boundaries. Part I of this paper describes the development of the local conductivity and stiffness matrices for a quadrilateral subcell which are then assembled into global matrices in an efficient manner following the finite-element assembly procedure. Part II verifies the parametric formulation through comparison with analytical solutions for homogeneous curved structural components and graded components where grading is modeled using piecewise uniform thermoelastic moduli assigned to each discretized region. Results for a heterogeneous microstructure in the form of a single inclusion embedded in a matrix phase are also generated and compared with the exact analytical solution, as well as with the results obtained using the original reconstructed theory based on rectangular discretization and finite-element analysis. [DOI: 10.1115/1.2722312]

Keywords: functionally graded materials, finite-volume theory, parametric formulation

1 Introduction

Functionally graded materials (FGMs) are multiphase materials with engineered microstructures which produce property gradients aimed at optimizing structural response under different types of loads (thermal, mechanical, electrical, optical, etc.). These property gradients are produced in several ways, for example by gradual variation of the content of one phase (ceramic) relative to other (metallic) used in thermal barrier coatings, or by using a sufficiently large number of constituent phases with different properties. Developed by Japanese researchers in the mid-1980s, these materials continue to evolve and to find new applications in areas other than the thermal protection/management structures for which they had been originally developed, cf. Suresh and Mortensen [1], Miyamoto et al. [2], Paulino [3], Chatzigeorgiou and Charalambakis [4].

The use of graded material concepts in structural design and optimization requires the development of appropriate analysis techniques which account for the spatially variable microstructures in this class of materials. Presently, there are two approaches available to analyze the response of FGMs to thermomechanical loads, called coupled and uncoupled approaches, Pindera et al. [5]. In the uncoupled approach, the graded material's microstructure is replaced by equivalent homogenized properties which are either determined from micromechanics considerations or assumed a priori. This results in a boundary-value problem with either continuously or discretely variable elastic moduli at the scale at which the analysis is conducted, called macroscale. In the coupled approach originally proposed by Aboudi et al. [6], and summarized in a review paper by Aboudi et al. [7], the material's microstructure is explicitly taken into account by performing the

analysis at the microscale. In particular, in the original formulation of this so-called higher-order theory, a two-step discretization involving generic cells and subcells is employed to capture the graded material's heterogeneous microstructure. Subsequently, thermal and displacement fields within each subcell are approximated using quadratic expansions in local coordinates, and the unknown coefficients associated with the different-order terms are obtained by satisfying various moments of the field equations in a volume-averaged sense in each subcell, followed by the application of continuity conditions within each generic cell, and between adjacent cells, in a surface-average sense together with the imposed boundary conditions. We mention that surface averaging of continuity conditions was proposed by Achenbach [8] in the context of the author's cell model for unidirectional composites.

This approach has recently been reconstructed by Bansal and Pindera [9] and Zhong et al. [10] based on a simplified volume discretization using subcells as the fundamental subvolumes, in place of the two-level discretization employed in the original construction. The use of subcells as the fundamental subvolumes, in turn, facilitated the implementation of the local/global stiffness matrix formulation, Bufler [11], Pindera [12], into the solution procedure for the unknown subcell surface-averaged interfacial displacements which became the primary unknown quantities in the reconstructed theory. The reconstruction has also revealed that the model's theoretical framework is based on direct satisfaction of the field equations within each subcell, in contrast to the original construction wherein higher-order moments of the equilibrium equations were also satisfied, thereby erroneously suggesting this model to be a version of a micropolar continuum theory. The significantly simplified theoretical structure of this so-called higher-order theory in conjunction with the implementation of the local/global stiffness matrix approach also resulted in a substantial reduction in the final system of equations for the unknown quantities, thereby making it possible to analyze realistic graded microstructures that required extensive discretization not possible

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received June 6, 2006; final manuscript received December 22, 2006. Review conducted by Robert M. McMeeking.

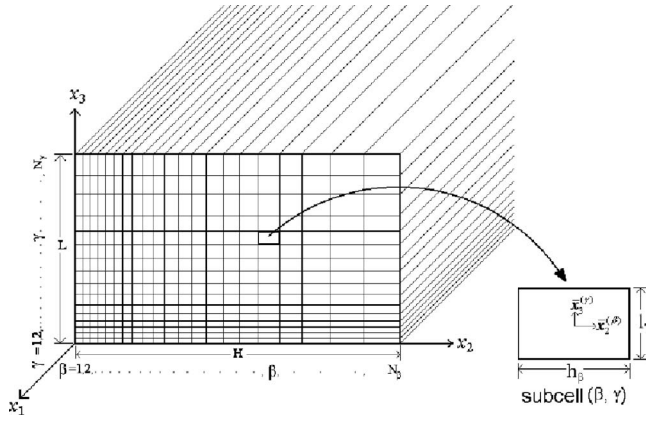


Fig. 1 Simplified discretization of a graded microstructure (left) into rectangular subcells with the local coordinate system $x_2^{(\beta)} - x_3^{(\gamma)}$ (right) used in the reconstructed finite-volume theory

with the theory's original formulation. Equally important, the reconstruction has revealed the method to be a finite-volume direct averaging technique with clearly discernible similarities and differences between the finite-volume technique used in fluid dynamics problems, cf. Versteeg and Malalasekera [13], and the finite-element method, as reported by Pindera et al. [14]. This reconstruction made it possible to further develop the method in order to increase its efficiency and applicability.

In this communication, we continue the development of the reconstructed Cartesian-version of the theory, herein referred to as the standard finite-volume theory, by incorporating parametric mapping capability in order to enable efficient modeling of microstructures that cannot be easily modeled using rectangular subcells employed in the standard theory. The parametric formulation offers the same flexibility as the finite-element method for modeling microstructures and geometries with curved and rectilinear features, while retaining its quick convergence characteristics in the presence of highly heterogeneous microstructures. We begin by providing a brief outline of the reconstructed finite-volume theory in order to set the stage for the parametric formulation.

2 The Finite-Volume Theory for Functionally Graded Materials

In the standard version of the finite-volume theory for FGMs, the material microstructure is discretized into N_β columns, each of width h_β , spanning the distance H along the x_2 -axis, and N_γ rows, each of height l_γ , spanning the distance L along the x_3 -axis, Fig. 1. This results in a grid of $N_\beta \times N_\gamma$ (β, γ) subcells of $h_\beta \times l_\gamma$ dimensions, which is used to approximate the heterogeneous microstructure by assigning appropriate moduli to each subcell. The temperature and displacement fields are approximated in each subcell using quadratic expansions in the local subcell coordinates $(\bar{x}_2^{(\beta)}, \bar{x}_3^{(\gamma)})$ attached to the subcell's center. For the thermal problem we have

$$T^{(\beta, \gamma)} = T_{(00)}^{(\beta, \gamma)} + \bar{x}_2^{(\beta)} T_{(10)}^{(\beta, \gamma)} + \bar{x}_3^{(\gamma)} T_{(01)}^{(\beta, \gamma)} + \frac{1}{2} \left(3\bar{x}_2^{(\beta)2} - \frac{h_\beta^2}{4} \right) T_{(20)}^{(\beta, \gamma)} + \frac{1}{2} \left(3\bar{x}_3^{(\gamma)2} - \frac{l_\gamma^2}{4} \right) T_{(02)}^{(\beta, \gamma)} \quad (1)$$

The heat flux components $q_i^{(\beta, \gamma)}$ at any point passing through a subcell (β, γ) are then obtained from the Fourier's law of heat conduction

$$q_i^{(\beta, \gamma)} = -k_i^{(\beta, \gamma)} \frac{\partial T^{(\beta, \gamma)}}{\partial \bar{x}_i^{(\cdot)}} \quad (i = 2, 3; \text{no sum}) \quad (2)$$

where $k_i^{(\beta, \gamma)}$ are the heat conductivity coefficients of the material in the subcell (β, γ) . For the mechanical problem

$$u_i^{(\beta, \gamma)} = W_{i(00)}^{(\beta, \gamma)} + \bar{x}_2^{(\beta)} W_{i(10)}^{(\beta, \gamma)} + \bar{x}_3^{(\gamma)} W_{i(01)}^{(\beta, \gamma)} + \frac{1}{2} \left(3\bar{x}_2^{(\beta)2} - \frac{h_\beta^2}{4} \right) W_{i(20)}^{(\beta, \gamma)} + \frac{1}{2} \left(3\bar{x}_3^{(\gamma)2} - \frac{l_\gamma^2}{4} \right) W_{i(02)}^{(\beta, \gamma)} \quad (3)$$

where $i=2, 3$ for plane problems. The stress components at any point within the (β, γ) subcell are obtained, upon the use of strain-displacement equations, from the generalized Hooke's law

$$\sigma_{ij}^{(\beta, \gamma)} = C_{ijkl}^{(\beta, \gamma)} \epsilon_{kl}^{(\beta, \gamma)} - \sigma_{ij}^{th(\beta, \gamma)} \quad (4)$$

where $\sigma_{ij}^{th(\beta, \gamma)} = C_{ijkl}^{(\beta, \gamma)} \alpha_{kl}^{(\beta, \gamma)} \Delta T = \Gamma_{ij}^{(\beta, \gamma)} \Delta T$.

Subsequently, conductivity and stiffness matrices are constructed for each subcell by relating the surface-averaged temperature and displacement vectors, $\bar{\mathbf{T}}^{(\beta, \gamma)}$ and $\bar{\mathbf{u}}^{(\beta, \gamma)}$, to the corresponding surface-averaged heat flux and traction vectors, $\bar{\mathbf{Q}}^{(\beta, \gamma)}$ and $\bar{\mathbf{t}}^{(\beta, \gamma)}$, whose components are

$$\begin{aligned} \bar{\mathbf{T}}^{(\beta, \gamma)} &= [\bar{T}^{2+}, \bar{T}^{2-}, \bar{T}^{3+}, \bar{T}^{3-}]^{(\beta, \gamma)T} \\ \bar{\mathbf{Q}}^{(\beta, \gamma)} &= [\bar{Q}_2^{2+}, \bar{Q}_2^{2-}, \bar{Q}_3^{3+}, \bar{Q}_3^{3-}]^{(\beta, \gamma)T} \\ \bar{\mathbf{u}}^{(\beta, \gamma)} &= [\bar{u}_2^{2+}, \bar{u}_3^{2+}, \bar{u}_2^{2-}, \bar{u}_3^{2-}, \bar{u}_2^{3+}, \bar{u}_3^{3+}, \bar{u}_2^{3-}, \bar{u}_3^{3-}]^{(\beta, \gamma)T} \\ \bar{\mathbf{t}}^{(\beta, \gamma)} &= [\bar{t}_2^{2+}, \bar{t}_3^{2+}, \bar{t}_2^{2-}, \bar{t}_3^{2-}, \bar{t}_2^{3+}, \bar{t}_3^{3+}, \bar{t}_2^{3-}, \bar{t}_3^{3-}]^{(\beta, \gamma)T} \end{aligned} \quad (5)$$

where the superscripts 2, 3 identify the direction of the unit normal to the given face relative to the fixed subcell coordinates $(\bar{x}_2^{(\beta)}, \bar{x}_3^{(\gamma)})$, and $\pm$ denote the actual sense. The subscripts 2, 3 denote a vectorial quantity's component. Traction components are expressed in terms of stress components associated with a particular subcell face through Cauchy's relations

$$t_i^{\mathbf{n}^{(\beta, \gamma)}} = \sigma_{ji}^{(\beta, \gamma)} n_j^{(\beta, \gamma)} \quad (6)$$

and $\mathbf{n}^{(\beta, \gamma)}$ is the unit normal to a given face of the (β, γ) subcell. Finally, the surface averages are defined in the standard way, as for example

$$\begin{aligned} \bar{T}^{2\pm(\beta, \gamma)} &= \frac{1}{l_\gamma} \int_{-l_\gamma/2}^{l_\gamma/2} T^{(\beta, \gamma)} \left(\pm \frac{h_\beta}{2}, \bar{x}_3^{(\gamma)} \right) d\bar{x}_3^{(\gamma)} \\ \bar{T}^{3\pm(\beta, \gamma)} &= \frac{1}{h_\beta} \int_{-h_\beta/2}^{h_\beta/2} T^{(\beta, \gamma)} \left(\bar{x}_2^{(\beta)}, \pm \frac{l_\gamma}{2} \right) d\bar{x}_2^{(\beta)} \end{aligned} \quad (7)$$

for the surface-averaged temperatures, with similar expressions for the remaining field variables.

For the thermal problem, there are five coefficients which must be related to the four surface-averaged temperatures. Four relations are provided by the definitions of the surface-averaged temperatures. Satisfaction of the heat conduction equation in the large

$$\int_{S^{(\beta, \gamma)}} q_i^{(\beta, \gamma)} n_i^{(\beta, \gamma)} dS = 0 \quad (8)$$

provides the additional relation required in the local conductivity matrix construction carried out by averaging the heat flux equations along each of the four faces, which yields

$$\bar{\mathbf{Q}}^{(\beta, \gamma)} = \boldsymbol{\kappa}^{(\beta, \gamma)} \bar{\mathbf{T}}^{(\beta, \gamma)} \quad (9)$$

For the mechanical problem, there are ten coefficients which must be related to the eight surface-averaged displacement components. Eight relations are provided by the definitions of the

surface-averaged displacement components. The additional two equations are obtained by satisfying the stress equilibrium equations in each subcell in the large

$$\int_{S^{(\beta,\gamma)}} t_i^{n^{(\beta,\gamma)}} dS = 0, \quad i = 2, 3 \quad (10)$$

Evaluating the surface-averaged traction components on each face of the (β, γ) subcell leads to the local stiffness matrix relation of the form

$$\bar{\mathbf{t}}^{(\beta,\gamma)} = \mathbf{K}^{(\beta,\gamma)} \bar{\mathbf{u}}^{(\beta,\gamma)} + \mathbf{\Gamma}^{(\beta,\gamma)} \bar{\mathbf{T}}^{(\beta,\gamma)} \quad (11)$$

Assembly of the local conductivity and stiffness matrices by enforcing the continuity of both surface-averaged temperatures and displacements, and heat fluxes and tractions, together with the specified boundary conditions, determines the unknown surface-averaged temperatures and displacements from which local fields are obtained through constitutive and gradient relations. In this approach, the redundant temperature and displacement continuity equations are eliminated by setting the surface-averaged temperatures and displacements at the interfaces associated with the adjacent subcells (β, γ) , $(\beta+1, \gamma)$, and (β, γ) , $(\beta, \gamma+1)$ to common unknowns,

$$\begin{aligned} \bar{T}^{2+(\beta,\gamma)} = \bar{T}^{2-(\beta+1,\gamma)} = \bar{T}^{2+(\beta+1,\gamma)} \quad \bar{T}^{3+(\beta,\gamma)} = \bar{T}^{3-(\beta,\gamma+1)} = \bar{T}^{3+(\beta,\gamma+1)} \\ \bar{u}_i^{2+(\beta,\gamma)} = \bar{u}_i^{2-(\beta+1,\gamma)} = \bar{u}_i^{2+(\beta+1,\gamma)} \quad \bar{u}_i^{3+(\beta,\gamma)} = \bar{u}_i^{3-(\beta,\gamma+1)} = \bar{u}_i^{3+(\beta,\gamma+1)} \\ i = 2, 3 \end{aligned} \quad (12)$$

upon application of the heat flux and traction continuity conditions at these common interfaces

$$\begin{aligned} \bar{Q}_2^{+(\beta,\gamma)} - \bar{Q}_2^{-(\beta+1,\gamma)} = 0 \quad \bar{Q}_3^{+(\beta,\gamma)} - \bar{Q}_3^{-(\beta,\gamma+1)} = 0 \\ \bar{t}_i^{2+(\beta,\gamma)} + \bar{t}_i^{2-(\beta+1,\gamma)} = 0 \quad \bar{t}_i^{3+(\beta,\gamma)} + \bar{t}_i^{3-(\beta,\gamma+1)} = 0 \quad i = 2, 3 \end{aligned} \quad (13)$$

The resulting global thermal conductivity matrix relates the unknown interfacial surface-averaged temperatures (including those at the external boundaries) represented by $\bar{\mathbf{T}}$ to the corresponding surface-averaged heat fluxes represented by $\bar{\mathbf{Q}}$

$$\kappa \bar{\mathbf{T}} = \bar{\mathbf{Q}} \quad (14)$$

The vector $\bar{\mathbf{Q}}$ consists mainly of zeros which are obtained after applying interfacial heat flux continuity conditions across each interface separating adjacent subcells, with the nonzero terms representing surface-averaged heat fluxes along each portion of the discretized boundary. The above system of equations is modified accordingly when the boundary conditions are specified in terms of applied temperatures. Similarly, the resulting global stiffness matrix relates the unknown interfacial surface-averaged displacements (including those at the external boundaries) to the surface-averaged tractions prescribed at the external boundaries

$$\mathbf{K} \bar{\mathbf{U}} = \bar{\mathbf{t}} + \mathbf{\Gamma} \bar{\mathbf{T}} \quad (15)$$

with the known surface-averaged temperatures obtained from the solution of the thermal problem. The above system of equations is then reduced to eliminate rigid body motion, and modified according to the specified boundary conditions.

3 Parametric Formulation of the Finite-Volume Theory

The parametric formulation enables the use of quadrilateral subcells in approximating the heterogeneous microstructure of a graded material, thereby making the modeling of microstructural details with curvilinear boundaries more efficient, as illustrated in Fig. 2 for a single inclusion in a matrix phase approximated by

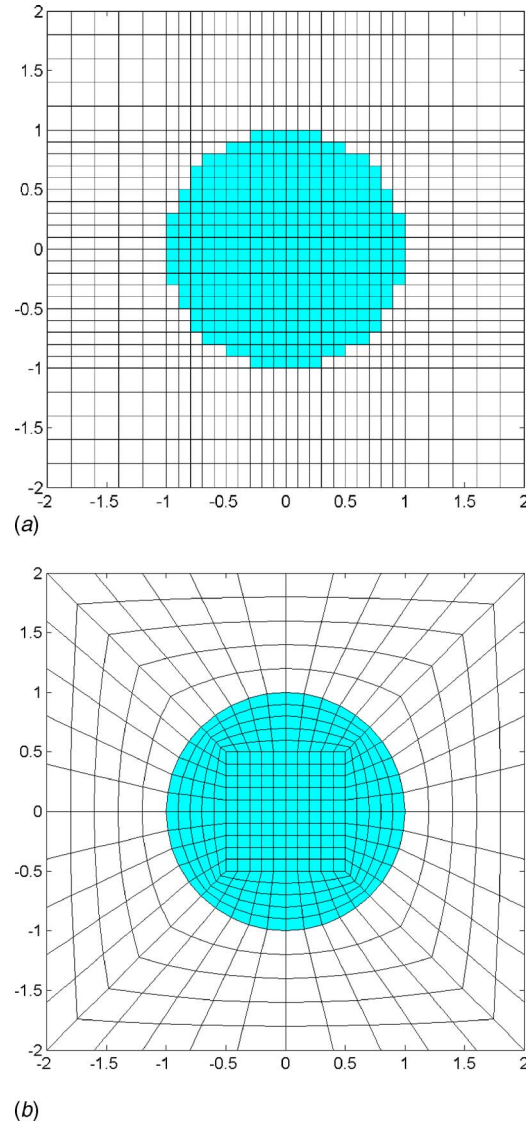


Fig. 2 Discretization of a square region containing a circular inclusion: (a) discretization based on 900 rectangular subcells; (b) discretization based on 500 quadrilateral subcells

rectangular and quadrilateral subcells. It is based on a mapping of a reference square subcell in the η - ξ plane onto a subcell in the x - y plane of the actual microstructure. The mapping facilitates the development of local conductivity and stiffness matrices for a quadrilateral subcell situated in the actual microstructure that relate surface-averaged temperatures and displacements to the corresponding heat fluxes and tractions acting on arbitrarily oriented rectilinear surfaces.

The important distinction between the parametric formulation and the version based on rectangular subcells described in Sec. 2 is the absence of a local coordinate system associated with each subcell in the actual microstructure. Rather, global coordinates are employed to describe the locations of quadrilateral subcell vertices in the actual microstructure, and thus the subcell's placement. Global reference indexes are also assigned to the four faces of each subcell in the actual microstructure which are employed in the construction of the connectivity matrix used to apply interfacial continuity and balance conditions in the assembly of the global conductivity and stiffness matrices. These indexes define the location of the local conductivity and stiffness matrix elements in the global systems of equations. We begin by describing the co-

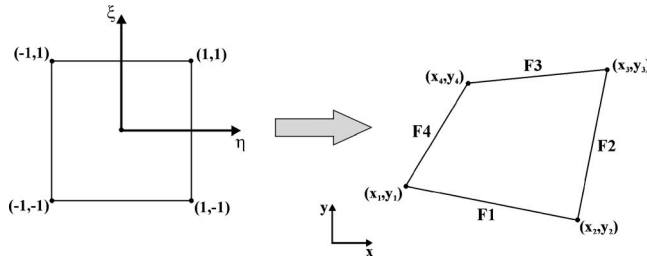


Fig. 3 Mapping of the reference subcell in the η - ξ plane onto a quadrilateral subcell in the x - y plane of the actual microstructure.

ordinate transformations associated with the mapping, and then outline transient and mechanical analyses based on this parametric mapping technique in the following sections.

In the parametric formulation, the reference subcell is a square in the η - ξ plane bounded by $-1 \leq \xi \leq 1$ and $-1 \leq \eta \leq 1$. As shown in Fig. 3, the vertices $(-1, -1)$ of the reference subcell correspond to the vertices (x_1, y_1) of the j th subcell in the actual discretized microstructure. The vertices are numbered such that the first set (x_1, y_1) is at the lower left corner and the numbering convention increases in a counterclockwise fashion. The faces are numbered similarly such that the face F_i lies between the vertices (x_i, y_i) and (x_{i+1}, y_{i+1}) with $i+1 \rightarrow 1$ when $i=4$. Thus the components of the unit normal vector $\mathbf{n}^{(i)} = [n_x^{(i)}, n_y^{(i)}]$ to the face F_i are given by

$$n_x^{(i)} = \frac{y_{i+1} - y_i}{L_i} \quad n_y^{(i)} = \frac{x_i - x_{i+1}}{L_i}$$

where

$$L_i = \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2} \quad (16)$$

The mapping of the point (η, ξ) in the reference subcell to the corresponding point (x, y) in the subcell of the actual discretized microstructure is given by, Cavalcante [15],

$$x(\eta, \xi) = N_1(\eta, \xi)x_1 + N_2(\eta, \xi)x_2 + N_3(\eta, \xi)x_3 + N_4(\eta, \xi)x_4$$

$$y(\eta, \xi) = N_1(\eta, \xi)y_1 + N_2(\eta, \xi)y_2 + N_3(\eta, \xi)y_3 + N_4(\eta, \xi)y_4 \quad (17)$$

where

$$N_1(\eta, \xi) = \frac{1}{4}(1 - \eta)(1 - \xi) \quad N_2(\eta, \xi) = \frac{1}{4}(1 + \eta)(1 - \xi)$$

$$N_3(\eta, \xi) = \frac{1}{4}(1 + \eta)(1 + \xi) \quad N_4(\eta, \xi) = \frac{1}{4}(1 - \eta)(1 + \xi) \quad (18)$$

In both the thermal and mechanical problems, the field variables are approximated by the same functional form based on a second-order representation in the local coordinates of the reference subcell in the η - ξ plane, given by

$$f(\eta, \xi) = F_{(00)} + \eta F_{(10)} + \xi F_{(01)} + \frac{1}{2}(3\eta^2 - 1)F_{(20)} + \frac{1}{2}(3\xi^2 - 1)F_{(02)} \quad (19)$$

For the thermal problem, $f(\eta, \xi)$ represents the local temperature field, whereas in the mechanical problem the above expression becomes a vector function representing the two displacement components in the η - ξ plane with two sets of coefficients associated with each term in the expansion. Further, both sets of field variables are governed by similar differential equations and constitutive equations relating flux-like quantities to the partial derivatives of the field variables. Therefore, in this section we develop

unifying relations containing common transformation matrices between surface-averaged quantities in the two planes in order to avoid duplication in the subsequent sections dealing with the thermal and mechanical problems.

First, we evaluate surface-averaged functions on each face of the j th subcell in the actual microstructure by performing integrations along the corresponding sides $\eta = \pm 1$ and $\xi = \pm 1$ in the η - ξ plane given the mapping in Eqs. (17) and (18). Denoting the surface-averaged functions associated with each face F_i referred to the global coordinates by $\bar{f}_i$, we obtain

$$\bar{f}_{1,3} = \frac{1}{2} \int_{-1}^{+1} f(\eta, \xi = \mp 1) d\eta = F_{(00)} \mp F_{(01)} + F_{(02)}$$

$$\bar{f}_{2,4} = \frac{1}{2} \int_{-1}^{+1} f(\eta = \pm 1, \xi) d\xi = F_{(00)} \pm F_{(10)} + F_{(20)} \quad (20)$$

where superscripts denoting the j th subcell have been suppressed for clarity of notation. Solving for the first and second order coefficients $F_{(10)}, \dots, F_{(02)}$ in terms of the surface-averaged functions and the zeroth order coefficient $F_{(00)}$ we have

$$\begin{bmatrix} F_{(10)} \\ F_{(01)} \\ F_{(20)} \\ F_{(02)} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 0 & 1 & 0 & -1 \\ -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \bar{f}_1 - F_{(00)} \\ \bar{f}_2 - F_{(00)} \\ \bar{f}_3 - F_{(00)} \\ \bar{f}_4 - F_{(00)} \end{bmatrix} \quad (21)$$

Derivation of the relations between surface-averaged flux-like variables and $\bar{f}_1, \dots, \bar{f}_4$ requires the relationship between first partial derivatives of the function $f(\cdot, \cdot)$ in the two planes η - ξ and x - y . These two sets of partial derivatives are related through the Jacobian $\mathbf{J}$ and its inverse $\mathbf{J}^{-1}$

$$\begin{bmatrix} \frac{\partial f}{\partial \eta} \\ \frac{\partial f}{\partial \xi} \end{bmatrix} = \mathbf{J} \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} = \mathbf{J}^{-1} \begin{bmatrix} \frac{\partial f}{\partial \eta} \\ \frac{\partial f}{\partial \xi} \end{bmatrix} \quad (22)$$

where $\mathbf{J}$ is obtained from the transformation equations, Eqs. (17) and (18), in the form

$$\mathbf{J} = \begin{bmatrix} \frac{\partial x}{\partial \eta} & \frac{\partial y}{\partial \eta} \\ \frac{\partial x}{\partial \xi} & \frac{\partial y}{\partial \xi} \end{bmatrix} = \begin{bmatrix} A_1 + A_2\xi & A_4 + A_5\xi \\ A_3 + A_2\eta & A_6 + A_5\eta \end{bmatrix} \quad (23)$$

with $A_1, \dots, A_6$ given below in terms of the vertex coordinates (x_i, y_i) ,

$$A_1 = \frac{1}{4}(-x_1 + x_2 + x_3 - x_4) \quad A_2 = \frac{1}{4}(x_1 - x_2 + x_3 - x_4)$$

$$A_3 = \frac{1}{4}(-x_1 - x_2 + x_3 + x_4) \quad A_4 = \frac{1}{4}(-y_1 + y_2 + y_3 - y_4)$$

$$A_5 = \frac{1}{4}(y_1 - y_2 + y_3 - y_4) \quad A_6 = \frac{1}{4}(-y_1 - y_2 + y_3 + y_4)$$

In the spirit of the present finite-volume direct averaging technique, we relate the partial derivatives of the function $f(\cdot, \cdot)$ with respect to the coordinates (x, y) to those with respect to the coordinates (η, ξ) through the inverse of the volume-averaged Jacobian $\bar{\mathbf{J}}$

$$\bar{\mathbf{J}} = \frac{1}{4} \int_{-1}^{+1} \int_{-1}^{+1} \mathbf{J} d\eta d\xi \quad (24)$$

so that its inverse $\bar{\mathbf{J}}^{-1}$, which replaces $\mathbf{J}^{-1}$ in Eq. (22), is

$$\bar{\mathbf{J}}^{-1} = \hat{\mathbf{J}} = \frac{1}{A_7} \begin{bmatrix} A_6 & -A_4 \\ -A_3 & A_1 \end{bmatrix} \quad (25)$$

where $A_7 = A_1 A_6 - A_3 A_4$. The discussion of this simplification is deferred to Sec. 6.

The surface-averaged partial derivatives of the function $f(\cdot, \cdot)$ with respect to the (x, y) coordinates are then linearly related to the corresponding surface-averaged partial derivatives with respect to the (η, ξ) coordinates through $\hat{\mathbf{J}}$. Evaluating partial derivatives of $f(\eta, \xi)$ along the four faces of the reference subcell, and taking the surface average of the resulting expressions on each face, the surface-averaged derivatives with respect to the (x, y) coordinates are obtained in terms of the first and second order coefficients $F_{(10)}, \dots, F_{(02)}$,

$$\begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}_{\eta=\pm 1} = \hat{\mathbf{J}} \begin{bmatrix} \frac{\partial f}{\partial \eta} \\ \frac{\partial f}{\partial \xi} \end{bmatrix}_{\eta=\pm 1} = \hat{\mathbf{J}} \mathbf{A}_{1,2} \begin{bmatrix} F_{(10)} \\ F_{(01)} \\ F_{(20)} \\ F_{(02)} \end{bmatrix} \quad (26)$$

$$\begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}_{\xi=\pm 1} = \hat{\mathbf{J}} \begin{bmatrix} \frac{\partial f}{\partial \eta} \\ \frac{\partial f}{\partial \xi} \end{bmatrix}_{\xi=\pm 1} = \hat{\mathbf{J}} \mathbf{A}_{3,4} \begin{bmatrix} F_{(10)} \\ F_{(01)} \\ F_{(20)} \\ F_{(02)} \end{bmatrix} \quad (27)$$

where

$$\mathbf{A}_{1,2} = \begin{bmatrix} 1 & 0 & \pm 3 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \quad \mathbf{A}_{3,4} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & \pm 3 \end{bmatrix} \quad (28)$$

The first and second order coefficients in Eqs. (26) and (27) are directly related to the surface-averaged values of the function and the remaining unknown zeroth order coefficient $F_{(00)}$, see Eq. (21). This coefficient is obtained from the volume-averaged governing differential equation in terms of the surface-averaged functions. Satisfaction of the field equations within subcells of the actual microstructure requires the relationship between the second partial derivatives of the function $f(\cdot, \cdot)$ in the two planes. This relationship is obtained in terms of the products of the $\hat{\mathbf{J}}$ elements as follows:

$$\begin{bmatrix} \frac{\partial^2 F}{\partial x^2} \\ \frac{\partial^2 F}{\partial x \partial y} \\ \frac{\partial^2 F}{\partial y^2} \end{bmatrix} = \begin{bmatrix} (\hat{J}_{11})^2 & 2\hat{J}_{11}\hat{J}_{12} & (\hat{J}_{12})^2 \\ \hat{J}_{11}\hat{J}_{21} & \hat{J}_{11}\hat{J}_{22} + \hat{J}_{12}\hat{J}_{21} & \hat{J}_{12}\hat{J}_{22} \\ (\hat{J}_{21})^2 & 2\hat{J}_{21}\hat{J}_{22} & (\hat{J}_{22})^2 \end{bmatrix} \begin{bmatrix} \frac{\partial^2 F}{\partial \eta^2} \\ \frac{\partial^2 F}{\partial \eta \partial \xi} \\ \frac{\partial^2 F}{\partial \xi^2} \end{bmatrix} \quad (29)$$

The relationships derived in this section are employed in the following sections to develop local conductivity and stiffness matrices for the j th subcell in the actual microstructure. The actual details differ due to the differences in the number of field variables, constitutive equations, and governing differential equations.

4 Thermal Analysis

The local conductivity matrix relates the surface-averaged temperatures on each rectilinear face of the j th subcell in the actual

microstructure to the corresponding surface-averaged heat fluxes. The construction of this matrix involves direct satisfaction of the heat conduction equation within each j th subcell

$$\frac{\partial q_x}{\partial x} + \frac{\partial q_y}{\partial y} = -\rho C \frac{\partial T}{\partial t} \quad (30)$$

in a volume-average sense given the pointwise relationship between the heat flux and temperature gradient components through the Fourier heat conduction law. Assuming that each subcell contains isotropic material, the heat flux components are related to the partial derivatives of the temperature field in the x - y plane as follows:

$$\begin{bmatrix} q_x \\ q_y \end{bmatrix} = \begin{bmatrix} -k & 0 \\ 0 & -k \end{bmatrix} \begin{bmatrix} \frac{\partial T}{\partial x} \\ \frac{\partial T}{\partial y} \end{bmatrix} = \bar{\mathbf{k}} \begin{bmatrix} \frac{\partial T}{\partial x} \\ \frac{\partial T}{\partial y} \end{bmatrix} \quad (31)$$

where k is the thermal conductivity coefficient. In the absence of spatial dependence of this coefficient, Eq. (30) is expressed directly in terms of the temperature field partial derivatives using the above Fourier law

$$k \frac{\partial^2 T}{\partial x^2} + k \frac{\partial^2 T}{\partial y^2} = \rho C \frac{\partial T}{\partial t} \quad (32)$$

We begin the local conductivity matrix construction by approximating the temperature field in the reference subcell, following Eq. (19), by

$$T(\eta, \xi) = T_{(00)} + \eta T_{(10)} + \xi T_{(01)} + \frac{1}{2}(3\eta^2 - 1)T_{(20)} + \frac{1}{2}(3\xi^2 - 1)T_{(02)} \quad (33)$$

The surface-averaged temperatures $\bar{T}_i$ on each face F_i of the j th subcell in the actual microstructure are then obtained in terms of the zeroth, first, and second order coefficients $T_{(00)}, T_{(10)}, \dots, T_{(02)}$ in the form given by Eq. (20). Solving for the first and second order coefficients in terms of the surface-averaged temperatures and the zeroth order coefficient we have, according to Eq. (21),

$$\begin{bmatrix} T_{(10)} \\ T_{(01)} \\ T_{(20)} \\ T_{(02)} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 0 & 1 & 0 & -1 \\ -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \bar{T}_1 - T_{(00)} \\ \bar{T}_2 - T_{(00)} \\ \bar{T}_3 - T_{(00)} \\ \bar{T}_4 - T_{(00)} \end{bmatrix} \quad (34)$$

To construct the local conductivity matrix, we first determine the surface-averaged heat flux components on each face of the j th subcell in the x - y plane using Eq. (31). Taking the surface averages of this equation on each face of the reference subcell in the η - ξ plane, and expressing the surface-averaged partial derivatives of the temperature field with respect to the (x, y) coordinates in terms of the corresponding surface-averaged partial derivatives with respect to the (η, ξ) coordinates, and thus the first and second order coefficients $T_{(10)}, \dots, T_{(02)}$ using Eqs. (26) and (27), we obtain

$$\begin{bmatrix} \bar{q}_x \\ \bar{q}_y \end{bmatrix}_{\eta=\pm 1} = \bar{\mathbf{k}} \begin{bmatrix} \frac{\partial T}{\partial x} \\ \frac{\partial T}{\partial y} \end{bmatrix}_{\eta=\pm 1} = \bar{\mathbf{k}} \hat{\mathbf{J}} \begin{bmatrix} \frac{\partial T}{\partial \eta} \\ \frac{\partial T}{\partial \xi} \end{bmatrix}_{\eta=\pm 1} = \bar{\mathbf{k}} \hat{\mathbf{J}} \mathbf{A}_{1,2} \mathbf{T} \quad (35)$$

$$\begin{bmatrix} \bar{q}_x \\ \bar{q}_y \end{bmatrix}_{\xi=\pm 1} = \bar{\mathbf{k}} \begin{bmatrix} \frac{\partial T}{\partial x} \\ \frac{\partial T}{\partial y} \end{bmatrix}_{\xi=\pm 1} = \bar{\mathbf{k}} \hat{\mathbf{J}} \begin{bmatrix} \frac{\partial T}{\partial \eta} \\ \frac{\partial T}{\partial \xi} \end{bmatrix}_{\xi=\pm 1} = \bar{\mathbf{k}} \hat{\mathbf{J}} \mathbf{A}_{3,4} \mathbf{T} \quad (36)$$

where $\mathbf{T} = [T_{(10)}, T_{(01)}, T_{(20)}, T_{(02)}]^T$. These surface-averaged heat fluxes are then projected onto the unit normals to each subcell face in the actual microstructure

$$\bar{q}_{1,3} = [n_x^{(1,3)} \ n_y^{(1,3)}] \begin{bmatrix} \bar{q}_x \\ \bar{q}_y \end{bmatrix}_{\xi=\mp 1} \quad \text{and} \quad \bar{q}_{2,4} = [n_x^{(2,4)} \ n_y^{(2,4)}] \begin{bmatrix} \bar{q}_x \\ \bar{q}_y \end{bmatrix}_{\eta=\pm 1} \quad (37)$$

This leads to the normal surface-averaged heat fluxes expressed directly in terms of the first and second order temperature coefficients $T_{(10)}, \dots, T_{(02)}$

$$\begin{bmatrix} \bar{q}_1 \\ \bar{q}_2 \\ \bar{q}_3 \\ \bar{q}_4 \end{bmatrix} = \bar{\mathbf{A}} \begin{bmatrix} T_{(10)} \\ T_{(01)} \\ T_{(20)} \\ T_{(02)} \end{bmatrix} \quad (38)$$

where the matrix $\bar{\mathbf{A}}$ is expressed as a product of four matrices containing the submatrices $\bar{\mathbf{k}}, \hat{\mathbf{J}}, \mathbf{A}_1, \dots, \mathbf{A}_4$ and $\mathbf{n}^{(1)}, \dots, \mathbf{n}^{(4)}$ obtained from Eqs. (35)–(37), see Appendix A.

The temperature coefficients $T_{(10)}, \dots, T_{(02)}$ are related to the surface-averaged temperatures on each face of the j th subcell and the zeroth order coefficient $T_{(00)}$ through Eq. (34). In order to complete the construction of the local conductivity matrix for the j th subcell, the volume-averaged heat conduction equation is employed to relate the zeroth order coefficient $T_{(00)}$ to the surface-averaged temperatures. Using Eq. (29) to relate the second partial derivatives of the temperature field in the x - y plane to those in the η - ξ plane in the heat conduction equation given by Eq. (32), and explicitly evaluating these partial derivatives

$$\frac{\partial^2 T}{\partial \eta^2} = 3T_{(20)} \quad \frac{\partial^2 T}{\partial \eta \partial \xi} = 0 \quad \frac{\partial^2 T}{\partial \xi^2} = 3T_{(02)} \quad (39)$$

the volume-averaged heat conduction equation becomes

$$3[k(\hat{J}_{11})^2 + k(\hat{J}_{21})^2]T_{(20)} + 3[k(\hat{J}_{12})^2 + k(\hat{J}_{22})^2]T_{(02)} = \rho C \frac{\partial T}{\partial t} \quad (40)$$

which is then expressed in terms of the four surface-averaged temperatures and the zeroth order coefficient $T_{(00)}$ through the use of Eq. (34). The solution for $T_{(00)}$ depends on whether a steady-state or transient thermal conduction problem is considered. For both cases, the solution for $T_{(00)}$ can be symbolically represented by the expression

$$T_{(00)} = \lambda(\bar{T}_2 + \bar{T}_4) + \omega(\bar{T}_1 + \bar{T}_3) + \gamma \quad (41)$$

where the parameters λ , ω , and γ take different forms in each case. Using this representation for $T_{(00)}$ in Eq. (34), the first and second order temperature coefficients are directly related to the surface-averaged temperatures and a vector containing γ as follows:

$$\begin{bmatrix} T_{(10)} \\ T_{(01)} \\ T_{(20)} \\ T_{(02)} \end{bmatrix} = \begin{bmatrix} 0 & 1/2 & 0 & -1/2 \\ -1/2 & 0 & 1/2 & 0 \\ -\omega & 1/2 - \lambda & -\omega & 1/2 - \lambda \\ 1/2 - \omega & -\lambda & 1/2 - \omega & -\lambda \end{bmatrix} \begin{bmatrix} \bar{T}_1 \\ \bar{T}_2 \\ \bar{T}_3 \\ \bar{T}_4 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ \gamma \\ \gamma \end{bmatrix} \\ = \bar{\mathbf{B}} \begin{bmatrix} \bar{T}_1 \\ \bar{T}_2 \\ \bar{T}_3 \\ \bar{T}_4 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ \gamma \\ \gamma \end{bmatrix} \quad (42)$$

When transient effects are present or cannot be ignored, the temperature change with respect to time on the right-hand side of Eq. (40) is approximated by

$$\frac{\partial T}{\partial t} \approx \frac{\Delta T}{\Delta t} \approx \frac{\bar{\Delta T}}{\Delta t} = \frac{T_{(00)}^k - T_{(00)}^{k-1}}{\Delta t}$$

Therefore, the volume-averaged heat-conduction is discretized as follows:

$$[k(\hat{J}_{11})^2 + k(\hat{J}_{21})^2]T_{(20)}^k + [k(\hat{J}_{12})^2 + k(\hat{J}_{22})^2]T_{(02)}^k \\ = \frac{\rho C}{3} \left(\frac{T_{(00)}^k - T_{(00)}^{k-1}}{\Delta t} \right) \quad (43)$$

where $T_{(00)}^{k-1}$ is known from the $(k-1)$ th time step and temperature dependence of thermal (and mechanical) properties, which is often important and can be easily incorporated, is not considered herein. In this case, using Eq. (34) to express the second order coefficients in terms of the surface-averaged temperatures, and then solving for $T_{(00)}$ we obtain the following expressions for λ , ω , and γ in Eq. (41):

$$\psi = \frac{\rho C}{\Delta t} + 3k[(\hat{J}_{11})^2 + (\hat{J}_{12})^2] + 3k[(\hat{J}_{21})^2 + (\hat{J}_{22})^2]$$

$$\lambda = \frac{3}{2\psi} [k(\hat{J}_{11})^2 + k(\hat{J}_{21})^2]$$

$$\omega = \frac{3}{2\psi} [k(\hat{J}_{12})^2 + k(\hat{J}_{22})^2]$$

$$\gamma = \frac{1}{\psi} \frac{\rho C}{\Delta t} T_{(00)}^{k-1} \quad (44)$$

In the absence of time dependence of the temperature field, such as may occur when the transients have died out, or when the thermal conductivity is very large with respect to the heat capacity, or the thermal boundary conditions are applied very slowly, the right-hand side of Eq. (40) is zero and the above expressions for ψ and γ in the absence of time effects simplify to

$$\psi = 3k[(\hat{J}_{11})^2 + (\hat{J}_{12})^2] + 3k[(\hat{J}_{21})^2 + (\hat{J}_{22})^2]$$

$$\gamma = 0 \quad (45)$$

with λ and ω retaining the forms given in Eqs. (44).

4.1 Local Conductivity Matrix. The local conductivity matrices for the transient and steady-state heat conduction problems are then constructed by replacing the first and second order temperature coefficients on the right-hand side of Eq. (38) by the surface-averaged temperatures and the vector containing γ using Eq. (42). For both types of heat conduction problems, the local conductivity matrix for the j th subcell can be represented using the same symbolic notation

$$\begin{bmatrix} \bar{q}_1 \\ \bar{q}_2 \\ \bar{q}_3 \\ \bar{q}_4 \end{bmatrix}^{(j)} = \kappa^{(j)} \begin{bmatrix} \bar{T}_1 \\ \bar{T}_2 \\ \bar{T}_3 \\ \bar{T}_4 \end{bmatrix}^{(j)} - \mathbf{q}_0^{(j)} \quad (46)$$

where for the transient case the local conductivity matrix $\kappa^{(j)}$ and the initial heat flux vector $\mathbf{q}_0^{(j)}$ are given by

$$\kappa^{(j)} = \bar{\mathbf{A}} \bar{\mathbf{B}}^{\text{tr}} \quad \mathbf{q}_0^{(j)} = \bar{\mathbf{A}} [0 \ 0 \ \gamma \ \gamma]^T \quad (47)$$

with λ and ω appearing in the matrix $\bar{\mathbf{B}}^{\text{tr}}$ given by Eqs. (44). For the steady-state case,

$$\kappa^{(j)} = \bar{\mathbf{A}} \bar{\mathbf{B}}^{\text{ss}} \quad \mathbf{q}_0^{(j)} = \mathbf{0} \quad (48)$$

with λ and ω appearing in the matrix $\bar{\mathbf{B}}^{\text{ss}}$ specialized according to Eqs. (45).

4.2 Global Conductivity Matrix. The local conductivity matrices are assembled into a global system of equations by applying the surface-averaged interfacial temperature continuity and heat flux balance conditions, followed by the specified boundary conditions. In this approach, redundant temperature continuity equations are eliminated by setting surface-averaged temperatures at interfaces associated with adjacent subcells to common unknowns in conjunction with the application of the heat flux balance conditions at the common interfaces. The assembly is similar to that used in the finite-element algorithms, in contrast with the assembly used in the reconstructed finite-volume theory based on the rectangular discretization which produces distinct rows and columns of subcells absent in the parametric formulation. In particular, in the parametric formulation the position of the j th subcell within the entire structure is defined by the global subcell vertices x_i, y_i and by the global reference indices 1, 2, 3, 4 assigned to the subcell faces F_i . The connectivity matrix for the subcell faces F_i defines the position of the individual local conductivity matrix elements in the global conductivity matrix. The assembly procedure allows to utilize existing algorithms developed for the finite-element method.

The resulting global thermal conductivity matrix relates the unknown interfacial and boundary surface-averaged temperatures to the surface-averaged heat fluxes prescribed at the external boundaries of the heterogeneous material/structure

$$\kappa \bar{\mathbf{T}} - \bar{\mathbf{Q}}_0 = \bar{\mathbf{Q}} \quad (49)$$

where $\bar{\mathbf{T}}$ contains all the unknown common interfacial and boundary surface-averaged temperatures, $\bar{\mathbf{Q}}_0$ contains the initial heat flux distribution within the investigated component, and $\bar{\mathbf{Q}}$ contains information on the surface-averaged heat fluxes along the internal interfaces and the discretized boundary. It consists mainly of zeros which are obtained after applying the interfacial heat flux balance conditions across each interface separating adjacent subcells, with the nonzero terms representing the surface-averaged heat fluxes along each portion of the discretized boundary. The above system of equations is modified accordingly when the boundary conditions are specified in terms of applied temperatures (either directly or through convective relations), Cavalcante [15].

5 Mechanical Analysis

The local stiffness matrix relates the surface-averaged displacements on each rectilinear face of a j th subcell in the actual microstructure to the corresponding surface-averaged tractions. The construction of this matrix parallels that of the local conductivity matrix with modifications that account for vectorial rather than scalar field variables and the attendant governing differential and constitutive equations. For plane problems in the x - y plane and in

the absence of body forces, the stress equilibrium equations that are directly satisfied in the volume-average sense during the local stiffness matrix construction are

$$\begin{aligned} \frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xy}}{\partial y} &= 0 \\ \frac{\partial \sigma_{xy}}{\partial x} + \frac{\partial \sigma_{yy}}{\partial y} &= 0 \end{aligned} \quad (50)$$

For an isotropic elastic material occupying the j th subcell, stress components are related to strain components through the familiar Hooke's law

$$\begin{bmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \sigma_{xy} \end{bmatrix} = \begin{bmatrix} \bar{C}_{xx} & \bar{C}_{xy} & 0 \\ \bar{C}_{xy} & \bar{C}_{xx} & 0 \\ 0 & 0 & \frac{1}{2}(\bar{C}_{xx} - \bar{C}_{xy}) \end{bmatrix} \begin{bmatrix} \varepsilon_{xx} \\ \varepsilon_{yy} \\ 2\varepsilon_{xy} \end{bmatrix} - \begin{bmatrix} \Gamma \\ \Gamma \\ 0 \end{bmatrix} \Delta T = \bar{\mathbf{C}} \begin{bmatrix} \varepsilon_{xx} \\ \varepsilon_{yy} \\ 2\varepsilon_{xy} \end{bmatrix} - \Gamma \Delta T \quad (51)$$

where the stiffness matrix elements and thermal contributions for the plane strain and plane stress cases are given in terms of the thermoelastic constants in different forms, cf. Sokolnikoff [16]. For the plane strain case $\bar{\varepsilon}_{33}=0$ and for the plane stress case $\bar{\sigma}_{33}$ is determined from the condition $\bar{\sigma}_{33}=0$. Using the strain-displacement relations, in conjunction with the constitutive equations, in the stress equilibrium equations we obtain the Navier's equations for plane problems

$$\begin{aligned} \bar{C}_{xx} \frac{\partial^2 u}{\partial x^2} + \frac{1}{2}(\bar{C}_{xx} - \bar{C}_{xy}) \frac{\partial^2 u}{\partial y^2} + \frac{1}{2}(\bar{C}_{xx} + \bar{C}_{xy}) \frac{\partial^2 v}{\partial x \partial y} &= \frac{\partial}{\partial x} (\Gamma \Delta T) \\ \frac{1}{2}(\bar{C}_{xx} + \bar{C}_{xy}) \frac{\partial^2 u}{\partial x \partial y} + \frac{1}{2}(\bar{C}_{xx} - \bar{C}_{xy}) \frac{\partial^2 v}{\partial x^2} + \bar{C}_{xx} \frac{\partial^2 v}{\partial y^2} &= \frac{\partial}{\partial y} (\Gamma \Delta T) \end{aligned} \quad (52)$$

We begin the local stiffness matrix construction by approximating the displacement field in the reference subcell, following Eq. (19), by

$$\begin{aligned} u(\eta, \xi) &= U_{1(00)} + \eta U_{1(10)} + \xi U_{1(01)} + \frac{1}{2}(3\eta^2 - 1)U_{1(20)} \\ &\quad + \frac{1}{2}(3\xi^2 - 1)U_{1(02)} \\ v(\eta, \xi) &= U_{2(00)} + \eta U_{2(10)} + \xi U_{2(01)} + \frac{1}{2}(3\eta^2 - 1)U_{2(20)} \\ &\quad + \frac{1}{2}(3\xi^2 - 1)U_{2(02)} \end{aligned} \quad (53)$$

The surface-averaged displacement components $\bar{u}_i$ and $\bar{v}_i$ on each face F_i of the j th subcell in the actual microstructure are then obtained in terms of the zeroth, first, and second-order coefficients of the displacement field in the form given by Eq. (20) for each set. Solving for the first and second order coefficients in terms of the surface-averaged displacement components and the zeroth order coefficient we have, according to Eq. (21),

$$\begin{bmatrix} U_{1(10)} \\ U_{1(01)} \\ U_{1(20)} \\ U_{1(02)} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 0 & 1 & 0 & -1 \\ -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \bar{u}_1 - U_{1(00)} \\ \bar{u}_2 - U_{1(00)} \\ \bar{u}_3 - U_{1(00)} \\ \bar{u}_4 - U_{1(00)} \end{bmatrix} \quad (54)$$

and similarly

$$\begin{bmatrix} U_{2(10)} \\ U_{2(01)} \\ U_{2(20)} \\ U_{2(02)} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 0 & 1 & 0 & -1 \\ -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \bar{v}_1 - U_{2(00)} \\ \bar{v}_2 - U_{2(00)} \\ \bar{v}_3 - U_{2(00)} \\ \bar{v}_4 - U_{2(00)} \end{bmatrix} \quad (55)$$

To construct the local stiffness matrix, we first determine the surface-averaged in-plane stress components on each face of the j th subcell in the x - y plane using Eq. (51). Taking the surface averages of this equation on each face of the reference subcell in the η - ξ plane, introducing the matrix $\bar{\mathbf{E}}$ which relates in-plane strains to all four derivatives of the displacement components

$$\begin{bmatrix} \varepsilon_{xx} \\ \varepsilon_{yy} \\ 2\varepsilon_{xy} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{\partial u}{\partial x} \\ \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} \\ \frac{\partial v}{\partial y} \end{bmatrix} = \bar{\mathbf{E}} \begin{bmatrix} \frac{\partial u}{\partial x} \\ \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} \\ \frac{\partial v}{\partial y} \end{bmatrix} \quad (56)$$

and expressing the surface-averaged partial derivatives of the displacement field with respect to the (x, y) coordinates in terms of the corresponding surface-averaged partial derivatives with respect to the (η, ξ) coordinates, we have

$$\begin{bmatrix} \bar{\sigma}_{xx} \\ \bar{\sigma}_{yy} \\ \bar{\sigma}_{xy} \end{bmatrix}_{\eta=\pm 1} = \bar{\mathbf{C}} \bar{\mathbf{E}} \begin{bmatrix} \frac{\partial u}{\partial x} \\ \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} \\ \frac{\partial v}{\partial y} \end{bmatrix}_{\eta=\pm 1} - \Gamma \begin{bmatrix} \Delta \bar{T} \\ \Delta \bar{T} \\ 0 \end{bmatrix}_{\eta=\pm 1} = \bar{\mathbf{C}} \bar{\mathbf{E}} \begin{bmatrix} \hat{\mathbf{J}} & \mathbf{0} \\ \mathbf{0} & \hat{\mathbf{J}} \end{bmatrix} \begin{bmatrix} \frac{\partial u}{\partial \eta} \\ \frac{\partial u}{\partial \xi} \\ \frac{\partial v}{\partial \eta} \\ \frac{\partial v}{\partial \xi} \end{bmatrix}_{\eta=\pm 1} - \Gamma \begin{bmatrix} \Delta \bar{T} \\ \Delta \bar{T} \\ 0 \end{bmatrix}_{\eta=\pm 1} \quad (57)$$

$$\begin{bmatrix} \bar{\sigma}_{xx} \\ \bar{\sigma}_{yy} \\ \bar{\sigma}_{xy} \end{bmatrix}_{\xi=\pm 1} = \bar{\mathbf{C}} \bar{\mathbf{E}} \begin{bmatrix} \frac{\partial u}{\partial x} \\ \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} \\ \frac{\partial v}{\partial y} \end{bmatrix}_{\xi=\pm 1} - \Gamma \begin{bmatrix} \Delta \bar{T} \\ \Delta \bar{T} \\ 0 \end{bmatrix}_{\xi=\pm 1} = \bar{\mathbf{C}} \bar{\mathbf{E}} \begin{bmatrix} \hat{\mathbf{J}} & \mathbf{0} \\ \mathbf{0} & \hat{\mathbf{J}} \end{bmatrix} \begin{bmatrix} \frac{\partial u}{\partial \eta} \\ \frac{\partial u}{\partial \xi} \\ \frac{\partial v}{\partial \eta} \\ \frac{\partial v}{\partial \xi} \end{bmatrix}_{\xi=\pm 1} - \Gamma \begin{bmatrix} \Delta \bar{T} \\ \Delta \bar{T} \\ 0 \end{bmatrix}_{\xi=\pm 1} \quad (58)$$

The surface-averaged displacement derivatives with respect to the (η, ξ) coordinates are then expressed in terms of the first and second order coefficients $U_{1(10)}, \dots, U_{1(02)}$ and $U_{2(10)}, \dots, U_{2(02)}$ using Eqs. (26) and (27),

$$\begin{bmatrix} \bar{\sigma}_{xx} \\ \bar{\sigma}_{yy} \\ \bar{\sigma}_{xy} \end{bmatrix}_{\eta=\pm 1} = \bar{\mathbf{C}} \bar{\mathbf{E}} \begin{bmatrix} \hat{\mathbf{J}} & \mathbf{0} \\ \mathbf{0} & \hat{\mathbf{J}} \end{bmatrix} \begin{bmatrix} \mathbf{A}_{1,2} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{1,2} \end{bmatrix} \begin{bmatrix} \mathbf{U}_1 \\ \mathbf{U}_2 \end{bmatrix} - \Gamma \begin{bmatrix} \Delta \bar{T} \\ \Delta \bar{T} \\ 0 \end{bmatrix}_{\eta=\pm 1} \quad (59)$$

$$\begin{bmatrix} \bar{\sigma}_{xx} \\ \bar{\sigma}_{yy} \\ \bar{\sigma}_{xy} \end{bmatrix}_{\xi=\pm 1} = \bar{\mathbf{C}} \bar{\mathbf{E}} \begin{bmatrix} \hat{\mathbf{J}} & \mathbf{0} \\ \mathbf{0} & \hat{\mathbf{J}} \end{bmatrix} \begin{bmatrix} \mathbf{A}_{3,4} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{3,4} \end{bmatrix} \begin{bmatrix} \mathbf{U}_1 \\ \mathbf{U}_2 \end{bmatrix} - \Gamma \begin{bmatrix} \Delta \bar{T} \\ \Delta \bar{T} \\ 0 \end{bmatrix}_{\xi=\pm 1} \quad (60)$$

where

$$\mathbf{U}_1 = [U_{1(10)}, U_{1(01)}, U_{1(20)}, U_{1(02)}]^T \text{ and}$$

$$\mathbf{U}_2 = [U_{2(10)}, U_{2(01)}, U_{2(20)}, U_{2(02)}]^T$$

The surface-averaged tractions $\mathbf{t}^{(i)}$ on each face F_i of a j th subcell with the unit normal $\mathbf{n}^{(i)}$ are then obtained in terms of the corresponding stresses using Cauchy's stress relations $\mathbf{t}^{(i)} = \boldsymbol{\sigma} \cdot \mathbf{n}^{(i)}$ which can be written in a unified manner in terms of the individual components as follows:

$$\begin{bmatrix} \bar{t}_x \\ \bar{t}_y \end{bmatrix}^{(1,3)} = \begin{bmatrix} n_x & 0 & n_y \\ 0 & n_y & n_x \end{bmatrix}^{(1,3)} \begin{bmatrix} \bar{\sigma}_{xx} \\ \bar{\sigma}_{yy} \\ \bar{\sigma}_{xy} \end{bmatrix}_{\xi=\mp 1} \quad \text{and} \quad \begin{bmatrix} \bar{t}_x \\ \bar{t}_y \end{bmatrix}^{(2,4)} = \begin{bmatrix} n_x & 0 & n_y \\ 0 & n_y & n_x \end{bmatrix}^{(2,4)} \begin{bmatrix} \bar{\sigma}_{xx} \\ \bar{\sigma}_{yy} \\ \bar{\sigma}_{xy} \end{bmatrix}_{\eta=\pm 1} \quad (61)$$

where the surface-averaged stress components on the faces $F_{1,3}$ and $F_{2,4}$ are evaluated at $\xi = \mp 1$ and $\eta = \pm 1$, respectively. This

leads to the surface-averaged traction components acting on each face expressed directly in terms of the first and second order displacement coefficients $U_{1(10)}, \dots, U_{1(02)}$ and $U_{2(10)}, \dots, U_{2(02)}$

$$\begin{bmatrix} \bar{t}_x^{(1)} \\ \bar{t}_y^{(1)} \\ \bar{t}_x^{(2)} \\ \bar{t}_y^{(2)} \\ \bar{t}_x^{(3)} \\ \bar{t}_y^{(3)} \\ \bar{t}_x^{(4)} \\ \bar{t}_y^{(4)} \end{bmatrix} = \bar{\mathbf{A}} \begin{bmatrix} U_{1(10)} \\ U_{1(01)} \\ U_{1(20)} \\ U_{1(02)} \\ U_{2(10)} \\ U_{2(01)} \\ U_{2(20)} \\ U_{2(02)} \end{bmatrix} - \Gamma \mathbf{D} \Delta \mathbf{T} \quad (62)$$

where $\Delta \mathbf{T} = \Delta[\bar{T}_1 \bar{T}_1 \ 0 \ \bar{T}_2 \bar{T}_2 \ 0 \ \bar{T}_3 \bar{T}_3 \ 0 \ \bar{T}_4 \bar{T}_4 \ 0]^T$, $\mathbf{D}$ contains the unit normal components associated with the four rectilinear faces, Eq. (61), and the matrix $\bar{\mathbf{A}}$ is a product of $\mathbf{D}$ and four other matrices containing the submatrices $\bar{\mathbf{C}}, \bar{\mathbf{E}}, \hat{\mathbf{J}}$ and $\mathbf{A}_1, \dots, \mathbf{A}_4$ obtained from Eqs. (59) and (60) which are given explicitly in Appendix B.

The first and second order displacement coefficients are related to the surface-averaged displacements on each face of the j th subcell and the zeroth order displacements through Eqs. (54) and (55). In order to complete the construction of the local stiffness matrix for the j th subcell, the volume-averaged stress equilibrium equations are employed to relate the zeroth order coefficients to the surface-averaged displacements. Using Eq. (29) to relate the second partial derivatives of the displacement field in the x - y plane to those in the η - ξ plane in the stress equilibrium equations given by Eqs. (52), and explicitly evaluating these partial derivatives

$$\begin{aligned} \frac{\partial^2 u}{\partial \eta^2} &= 3U_{1(20)} & \frac{\partial^2 u}{\partial \eta \partial \xi} &= 0 & \frac{\partial^2 u}{\partial \xi^2} &= 3U_{1(02)} \\ \frac{\partial^2 v}{\partial \eta^2} &= 3U_{2(20)} & \frac{\partial^2 v}{\partial \eta \partial \xi} &= 0 & \frac{\partial^2 v}{\partial \xi^2} &= 3U_{2(02)} \end{aligned} \quad (63)$$

the volume-averaged equilibrium equations become

$$\begin{aligned} & \left[\bar{C}_{xx}(\hat{J}_{11})^2 + \frac{1}{2}(\bar{C}_{xx} - \bar{C}_{xy})(\hat{J}_{21})^2 \right] U_{1(20)} + \left[\bar{C}_{xx}(\hat{J}_{12})^2 + \frac{1}{2}(\bar{C}_{xx} - \bar{C}_{xy}) \right. \\ & \quad \times (\hat{J}_{22})^2 \left. \right] U_{1(02)} + \frac{1}{2}(\bar{C}_{xx} + \bar{C}_{xy})\hat{J}_{11}\hat{J}_{21}U_{2(20)} \\ & \quad + \frac{1}{2}(\bar{C}_{xx} + \bar{C}_{xy})\hat{J}_{12}\hat{J}_{22}U_{2(02)} = \frac{\Gamma}{3}[\hat{J}_{11}T_{(10)} + \hat{J}_{12}T_{(01)}] \\ & \left[\bar{C}_{xx}(\hat{J}_{21})^2 + \frac{1}{2}(\bar{C}_{xx} - \bar{C}_{xy})(\hat{J}_{11})^2 \right] U_{2(20)} + \left[\bar{C}_{xx}(\hat{J}_{22})^2 + \frac{1}{2}(\bar{C}_{xx} - \bar{C}_{xy}) \right. \\ & \quad \times (\hat{J}_{12})^2 \left. \right] U_{2(02)} + \frac{1}{2}(\bar{C}_{xx} + \bar{C}_{xy})\hat{J}_{11}\hat{J}_{21}U_{1(20)} \\ & \quad + \frac{1}{2}(\bar{C}_{xx} + \bar{C}_{xy})\hat{J}_{12}\hat{J}_{22}U_{1(02)} = \frac{\Gamma}{3}[\hat{J}_{21}T_{(10)} + \hat{J}_{22}T_{(01)}] \end{aligned} \quad (64)$$

where the volume averages of the temperature gradient components have been approximated as follows:

$$\frac{\partial T}{\partial x} = \frac{1}{4} \int_{-1}^{+1} \int_{-1}^{+1} \frac{\partial T}{\partial x} d\eta d\xi = \hat{J}_{11}T_{(10)} + \hat{J}_{12}T_{(01)}$$

$$\frac{\partial T}{\partial y} = \frac{1}{4} \int_{-1}^{+1} \int_{-1}^{+1} \frac{\partial T}{\partial y} d\eta d\xi = \hat{J}_{21}T_{(10)} + \hat{J}_{22}T_{(01)} \quad (65)$$

The above volume-averaged equilibrium equations are then expressed in terms of the eight surface-averaged displacement components and the zeroth order coefficients $U_{1(00)}$ and $U_{2(00)}$ through the use of Eqs. (54) and (55). The solution for these two zeroth order coefficients produces

$$\begin{bmatrix} U_{1(00)} \\ U_{2(00)} \end{bmatrix} = \Phi^{-1} \Theta \begin{bmatrix} \bar{u}_2 + \bar{u}_4 \\ \bar{u}_1 + \bar{u}_3 \\ \bar{v}_2 + \bar{v}_4 \\ \bar{v}_1 + \bar{v}_3 \end{bmatrix} + \Phi^{-1} \Omega \quad (66)$$

where the elements of the matrices Φ , Θ , and Ω are given in Appendix B. Using this representation for $U_{1(00)}$ and $U_{2(00)}$ in Eqs. (54) and (55), the first and second order displacement coefficients are directly related to the surface-averaged displacement components and a vector containing first order temperature coefficients as follows:

$$\begin{bmatrix} U_{1(10)} \\ U_{1(01)} \\ U_{1(20)} \\ U_{1(02)} \\ U_{2(10)} \\ U_{2(01)} \\ U_{2(20)} \\ U_{2(02)} \end{bmatrix} = \bar{\mathbf{B}} \begin{bmatrix} \bar{u}_1 \\ \bar{v}_1 \\ \bar{u}_2 \\ \bar{v}_2 \\ \bar{u}_3 \\ \bar{v}_3 \\ \bar{u}_4 \\ \bar{v}_4 \end{bmatrix} - \mathbf{N} \Phi^{-1} \Omega \quad (67)$$

where $\bar{\mathbf{B}} = \mathbf{P} - \mathbf{N} \Phi^{-1} \Theta \mathbf{M}$ and the matrices $\mathbf{M}$, $\mathbf{N}$, and $\mathbf{P}$ are given in Appendix B in explicit forms.

5.1 Local Stiffness Matrix. The local stiffness matrix for the transient and steady-state thermomechanical problems is then constructed by replacing the first and second order displacement coefficients on the right-hand side of Eq. (62) by the surface-averaged displacements and the vector containing thermal contributions using Eq. (67). For both types of thermomechanical problems, the local stiffness matrix for the j th subcell can be represented using the same symbolic notation

$$\begin{bmatrix} \bar{t}_1 \\ \bar{t}_2 \\ \bar{t}_3 \\ \bar{t}_4 \end{bmatrix}^{(j)} = \mathbf{K}^{(j)} \begin{bmatrix} \bar{u}_1 \\ \bar{u}_2 \\ \bar{u}_3 \\ \bar{u}_4 \end{bmatrix}^{(j)} - \mathbf{t}_0^{(j)} \quad (68)$$

where $\bar{\mathbf{t}}_i = [\bar{t}_x^{(i)} \ \bar{t}_y^{(i)}]^T$ and $\bar{\mathbf{u}}_i = [\bar{u}_i \ \bar{v}_i]^T$, and the local stiffness matrix $\mathbf{K}^{(j)}$ and the initial traction vector $\mathbf{t}_0^{(j)}$ are given by

$$\mathbf{K}^{(j)} = \bar{\mathbf{A}} \bar{\mathbf{B}} \quad \mathbf{t}_0^{(j)} = \Gamma \mathbf{D} \Delta \mathbf{T} + \bar{\mathbf{A}} \mathbf{N} \Phi^{-1} \Omega \quad (69)$$

5.2 Global Stiffness Matrix. The local stiffness matrices are assembled into a global system of equations by applying surface-averaged interfacial traction and displacement continuity conditions, followed by the specified boundary conditions in the same manner as the local conductivity matrices. In this approach, redundant displacement continuity equations are eliminated by setting surface-averaged displacement components at the interfaces associated with adjacent subcells to common unknowns, upon application of traction continuity conditions at these interfaces. The resulting global stiffness matrix relates the unknown interfacial and boundary surface-averaged displacements to the surface-averaged tractions prescribed at the external boundaries of the heterogeneous material/structure

$$\mathbf{K}\bar{\mathbf{U}} - \bar{\mathbf{t}}_0 = \bar{\mathbf{t}} \quad (70)$$

where $\bar{\mathbf{U}}$ contains all the unknown interfacial and boundary surface-averaged displacements, $\bar{\mathbf{t}}_0$ contains the initial thermal loading effects, and $\bar{\mathbf{t}}$ contains information on the surface-averaged tractions along the internal interfaces and the discretized boundary. It consists mainly of zeros which are obtained after applying interfacial traction continuity conditions across each interface separating adjacent subcells, with the nonzero terms representing surface-averaged tractions along each portion of the discretized boundary. The above system of equations is then modified according to the specified boundary conditions, Cavalcante [15].

6 Discussion

The parametric formulation of the finite-volume theory for functionally graded materials based on the local/global conductivity and stiffness matrix approach is a significant step in the continuing development of this theory for the modeling of materials with heterogeneous microstructures. In particular, modeling of microstructural details has been made substantially more efficient and accurate through the ability to more precisely capture continuous inclusions with arbitrarily-shaped cross sections, as has modeling of arbitrarily-shaped external boundaries of structural components. In this sense, the theory has been brought closer to the finite-element approach without the concomitant convergence problems at the interfaces separating heterogeneities with large material property mismatch, but with the added advantage of profiting from existing preprocessing and postprocessing algorithms developed for the finite-element method. Explicit representation of the temperature and displacement fields within the subcells used to discretize the material's microstructure, application of both the displacement and traction continuity conditions across interfaces separating adjacent heterogeneities, and satisfaction of field equations in the large within each subcell, make this approach distinct from the finite-element method and the finite-volume technique used in computational fluid dynamics. These features also produce better solution convergence relative to the finite-element method when continuous inclusions with large elastic moduli contrast in a heterogeneous microstructure are modeled by single subcells/elements with rectangular cross sections, as demonstrated by Bansal and Pindera [9]. The parametric formulation facilitates comparison with the finite-element method in a more direct way when curved boundaries or continuous inclusions are involved, as will be demonstrated in Part II.

The parametric formulation now makes the method competitive with the finite-element approach vis-à-vis its ability to model arbitrary two-dimensional geometries with the added advantage of accommodating continuously reinforced heterogeneous microstructures in a more efficient manner. The full advantage of the two-dimensional version will be realized upon incorporating spatially variable material properties at the subcell level, as was done by Kim and Paulino [17] in the context of the finite-element method and by Zhong et al. [10] in the context of the standard version of the finite-volume theory. This will enable unified multiscale analysis of graded microstructures with two levels of microstructural scales not easily amenable to analysis by the finite-element method. Incorporation of three-dimensional mapping capability into the parametric version will enable modeling of heterogeneous microstructures containing arbitrarily-shaped particulate inclusions. This is an active area of research spanning a wide range of applications ranging from particulate metal matrix composites to mechanism-based simulation of polycrystalline materials. Modeling of actual three-dimensional microstructures with sufficient fidelity using the standard finite-volume theory is prohibitively expensive and beyond the present computing capabilities.

The present development of the theory was simplified by using the volume-averaged Jacobian of the coordinate transformation to

relate flux-like quantities in the reference and actual coordinate systems. This, in turn, simplified the calculation of the corresponding surface-averaged quantities on the four faces of a subcell in the actual microstructure, as well as the satisfaction of heat conductivity and equilibrium equations, and ultimately the local conductivity and stiffness matrix constructions. These approximations may potentially affect the accuracy of the predicted results, and therefore their effect will be investigated in our future work. It is a straightforward, albeit tedious, procedure to account for the dependence of the coordinate transformation Jacobian on local coordinates when calculating surface-averaged heat fluxes and tractions at the expense of some loss in conciseness and ease of programming. Part II of this manuscript demonstrates that the above simplification produces acceptably accurate predictions of stress fields in curved homogeneous and graded structural components and heterogeneous materials with sufficient mesh refinement in comparison with the finite-element results.

7 Conclusions

A parametric formulation of the standard finite-volume theory for functionally graded materials has been developed based on the local/global conductivity and stiffness matrix approaches. Using parametric mapping of a reference subcell onto the corresponding subcell in the actual microstructure, easily programmable local conductivity and stiffness matrices have been obtained for quadrilateral subcells that can be used to more efficiently model microstructural details with curved boundaries previously approximated using rectangular subcells. The parametric version can also be used to model homogeneous or heterogeneous structural components with curved boundaries. Existing assembly algorithms developed for the finite-element method can be utilized for efficient construction and solution of the global systems of equations that govern thermomechanical problems of heterogeneous media. Verification of the implemented parametric formulation for homogeneous and heterogeneous components with curved and rectangular boundaries is presented in Part II, together with demonstration of this version's advantage over the standard theory based on rectangular subcell discretization.

Acknowledgment

M.A.A.C. and S.P.C.M. would like to gratefully acknowledge the financial support of the Brazilian federal agencies CNPq/PADCT and CAPES. M.J.P. acknowledges the support of the Engineered Materials Concepts, LLC.

Appendix A

The matrix $\bar{\mathbf{A}}$ in Eq. (38) is a product of the four matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{D}$ arranged in reverse order, namely $\bar{\mathbf{A}} = \mathbf{DCBA}$, which are given below

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_4 \\ \mathbf{A}_1 \\ \mathbf{A}_3 \\ \mathbf{A}_2 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} \hat{\mathbf{J}} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \hat{\mathbf{J}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \hat{\mathbf{J}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \hat{\mathbf{J}} \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} \bar{\mathbf{k}} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \bar{\mathbf{k}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \bar{\mathbf{k}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \bar{\mathbf{k}} \end{bmatrix}$$

$$\mathbf{D} = \begin{bmatrix} \mathbf{n}^{(1)} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{n}^{(2)} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{n}^{(3)} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{n}^{(4)} \end{bmatrix}$$

Appendix B

The matrix $\bar{\mathbf{A}}$ in Eq. (62) is a product of the five matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$, and $\mathbf{E}$ arranged as follows $\bar{\mathbf{A}} = \mathbf{DCEBA}$, which are given below

$$\begin{aligned}
\mathbf{A} &= \begin{bmatrix} \mathbf{A}_4 & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_4 \\ \mathbf{A}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_1 \\ \mathbf{A}_3 & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_3 \\ \mathbf{A}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 \end{bmatrix} & \mathbf{B} &= \begin{bmatrix} \hat{\mathbf{J}} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \hat{\mathbf{J}} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \hat{\mathbf{J}} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \hat{\mathbf{J}} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \hat{\mathbf{J}} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \hat{\mathbf{J}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \hat{\mathbf{J}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \hat{\mathbf{J}} \end{bmatrix} \\
\mathbf{C} &= \begin{bmatrix} \bar{\mathbf{C}} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \bar{\mathbf{C}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \bar{\mathbf{C}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \bar{\mathbf{C}} \end{bmatrix} & \mathbf{D} &= \begin{bmatrix} \hat{\mathbf{n}}^{(1)} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \hat{\mathbf{n}}^{(2)} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \hat{\mathbf{n}}^{(3)} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \hat{\mathbf{n}}^{(4)} \end{bmatrix} \\
& & \mathbf{E} &= \begin{bmatrix} \bar{\mathbf{E}} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \bar{\mathbf{E}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \bar{\mathbf{E}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \bar{\mathbf{E}} \end{bmatrix}
\end{aligned}$$

where the matrices $\hat{\mathbf{n}}^{(i)}$ are defined as

$$\hat{\mathbf{n}}^{(i)} = \begin{bmatrix} n_x & 0 & n_y \\ 0 & n_y & n_x \end{bmatrix}^{(i)}$$

The elements of the matrices Φ and Θ , and the vector Ω in Eq. (66) are

$$\Phi_{11} = \bar{C}_{xx}[(\hat{J}_{11})^2 + (\hat{J}_{12})^2] + \frac{1}{2}(\bar{C}_{xx} - \bar{C}_{xy})[(\hat{J}_{21})^2 + (\hat{J}_{22})^2]$$

$$\Phi_{12} = \frac{1}{2}(\bar{C}_{xx} + \bar{C}_{xy})[\hat{J}_{11}\hat{J}_{21} + \hat{J}_{12}\hat{J}_{22}]$$

$$\Phi_{21} = \Phi_{12}$$

$$\Phi_{22} = \bar{C}_{xx}[(\hat{J}_{21})^2 + (\hat{J}_{22})^2] + \frac{1}{2}(\bar{C}_{xx} - \bar{C}_{xy})[(\hat{J}_{11})^2 + (\hat{J}_{12})^2]$$

$$\Theta_{11} = \frac{1}{2} \left[\bar{C}_{xx}(\hat{J}_{11})^2 + \frac{1}{2}(\bar{C}_{xx} - \bar{C}_{xy})(\hat{J}_{21})^2 \right]$$

$$\Theta_{12} = \frac{1}{2} \left[\bar{C}_{xx}(\hat{J}_{12})^2 + \frac{1}{2}(\bar{C}_{xx} - \bar{C}_{xy})(\hat{J}_{22})^2 \right]$$

$$\Theta_{13} = \frac{1}{2}(\bar{C}_{xx} + \bar{C}_{xy})\hat{J}_{11}\hat{J}_{21}$$

$$\Theta_{14} = \frac{1}{2}(\bar{C}_{xx} + \bar{C}_{xy})\hat{J}_{12}\hat{J}_{22}$$

$$\Theta_{21} = \frac{1}{2}(\bar{C}_{xx} + \bar{C}_{xy})\hat{J}_{11}\hat{J}_{21}$$

$$\Theta_{22} = \frac{1}{2}(\bar{C}_{xx} + \bar{C}_{xy})\hat{J}_{12}\hat{J}_{22}$$

$$\Theta_{23} = \frac{1}{2} \left[\bar{C}_{xx}(\hat{J}_{21})^2 + \frac{1}{2}(\bar{C}_{xx} - \bar{C}_{xy})(\hat{J}_{11})^2 \right]$$

$$\Theta_{24} = \frac{1}{2} \left[\bar{C}_{xx}(\hat{J}_{22})^2 + \frac{1}{2}(\bar{C}_{xx} - \bar{C}_{xy})(\hat{J}_{12})^2 \right]$$

$$\Omega_1 = -\frac{1}{3}\Gamma[\hat{J}_{11}T_{(10)} + \hat{J}_{12}T_{(01)}]$$

$$\Omega_2 = -\frac{1}{3}\Gamma[\hat{J}_{21}T_{(10)} + \hat{J}_{22}T_{(01)}]$$

In order to write down the matrices $\mathbf{M}$, $\mathbf{N}$, and $\mathbf{P}$ that appear in Eq. (67) we define the following vectors $\mathbf{v}_i$:

$$\mathbf{v}_1 = [1, 0, 0, 0] \quad \mathbf{v}_2 = [0, 1, 0, 0] \quad \mathbf{v}_3 = [0, 0, 1, 0] \quad \mathbf{v}_4 = [0, 0, 0, 1]$$

so that

$$\begin{aligned}
\mathbf{M} &= \begin{bmatrix} \mathbf{v}_3 & \mathbf{v}_3 \\ \mathbf{v}_1 & \mathbf{v}_1 \\ \mathbf{v}_4 & \mathbf{v}_4 \\ \mathbf{v}_2 & \mathbf{v}_2 \end{bmatrix} & \mathbf{N} &= \begin{bmatrix} \mathbf{v}_3^T + \mathbf{v}_4^T & \mathbf{0} \\ \mathbf{0} & \mathbf{v}_3^T + \mathbf{v}_4^T \end{bmatrix} & \mathbf{P} &= \frac{1}{2} \begin{bmatrix} \mathbf{v}_3 & -\mathbf{v}_3 \\ -\mathbf{v}_1 & \mathbf{v}_1 \\ \mathbf{v}_3 & \mathbf{v}_3 \\ \mathbf{v}_1 & \mathbf{v}_1 \\ \mathbf{v}_4 & -\mathbf{v}_4 \\ -\mathbf{v}_2 & \mathbf{v}_2 \\ \mathbf{v}_4 & \mathbf{v}_4 \\ \mathbf{v}_2 & \mathbf{v}_2 \end{bmatrix}
\end{aligned}$$

References

- [1] Suresh, S., and Mortensen, A., 1997, *Fundamentals of Functionally Graded Materials*, Cambridge University Press, Cambridge, UK.
- [2] Miyamoto, Y., Kaysser, W. A., Rabin, B. H., Kawasaki, A., and Ford, R. G., 1999, *Functionally Graded Materials: Design, Processing and Applications*, Kluwer Academic, Dordrecht, The Netherlands.
- [3] Paulino, G. H., 2002, "Fracture of Functionally Graded Materials," *Eng. Fract. Mech.*, **69**(14–16), pp. 1519–1812.
- [4] Chatzigeorgiou, G., and Charalambakis, G., 2005, "Instability Analysis of Non-Homogeneous Materials Under Biaxial Loading," *Int. J. Plast.*, **21**(10), pp. 1970–1999.
- [5] Pindera, M.-J., Aboudi, J., and Arnold, S. M., 1995, "Limitations of the Uncoupled, RVE-Based Micromechanical Approach in the Analysis of Functionally Graded Composites," *Mech. Mater.*, **20**(1), pp. 77–94.
- [6] Aboudi, J., Pindera, M.-J., and Arnold, S. M., 1994, "Elastic Response of Metal Matrix Composites With Tailored Microstructures to Thermal Gradients," *Int. J. Solids Struct.*, **31**(10), pp. 1393–1428.
- [7] Aboudi, J., Pindera, M.-J., and Arnold, S. M., 1999, "Higher-Order Theory for Functionally Graded Materials," *Composites, Part B*, **30**(8), pp. 777–832.
- [8] Achenbach, J. D., 1975, *A Theory of Elasticity With Microstructure for Directionally Reinforced Composites*, Springer-Verlag, New York.
- [9] Bansal, Y., and Pindera, M.-J., 2003, "Efficient Reformulation of the Thermoelastic Higher-Order Theory for FGMs," *J. Therm. Stresses*, **26**(11/12), pp. 1055–1092; see also: NASA CR 2002-211909, November 2002.
- [10] Zhong, Y., Bansal, Y., and Pindera, M.-J., 2004, "Efficient Reformulation of the Thermal Higher-Order Theory for FGM's With Variable Thermal Conductivity," *Int. J. Comput. Eng. Sci.*, **5**(4), pp. 795–831; see also: NASA CR 2002-211910, November 2002.
- [11] Buefler, H., 1971, "Theory of Elasticity of a Multilayered Medium," *J. Elast.*, **1**, pp. 125–143.
- [12] Pindera, M.-J., 1991, "Local/Global Stiffness Matrix Formulation for Composite Materials and Structures," *Composites Eng.*, **1**(2), pp. 69–83.
- [13] Versteeg, H. K., and Malalasekera, W., 1995, *An Introduction to Computational Fluid Dynamics: The Finite Volume Method*, Prentice-Hall, New York.
- [14] Pindera, M.-J., Bansal, Y., and Zhong, Y., 2004, "Finite-Volume Direct Averaging Theory for Functionally Graded Materials (FVDAT-FGM)," NASA Disclosure of Invention and New Technology Form 1679, September, 14.
- [15] Cavalcante, M. A. A., 2006, "Modelling of the Transient Thermo-Mechanical Behavior of Composite Material Structures by the Finite-Volume Theory," M.S. thesis, Civil Engineering Department, Federal University of Alagoas, Maceio, Alagoas, Brazil, February 10.
- [16] Sokolnikoff, I. S., 1987, *Mathematical Theory of Elasticity*, Krieger, Malabar, FL (reprint edition).
- [17] Kim, J. H., and Paulino, G. H., 2002, "Isoparametric Graded Finite Elements for Nonhomogeneous Isotropic and Orthotropic Materials," *J. Appl. Mech.*, **69**(4), pp. 502–514.

Parametric Formulation of the Finite-Volume Theory for Functionally Graded Materials—Part II: Numerical Results

Marcio A. A. Cavalcante
Severino P. C. Marques

Center of Technology,
Federal University of Alagoas,
Maceio, Alagoas, Brazil

Marek-Jerzy Pindera
Civil Engineering Department,
University of Virginia,
Charlottesville, VA 22903

In Part I of this communication, the finite-volume theory for functionally graded materials was further extended to enable efficient analysis of structural components with curved boundaries, as well as efficient modeling of continuous inclusions with arbitrarily-shaped cross sections of a graded material's microstructure, previously approximated using discretizations by rectangular subcells. This was accomplished through a parametric formulation based on mapping of a reference square subcell onto a quadrilateral subcell resident in the actual microstructure. In Part II, the parametric formulation is verified through comparison with analytical solutions for homogeneous and graded curved structural components subjected to transient thermal and steady-state thermomechanical loading. Grading is modeled using piecewise uniform thermoelastic moduli assigned to each discretized region. Results for a heterogeneous microstructure in the form of a single inclusion embedded in the matrix phase of large dimensions are also generated and compared with the exact analytical solution, as well as with the results obtained using the standard version of the finite-volume theory based on rectangular discretization and the finite-element method. It is demonstrated that the parametric finite-volume theory is a very competitive alternative to the finite-element method based on the quality of results and execution time. [DOI: 10.1115/1.2722313]

Keywords: functionally graded materials, finite-volume theory, parametric formulation

1 Introduction

In Part I of this paper, a parametric formulation of the recently reconstructed finite-volume theory for functionally graded materials, Bansal and Pindera [1], Zhong et al. [2], and Pindera et al. [3], was developed in order to enable efficient modeling of microstructures and structural components that cannot be easily modeled using rectangular subvolumes (called subcells) employed in the original Cartesian-based theory. In the parametric formulation, the local conductivity and stiffness matrices of quadrilateral subcells that are used in the discretization of the actual microstructure were developed using a transformation that maps a reference square subcell onto the actual quadrilateral subcell. Such mapping capability offers the same flexibility as the finite-element method for modeling microstructures and geometries with curved and quadrilateral features, while retaining its quick convergence characteristics in the presence of highly heterogeneous microstructures.

In its most basic form, the parametric version of the finite-volume theory reduces to the standard version when the mapping preserves the reference subcell's mutually orthogonal faces. The numerical implementation of the parametric version for this special case was verified by Cavalcante [4] upon comparison of the predicted results for the transient temperature distribution in a rectangular slab graded in the through-thickness direction with the analytical solution provided by Sutradhar et al. [5]. This class of problems is important in its own right due to applications to thermal barrier coatings which are becoming widespread in many

technological areas, cf. Paulino [6] for some recent contributions to this area. Moreover, the transient case initially produces large temperature gradients leading to large stress gradients which require substantial discretization of the structural component to be captured accurately by the finite-volume theory. Identification of the required discretization for this case provides a guide for the discretization of structural components with curved boundaries under transient thermal loading.

In Part II of this paper, we illustrate the parametric formulation's capability to accurately capture thermomechanical fields in homogeneous and graded thick-walled cylinders subjected to steady-state and transient loading, as well as local stress fields in the vicinity of circular inclusions. Part II is organized as follows. Section 2 deals with homogeneous and graded thick-walled cylinders subjected to transient as well as steady-state thermomechanical loading, and verifies the parametric finite-volume theory's ability to accurately model structural components with curved boundaries. The so-called Eshelby problem, Eshelby [7], is considered in Sec. 3 and solved using the standard and parametric versions of the finite-volume theory, and the results are compared with the finite-element predictions. This comparison demonstrates that the stress concentration effects artificially introduced at the fiber/matrix interface by rectangular subcell approximation are eliminated by the parametric formulation with sufficiently fine discretization. Comparison with the finite-element results demonstrates the parametric theory's competitive advantage from the accuracy and processing time perspectives. Some observations on the predictive capabilities of the standard and parametric versions of the finite-volume theories relative to the finite-element method are made in Sec. 4 based on the results of Sec. 3, which point to new areas of investigation. Finally, Sec. 5 contains a summary and conclusions stemming from this investigation.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received June 6, 2006; final manuscript received December 22, 2006. Review conducted by Robert M. McMeeking.

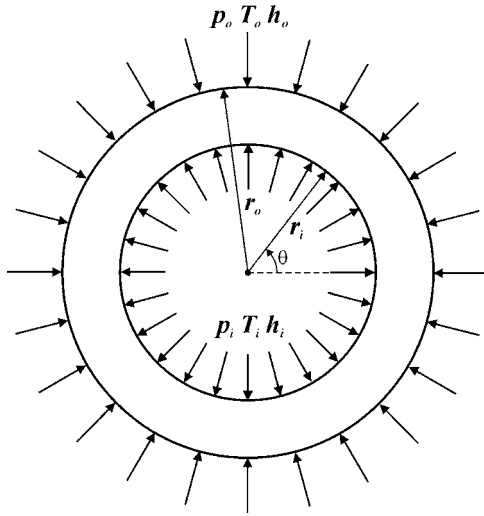


Fig. 1 Thick-walled cylinder subjected to thermomechanical loading

2 Thermomechanical Loading of Thick-Walled Cylinders

We consider the response of homogeneous and graded thick-walled cylinders subjected to steady-state and transient thermomechanical loading by the application of a pressure and temperature differential at the inner and outer surfaces, Fig. 1. The specified loading produces an axisymmetric state of deformation and stress for which one-dimensional analytical solutions are available. These analytical solutions are obtained by first solving the heat conduction equation in polar coordinates

$$\frac{\partial q_r}{\partial r} + \frac{q_r}{r} = -\rho C \frac{\partial T}{\partial t}$$

where

$$q_r = -k \frac{\partial T}{\partial r} \quad (1)$$

to generate the temperature field. The mechanical fields are then obtained by solving the surviving elasticity field equations using either the displacement or stress formulation. In the former approach employed here, the radial displacement field is determined from the solution of the surviving stress equilibrium equations

$$\frac{\partial \sigma_{rr}}{\partial r} + \frac{\sigma_{rr} - \sigma_{\theta\theta}}{r} = 0 \quad (2)$$

expressed in terms of the radial displacement using Hooke's law, which under plane strain condition has the form

$$\begin{aligned} \sigma_{rr} &= \frac{E}{(1+\nu)(1-2\nu)} [(1-\nu)\varepsilon_{rr} + \nu\varepsilon_{\theta\theta}] - \frac{E\alpha}{1-2\nu} T(r) \\ \sigma_{\theta\theta} &= \frac{E}{(1+\nu)(1-2\nu)} [(1-\nu)\varepsilon_{\theta\theta} + \nu\varepsilon_{rr}] - \frac{E\alpha}{1-2\nu} T(r) \end{aligned} \quad (3)$$

where $T(r)$ is the temperature deviation from the reference temperature at which all field variables are taken as zero, and the in-plane strain components ε_{rr} and $\varepsilon_{\theta\theta}$ are related to the radial displacement through the strain-displacement relations,

$$\varepsilon_{rr} = \frac{\partial u}{\partial r} \quad \varepsilon_{\theta\theta} = \frac{u}{r} \quad (4)$$

The thermoelastic moduli in Eqs. (3) vary with the radial coordinate for the case when the cylinder is graded.

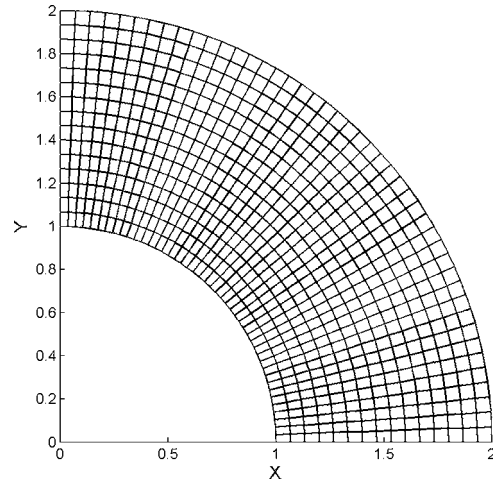


Fig. 2 Discretization of one quarter of a thick-walled cylinder into 15×45 subcells for steady-state analysis of the radially graded and homogeneous cylinder

We test the parametric formulation's accuracy by analyzing the response of one-quarter of the cylinder subject to zero heat flux and zero normal displacement boundary conditions at $\theta=0$ deg and 90 deg applied in the surface-averaged sense at the subcells' external faces. For the steady-state case, combined thermomechanical boundary conditions are applied at the inner and outer radii, while for the transient case purely thermal loading is employed. The quarter-cylinder configuration allows us to rigorously test the theory's predictive capability along different radial planes in order to verify that no circumferential dependence is detected in the predicted field variable distributions. Different discretizations were employed in a numerical convergence study to determine the solution's dependence on mesh refinement. For the analysis under steady-state loading, presented results are based on 15×45 subcell quarter-cylinder discretization with equal partitions along the radial and circumferential directions, respectively, which was employed for both the homogeneous and radially graded configurations, Fig. 2. For the transient analysis involving a homogeneous cylinder, presented results are based on equally-partitioned 60×45 subcell quarter-cylinder discretization (not shown), motivated by the one-dimensional transient analysis of the rectangular plate with through-thickness graded properties discussed by Cavalcante [4]. This case provides a more stringent test than its homogeneous counterpart.

2.1 Graded Cylinder: Steady-State Case. The solution to a radially graded cylinder under the type of loading considered here was obtained by Jabbari et al. [8] when the thermoelastic moduli vary with the radial coordinate in the manner shown below

$$E(r) = E_o r^{m_1} \quad \nu(r) = \nu_o \quad \alpha(r) = \alpha_o r^{m_2} \quad k(r) = k_o r^{m_3} \quad (5)$$

For such power-law dependence, the solution to the heat conduction equation (1), in the presence of convective thermal boundary conditions is obtained in the form

$$T(r) = -\frac{A_1}{m_3} r^{-m_3} + A_2 \quad (6)$$

where the coefficients A_1 and A_2 are given in terms of the cylinder geometry and the convective heat transfer coefficients and temperatures at the inner and outer radii in the Appendix. Use of this temperature field in the surviving Navier's equation leads to the solution for the radial displacement of the form

$$u(r) = B_1 r^{\eta_1} + B_2 r^{\eta_2} + D_1 r^{m_2+1} + D_2 r^{m_2-m_3+1} \quad (7)$$

where

Table 1 Thermal and elastic parameters of the graded cylinder used in the steady-state analysis

Conductivity $k(r)=k_0r^{m_3}$	Thermal expansion $\alpha(r)=\alpha_0r^{m_2}$	Modulus $E(r)=E_0r^{m_1}$	Poisson's ratio $\nu(r)=\nu_0$
$k_0=4.01 \times 10^{-4}$ BTU/in. s °F $m_3=2.0$	$\alpha_0=1.83 \times 10^{-6}$ / °F $m_2=2.0$	$E_0=43.5 \times 10^6$ psi $m_1=2.0$	$\nu_0=0.22$

$$\eta_{1,2} = -\frac{m_1}{2} \mp \sqrt{\frac{m_1^2}{4} - \frac{\nu_0 m_1}{1 - \nu_0} + 1}$$

and the coefficients $B_1, \dots, D_2$ are complicated functions of the geometry, thermoelastic moduli and the thermal coefficients A_1 and A_2 , and are also given in the Appendix. The stress field follows directly upon the use of Eqs. (3) and (4).

The analytical results were generated using thermoelastic material properties listed in Table 1 for the geometry and loading by internal and external pressure and temperature using the coefficients for convective thermal boundary conditions reported in Table 2. These properties are based on a related investigation of a homogeneous cylinder conducted by Arnold et al. [9] using the original unreconstructed higher-order theory of Aboudi et al. [10]. The chosen power-law exponents m_1 , m_2 , and m_3 produce thermoelastic moduli that increase parabolically with increasing radial distance in order to dramatize the effect of grading. The results based on the parametric finite-volume theory were generated by assigning constant thermoelastic properties to each of the 15 cylindrical shells into which the cylinder was discretized in the radial direction, which were obtained by evaluating the properties given in Eq. (5) at the centroid of each shell. The use of 15 cylindrical shells to approximate continuous radial grading by discrete grading was sufficient to generate accurate results for those field variables that are continuous across the subcell interfaces along radial paths as shown in the sequel.

Figure 3 presents temperature and in-plane σ_{rr} and $\sigma_{\theta\theta}$ normal stress distributions obtained from the parametric finite-volume theory along three radial paths oriented at $\theta=0$ deg, 45 deg, and 90 deg with respect to the horizontal axis. As observed, there is no angular dependence of the field quantities, and the predicted results for the temperature and radial stress distributions, which must be continuous along radial paths across adjacent subcell interfaces, are virtually identical with the exact analytical solution. In contrast, discontinuities are observed in the hoop stress distributions at adjacent subcell interfaces which are magnified towards the outer radius where the hoop stress gradient becomes greater. Since hoop stress is not a traction component in the radial direction, these discontinuities are due to the piecewise uniform manner of grading and their magnitude depends on the mesh refinement as observed in the numerical convergence study. Nonetheless, the analytical solution practically coincides with the results of the parametric finite-volume theory at the subcells' midpoints. As an additional check of the solution's accuracy, the in-plane $\sigma_{r\theta}$ shear stress distributions, which should be identically zero, were examined and found to be an insignificant fraction of the normal stress magnitudes (on the order of 0.1%).

In order to demonstrate the effect of grading, the corresponding

results for the homogeneous cylinder under the same thermo-mechanical loading are also included in Fig. 3. These results were obtained by setting the power-law exponents m_1 , m_2 , and m_3 to zero in the expressions for the thermoelastic moduli given in Eq. (5). The resulting properties are the same as those employed in the investigation based on the original unreconstructed higher-order theory by Arnold et al. [9] where substantial stress concentrations caused by the use of rectangular subcells were observed at the outer and inner radii of a similar homogeneous cylinder subjected to similar thermomechanical loading. This issue will be discussed in more detail in Sec. 3 in the context of the inclusion problem.

For this specialized homogeneous cylinder case, the solution of the heat conduction equation produces the temperature distribution of the form

$$T(r) = A_1 \ln(r) + A_2 \quad (8)$$

where the coefficients A_1 and A_2 , obtained upon application of the convective thermal boundary conditions, are

$$A_1 = \frac{T_i - T_o}{\ln(r_i/r_o) - k[1/(r_i h_i) + 1/(r_o h_o)]}$$

$$A_2 = \frac{T_i - T_o}{\ln(r_i/r_o) - k[1/(r_i h_i) + 1/(r_o h_o)]} \left(\frac{k}{r_i h_i} - \ln(r_i) \right) + T_i \quad (9)$$

The solution of the Navier's equation then yields radial displacement of the form, cf. Timoshenko and Goodier [11],

$$u(r) = \frac{(1+\nu)}{(1-\nu)} \frac{1}{r} \int_{r_i}^r T(r') r' dr' + B_1 r + \frac{B_2}{r} \quad (10)$$

from which the stress components are obtained upon use of Eqs. (3) and (4) in the form

$$\sigma_{rr}(r) = - \left(\frac{\alpha E}{1-\nu} \right) \frac{1}{r^2} \int_{r_i}^r T(r') r' dr' + \frac{E}{(1+\nu)} \left(\frac{B_1}{1-2\nu} - \frac{B_2}{r^2} \right)$$

$$\sigma_{\theta\theta}(r) = \left(\frac{\alpha E}{1-\nu} \right) \frac{1}{r^2} \int_{r_i}^r T(r') r' dr' - \frac{\alpha E T(r)}{1-\nu} + \frac{E}{(1+\nu)} \left(\frac{B_1}{1-2\nu} + \frac{B_2}{r^2} \right) \quad (11)$$

Application of the specified boundary conditions yields the coefficients B_1 and B_2 given below

Table 2 Graded cylinder geometry and loading used in the steady-state analysis

Radius r_i	Radius r_o	Temperature T_i	Temperature T_o	Pressure p_i	Pressure p_o
1 in.	2 in.	0 °F	3500 °F	10,000 psi	1000 psi
Convective coefficients					
$h_i=0.0385$ BTU/in. ² s °F					
$h_o=0.0003$ BTU/in. ² s °F					

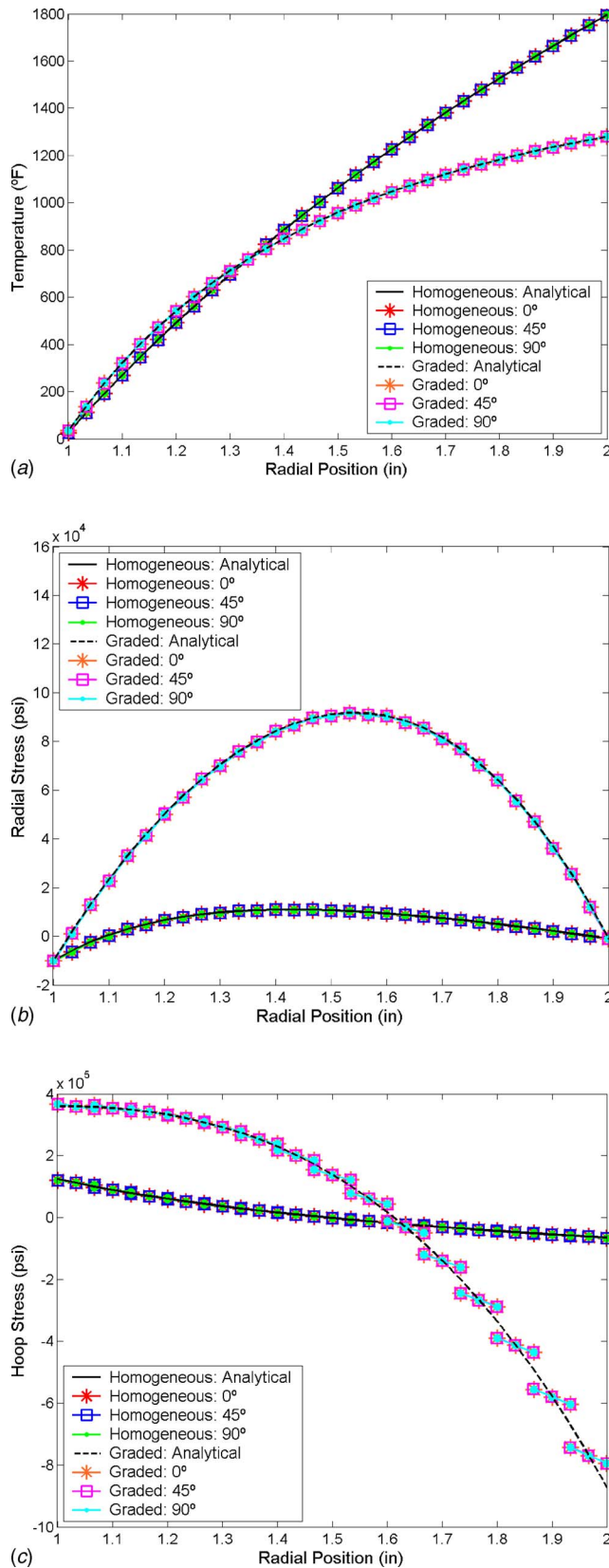


Fig. 3 Comparison of the analytical solution results with the parametric finite-volume predictions for steady-state thermo-mechanical loading of radially graded and homogeneous cylinders. Radial temperature and stress distributions in the three cross sections $\theta=0$ deg, 45 deg, 90 deg. (a) $T(r, \theta=0$ deg, 45 deg, 90 deg) distributions, (b) $\sigma_{rr}(r, \theta=0$ deg, 45 deg, 90 deg) distributions, (c) $\sigma_{\theta\theta}(r, \theta=0$ deg, 45 deg, 90 deg) distributions.

$$B_1 = \frac{(1-2\nu)(1+\nu)\alpha}{1-\nu} \frac{1}{r_o^2 - r_i^2} \int_{r_i}^{r_o} T(r') r' dr' - \frac{(1-2\nu)(1+\nu)}{E} \left(\frac{p_o r_o^2 - p_i r_i^2}{r_o^2 - r_i^2} \right)$$

$$B_2 = \frac{(1+\nu)\alpha}{1-\nu} \frac{r_i^2}{r_o^2 - r_i^2} \int_{r_i}^{r_o} T(r') r' dr' - \frac{(1+\nu)}{E} \left(\frac{p_o - p_i}{r_o^2 - r_i^2} \right) r_i^2 r_o^2$$
(12)

As observed in Fig. 3, the temperature and both stress distributions within the homogeneous cylinder predicted by the parametric finite-volume theory are continuous along radial paths across subcell interfaces due to the absence of material discontinuities. As before, no angular dependence is observed and the agreement with the analytical solution is excellent. Moreover, the stress concentrations at the external boundary subcell corners observed in the results obtained by Arnold et al. [9] have been eliminated by the present parametric mapping approach.

The effect of grading in the chosen manner decreases the temperature at the outer radius at the expense of significant increases in the magnitudes of the σ_{rr} and $\sigma_{\theta\theta}$ stress components. Specifically, the maximum tensile radial stress which occurs in the interior is shifted towards the outer radius and amplified substantially by grading, as are the maximum hoop stress magnitudes which occur at the inner (tensile) and outer (compressive) radii.

2.2 Homogeneous Cylinder: Transient Case. The temperature distribution in a homogeneous cylinder, initially at zero temperature, subjected to a sudden temperature change T_i at the inner radius r_i with the temperature T_o at the outer radius r_o kept at zero is given in closed form in terms of Bessel's functions of first and second type, J_o and Y_o , Carslaw and Jaeger [12],

$$T(r, t) = \frac{\ln(r_o/r)}{\ln(r_o/r_i)} T_i + \pi \sum_{n=1}^{\infty} \frac{J_o(r_o \alpha_n) J_o(r_i \alpha_n)}{J_o^2(r_i \alpha_n) - J_o^2(r_o \alpha_n)} [J_o(r \alpha_n) Y_o(r_o \alpha_n) - J_o(r_o \alpha_n) Y_o(r \alpha_n)] \exp\left(-\frac{k \alpha_n^2 t}{C \rho}\right)$$
(13)

where α_n 's satisfy the characteristic equation

$$J_o(r_i \alpha_n) Y_o(r_o \alpha_n) - J_o(r_o \alpha_n) Y_o(r_i \alpha_n) = 0$$
(14)

Since the thermal problem is uncoupled from the mechanical problem, once the temperature field is specified the in-plane stresses are obtained in the same form as for the steady-state problem given by Eqs. (11), with the temperature field $T(r)$ replaced by $T(r, t)$ in order to account for the transient effect, and the coefficients B_1 and B_2 specialized by setting the internal and external pressures to zero,

$$B_1 = \frac{(1-2\nu)(1+\nu)\alpha}{1-\nu} \frac{1}{r_o^2 - r_i^2} \int_{r_i}^{r_o} T(r', t) r' dr'$$

$$B_2 = \frac{(1+\nu)\alpha}{1-\nu} \frac{r_i^2}{r_o^2 - r_i^2} \int_{r_i}^{r_o} T(r', t) r' dr'$$

The results generated using the parametric finite-volume theory are compared with the analytical predictions based on the above equations in Fig. 4 when the sudden temperature change T_i at the inner radius is one. These results were generated using the dimensionless thermoelastic, geometric and loading parameters given in Tables 3 and 4, and an initial time increment of 5×10^{-5} sec which was geometrically increased by 1.05 at each time step to accelerate solution towards the steady state. The results are shown beginning with the time of 7×10^{-4} sec. As observed, the temperature and circumferential stress fields are reproduced with very

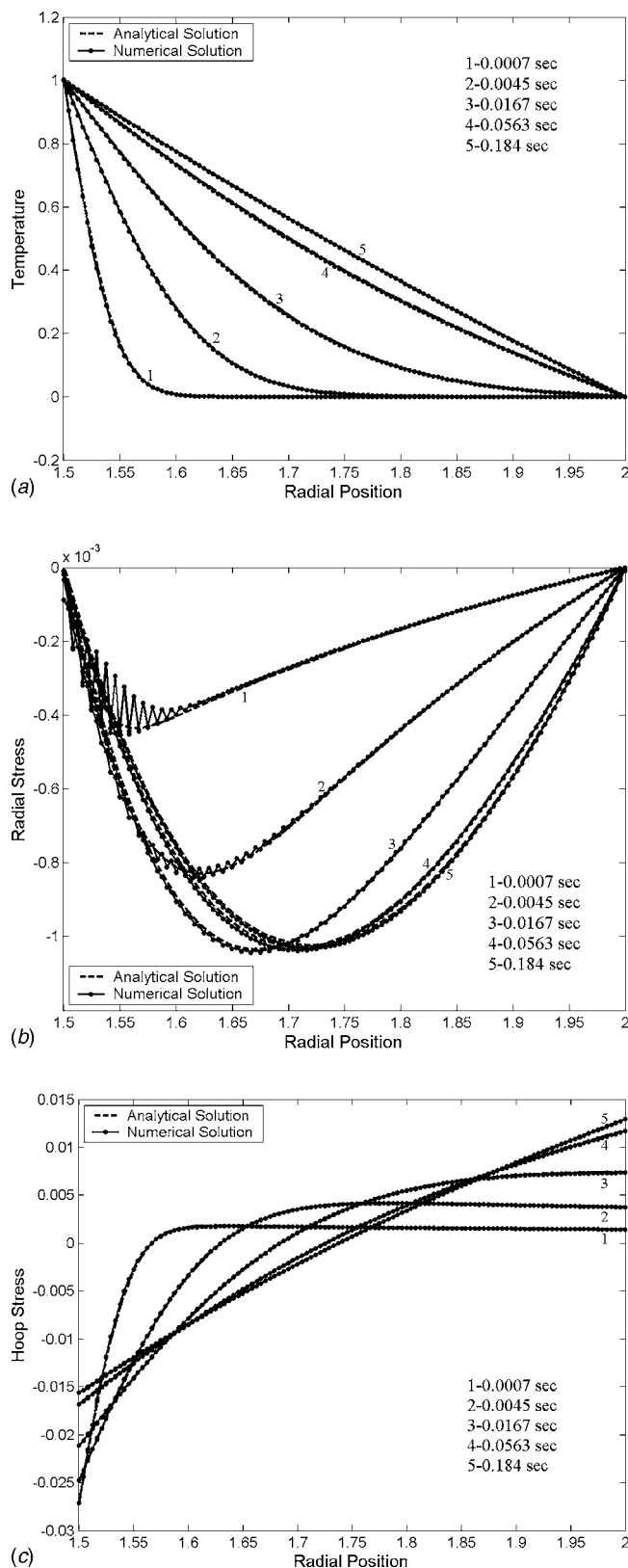


Fig. 4 Comparison of the analytical solution results with the parametric finite-volume predictions for transient thermomechanical loading of a homogeneous cylinder. Radial temperature and stress distributions in the cross section $\theta=45$ deg at different times. (a) $T(r, \theta=45 \text{ deg})$ distributions, (b) $\sigma_{rr}(r, \theta=45 \text{ deg})$ distributions, (c) $\sigma_{\theta\theta}(r, \theta=45 \text{ deg})$ distributions.

good accuracy by the parametric finite-volume theory for all times shown. The same was true for times less than 7×10^{-4} sec. Numerical convergence of the radial stress field, however, is more difficult to achieve during the initial rapid heating of the plate's interior with the given mesh, and particularly in the time range $5 \times 10^{-5} - 7 \times 10^{-4}$ sec. For times greater than 7×10^{-4} sec, some deviations from the analytical solution continue to be initially observed in the radial stress distribution close to the inner radius, which however do disappear with time. The oscillatory character of this stress distribution in the vicinity of the inner radius can be eliminated by increasing the number of subcells in this region as was verified (not shown). However, it should be noted that the radial stress is one order of magnitude lower than the hoop stress which is reproduced with excellent accuracy by the parametric finite-volume theory. It is also worth noting that despite the large gradients observed in the radial stress distributions, the traction-free boundary condition is satisfied at the outer cylinder radius at all times. At the inner radius, this condition is satisfied for times greater than 7×10^{-4} .

3 The Inclusion Problem

The parametric finite-volume theory enables more accurate modeling of different inclusion types of heterogeneous microstructures with fewer subcells than the standard finite-volume theory based on rectangular subcell discretization. Superior results are obtained due to the elimination of stress concentrations at subcell corners that form the fiber/matrix interface, which are introduced by the artificial modeling of curved boundaries using rectangular discretization. Moreover, increasing the mesh refinement produces results that approach those obtained from the finite-element analysis, including satisfaction of traction and displacement continuity at curved interfaces separating regions of large elastic property contrast.

These features are illustrated in this section by considering the classical problem of an elastic fiber embedded in a large matrix subjected to uniform far-field stress $\sigma_{xx}^0 = 100$ MPa along the horizontal direction, Fig. 5. When the matrix is infinite, exact analytical solution for this Eshelby problem is obtained in simple form, cf. Dugdale and Ruiz [13]. In our case, the fiber radius is 1 and the matrix region has dimensions of 20×20 which mimic the Eshelby problem. To illustrate the above features, we use two mesh discretizations based on rectangular subcells for use in the standard finite-volume theory analysis, and three discretizations based on quadrilateral subcells for use in the parametric version and the finite-element analysis. The elastic properties for the fiber and matrix phases used in the analyses based on these discretizations are given in Table 5.

The exact analytical solution to this problem serves as the gold standard against which the standard and parametric finite-volume and finite-element results are compared. It is given in compact form in terms of two sets of complex potentials ϕ and ψ , with each pair corresponding to the fiber and matrix phases. The stresses in the x - y plane are obtained from the formulas

$$\sigma_{xx} = 2 \operatorname{Re}(\phi') - \operatorname{Re}(\bar{z}\phi'' + \psi') \quad \sigma_{yy} = 2 \operatorname{Re}(\phi') + \operatorname{Re}(\bar{z}\phi'' + \psi')$$

$$\sigma_{xy} = \operatorname{Im}(\bar{z}\phi'' + \psi') \quad (15)$$

where the prime denotes the derivative with respect to the complex variable $z=x+iy$, and Re and Im denote the real and imaginary parts of the expressions within the parentheses, respectively. For the fiber phase, the two complex potentials are given by

$$\phi_f = a_1 z \quad \psi_f = p_1 z \quad (16)$$

and for the matrix phase they are

$$\phi_m = \tilde{a}_{-1} z^{-1} + \tilde{a}_1 z \quad \psi_m = \tilde{p}_{-3} z^{-3} + \tilde{p}_{-1} z^{-1} + \tilde{p}_1 z \quad (17)$$

Application of the interfacial displacement and traction continuity conditions, and the far-field boundary conditions gives the follow-

Table 3 Dimensionless thermal and elastic parameters of the homogeneous cylinder used in the transient analysis

Conductivity	Heat capacity	Density	Thermal expansion	Modulus	Poisson's ratio
1.0	1.0	1.0	0.02	1.0	0.3

Table 4 Dimensionless geometric and loading parameters used in the transient thermal analysis of the homogeneous cylinder

Radius r_i	Radius r_o	Temperature T_i	Temperature T_o
1.5	2.0	1	0

ing expressions for the coefficients appearing in the above equations in the case of plane strain loading ($\bar{\epsilon}_{zz}=0$)

$$a_1 = \frac{1}{4} \sigma_{xx}^{\infty} \frac{(\kappa_m + 1) \mu_f}{[2\mu_f + (\kappa_f - 1) \mu_m]}$$

$$p_1 = -\frac{1}{2} \sigma_{xx}^{\infty} \left[1 + \frac{\mu_f - \mu_m}{\mu_m + \kappa_m \mu_f} \right] \quad (18)$$

$$\tilde{a}_{-1} = -\frac{1}{2} \sigma_{xx}^{\infty} \left[\frac{\mu_f - \mu_m}{\mu_m + \kappa_m \mu_f} \right] \quad \tilde{a}_1 = \frac{1}{4} \sigma_{xx}^{\infty} \quad \tilde{p}_{-3} = \tilde{a}_{-1}$$

$$\tilde{p}_{-1} = 2a_1 - \frac{1}{2} \sigma_{xx}^{\infty} \quad \tilde{p}_1 = -\frac{1}{2} \sigma_{xx}^{\infty} \quad (19)$$

where the subscripts f and m denote fiber and matrix phases, respectively, and the corresponding κ 's are related to the Lamé's constants λ and μ as follows:

$$\kappa = \frac{\lambda + 3\mu}{\lambda + \mu}$$

and $\sigma_{xx}^{\infty} = \sigma_{xx}^o = 100$ MPa. Figure 5 includes the stress distributions obtained from the above solution for the current problem. As ob-

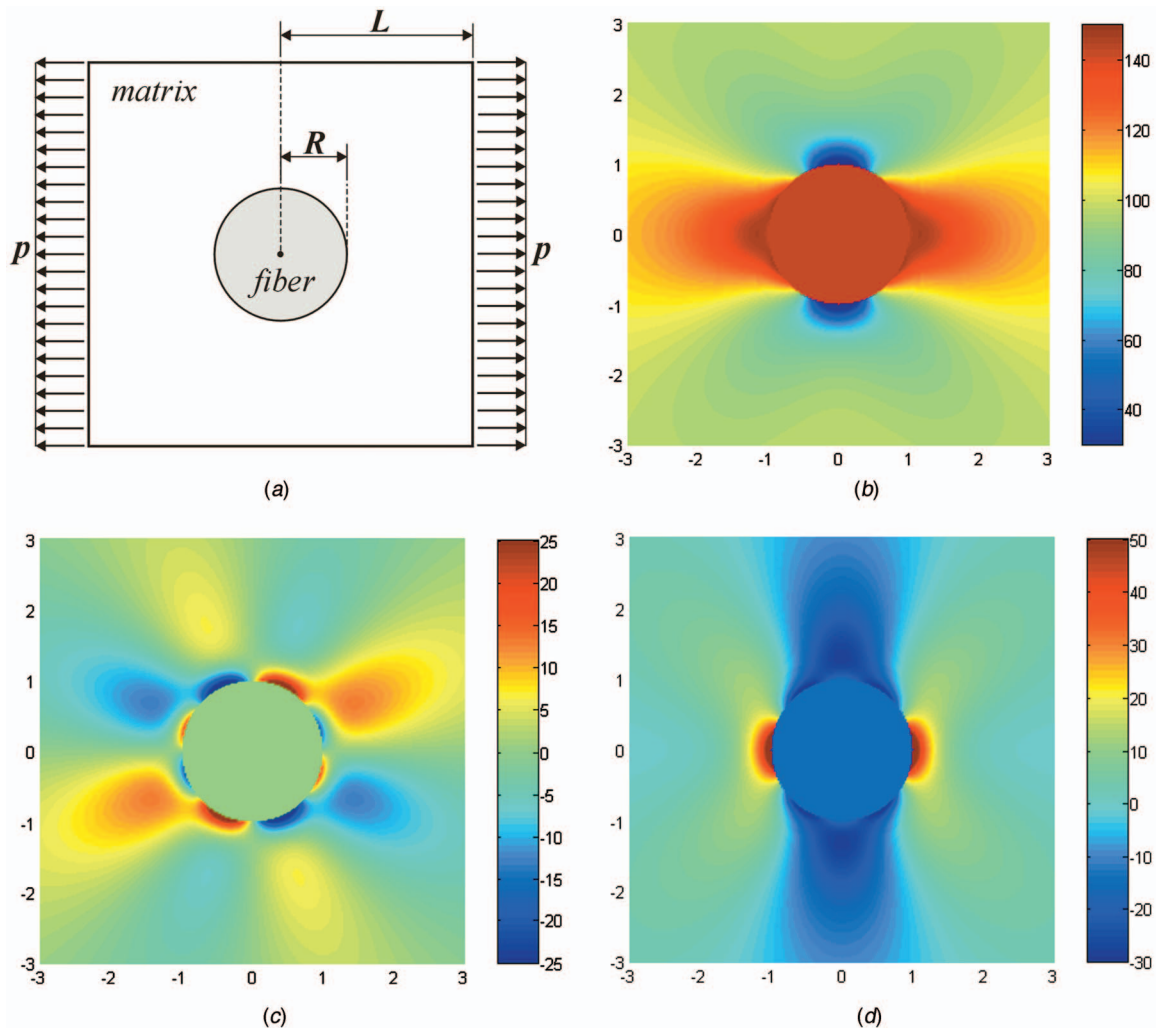


Fig. 5 A circular fiber of radius 1 embedded in a large matrix of dimensions 20×20 subjected to uniform far-field stress $\sigma_{xx}^o = 100$ MPa and the stress fields obtained from the exact analytical solution for the corresponding Eshelby problem. (a) A circular fiber in large matrix, (b) $\sigma_{xx}(x, y)$ distribution, (c) $\sigma_{xy}(x, y)$ distribution, (d) $\sigma_{yy}(x, y)$ distribution.

Table 5 Fiber and matrix elastic properties for the Eshelby problem

Material	E (GPa)	ν
Glass fiber	69.0	0.2
Epoxy matrix	4.8	0.34

served, the normal stresses σ_{xx} and σ_{yy} in the fiber are constant while the shear stress σ_{xy} vanishes. In the matrix, the stress field is quite complex with all stress components present, but approaches the uniform far-field stress σ_{xx}^o with increasing distance from the fiber.

First, we compare the predictions of the standard and parametric versions of the finite-volume theory for comparable mesh discretizations relative to the exact solution. Figure 6 compares the rectangular and quadrilateral subcell discretizations of the inner 2×2 region containing the fiber using the same number of divisions along the horizontal and vertical axes passing through the fiber center. For the same level of fiber discretization, 4×625 rectangular subcells were employed for the standard finite-volume analysis of the entire region, in contrast with just 4×325 quadrilateral subcells required for the parametric analysis. However, because of symmetry, the analyses were performed on one quarter of

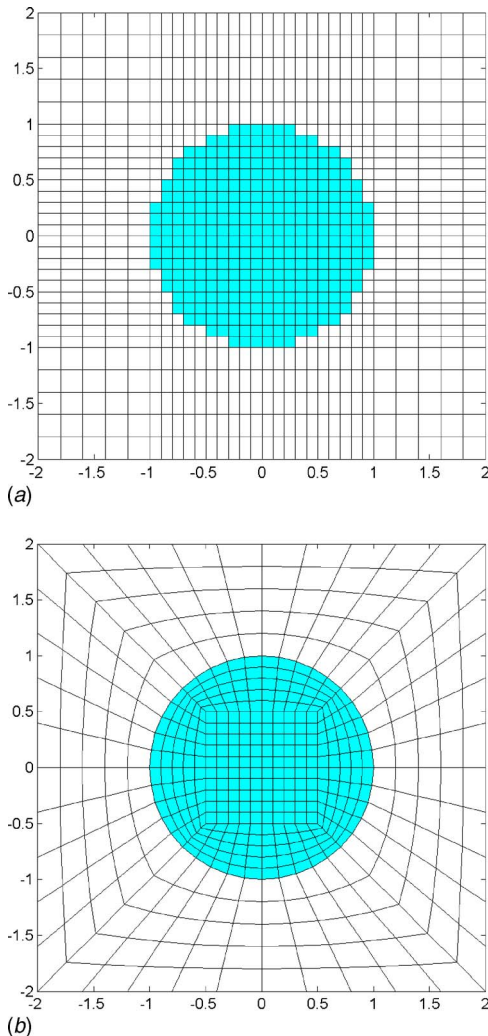


Fig. 6 Discretizations of the 2×2 region containing centered fiber used in the standard (a) 4×625 discretization, and parametric (b) 4×325 discretization finite-volume analyses

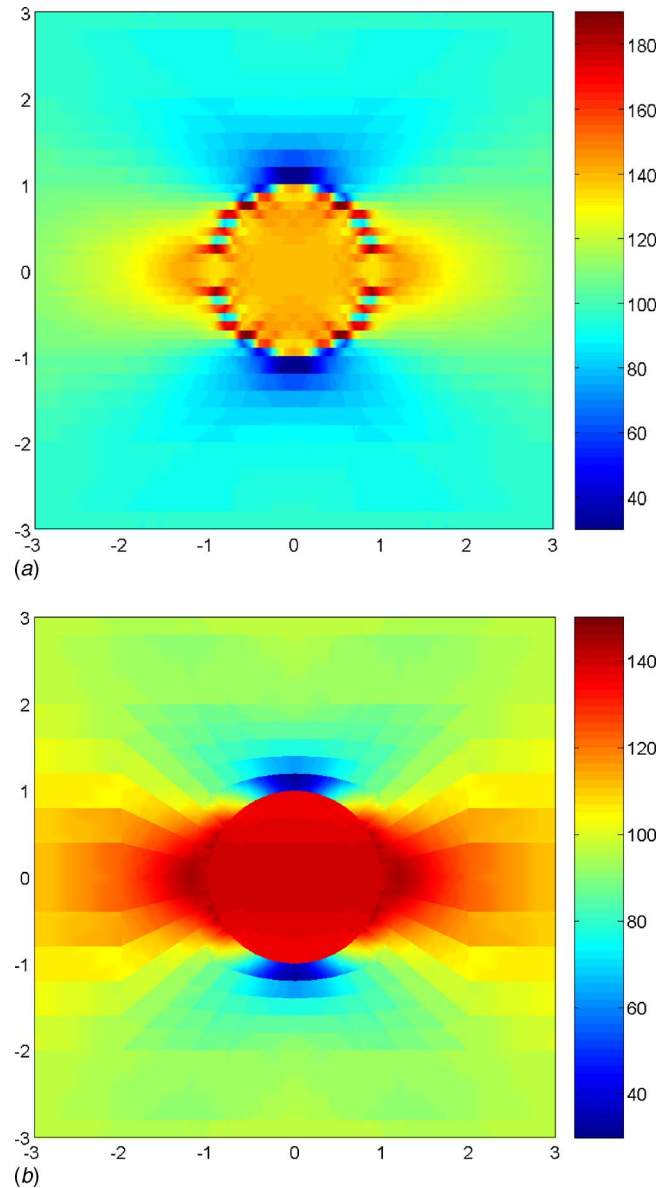


Fig. 7 Local $\sigma_{xx}(x,y)$ stress distributions in the 3×3 region containing the centered fiber predicted by the standard and parametric finite-volume theories based on $4 \times (25 \times 25) = 4 \times 625$ rectangular subcell and 4×325 quadrilateral subcell discretizations, respectively. (a) Standard finite-volume theory, (b) parametric finite-volume theory.

the entire region with appropriate symmetry conditions applied along horizontal and vertical axes passing through the fiber center. The predicted $\sigma_{xx}(x,y)$ and $\sigma_{xy}(x,y)$ stress distributions in the 3×3 region containing the centered fiber are compared in Fig. 7 and Fig. 8, respectively. Note the different color bar scales used for the $\sigma_{xx}(x,y)$ distributions in this figure, and in Fig. 5 showing the exact elasticity solution, that highlight the stress concentrations caused by the square subcells along the fiber/matrix interface. Clearly, even at this level of discretization these stress concentrations evident in the standard finite-volume predictions are eliminated through the use of quadrilateral subcells in the parametric version. The uniform normal and shear stress fields within the fiber observed in the Eshelby solution results are well captured by the parametric finite-volume predictions, as are the general stress patterns in the fiber's vicinity. It could be argued, however, that the standard finite-volume theory captures the exterior field

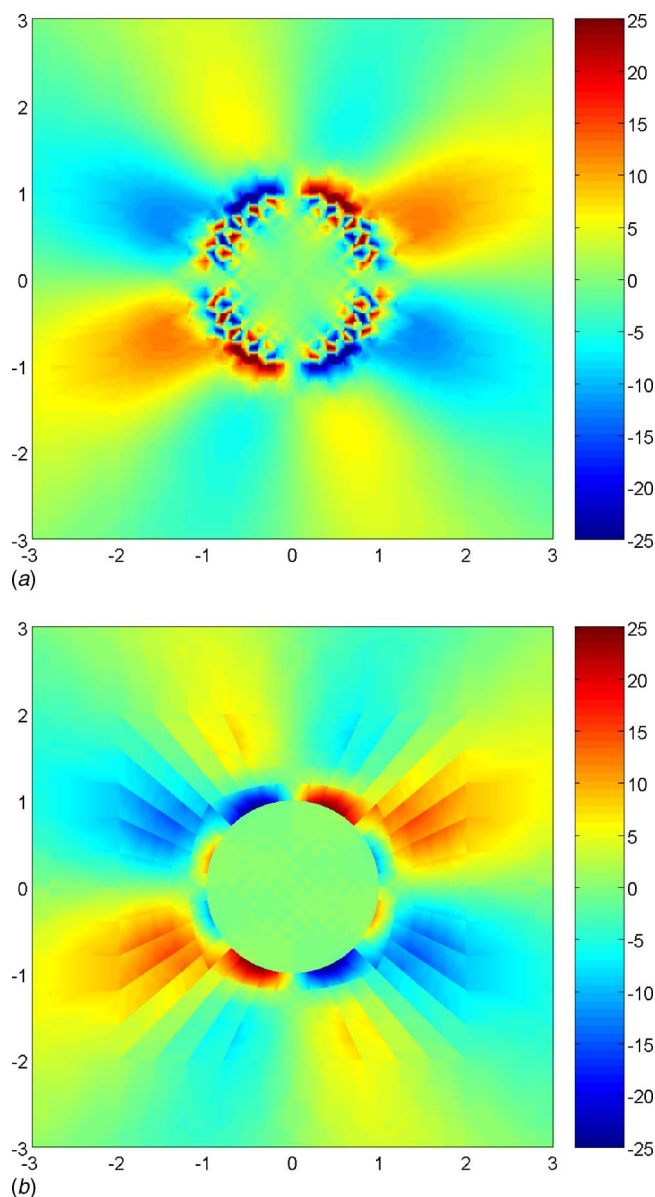


Fig. 8 Local $\sigma_{xy}(x,y)$ stress distributions in the 3×3 region containing the centered fiber predicted by the standard and parametric finite-volume theories based on $4 \times (25 \times 25) = 4 \times 625$ rectangular subcell and 4×325 quadrilateral subcell discretizations, respectively. (a) Standard finite-volume theory, (b) parametric finite-volume theory.

sufficiently removed from the fiber/matrix interface at least as well (if not better) as the parametric version even at this level of discretization. The subcell geometry outside of the fiber does influence the resulting stress distributions where a bias is observed in the parametric version results which produces traction discontinuities in the circumferential direction in contrast to the (nearly) uniform interior fields. The traction discontinuities observed in the exterior region become smaller with increasing mesh refinement as demonstrated in the sequel.

Next, we demonstrate that the extent of stress concentrations caused by the modeling of curved interfaces using rectangular subcells can be reduced with increasing mesh refinement, but not fully eliminated. Figure 9(a) illustrates a detail of the mesh in the inner 2×2 region containing the fiber employed to demonstrate this effect. The entire mesh in this study consisted of $4 \times (75 \times 75) = 4 \times 5625$ subcells. The predicted $\sigma_{xx}(x,y)$, $\sigma_{xy}(x,y)$, and

$\sigma_{yy}(x,y)$ stress distributions in the 3×3 region containing the centered fiber are presented in Figs. 9(b)–9(d). As observed, while the magnitude of stress concentrations has been slightly reduced, the extent of these concentrations has been substantially minimized producing nearly uniform stress fields within the fiber. We note that the use of nonuniformly dimensioned subcells that better mimic the circular fiber shape for the same overall subcell discretization of the entire region reduces these stress concentrations, Bansal and Pindera [1].

Finally, we compare stress fields predicted by the parametric finite-volume theory with the finite-element results. Three additional meshes were constructed using 4×1200 , 4×2600 , and 4×5700 quadrilateral subcells for the entire region. The same meshes were employed in the finite-element analysis using Q4 and Q8 elements with 2 and 6 degrees of freedom per element, respectively. This contrasts with 4 degrees of freedom per quadrilateral subcell. Figure 10 illustrates the quadrilateral subcell/element discretizations of the inner 2×2 region containing the centered fiber taken from 4×1200 and 4×2600 subcell/element meshes used to discretize the entire 20×20 region. Herein, we compare finite-element results generated with the 4×1200 Q8-element mesh with parametric finite-volume results obtained from the 4×2600 subcell mesh. The 4×1200 Q8-element mesh produced converged results as verified upon comparison with the results produced by the two finer 4×2600 and 4×5700 element meshes. Use of the Q4 elements in the 4×1200 element mesh produced inferior results suffering from discontinuities in the surrounding matrix. Increasing the mesh size to 4×5700 elements with the Q4 elements produced results which were not as good as those obtained from the 4×1200 Q8-element mesh.

This comparison is presented in Figs. 11–13 for the local stress distributions $\sigma_{xx}(x,y)$, $\sigma_{xy}(x,y)$, and $\sigma_{yy}(x,y)$, respectively, in the 3×3 region containing the centered fiber. These results were plotted in MATLAB using data generated by the finite-element and parametric finite-volume approaches at 5×5 and 3×3 equally-spaced points in each element/subcell of the 4×1200 element and 4×2600 subcell meshes, respectively. As observed, comparable results are obtained from both analyses. Increasing the mesh size to 4×5700 subcells does improve the quality of finite-volume results at the expense of a threefold increase in the execution time. These execution times compare very favorably with the execution time recorded for the 4×1200 Q8-element mesh which was approximately fifteen-fold and five-fold longer, respectively. Both the finite-volume and finite-element codes were written in MATLAB based on the same matrix solver, and executed on the same 2.79 GHz machine with 512 Mb of RAM for direct comparison.

Comparison of execution times for the different meshes used in generating the standard and parametric finite-volume theory results with the corresponding finite-element execution times is tentative, at present, because research codes not optimized for speed were employed in the calculations. In both the finite-volume and finite-element cases, only the assembly of the global matrices and the solution of the systems of equations was included in the execution time comparison. Most of the time was spent in assembling the global matrices for the two methods (greater than 90% for the investigated discretizations), which nonetheless does demonstrate the finite-volume theory's competitive potential.

The implemented finite-element procedure employed complete integration in the construction of local stiffness matrices (4 Gauss points for the Q4 elements and 9 Gauss points for the Q8 elements, each integration procedure with its respective Gaussian weights). In the case of the parametric finite-volume approach, the computation of subcell local stiffness matrix is much faster because no integration is necessary owing to the direct enforcement of the equilibrium equations in the large, or in the volume-average sense. Therefore, the calculation of the local stiffness matrix could be interpreted in the finite-element sense as if there was only one Gauss point, providing a partial explanation for the observed ex-

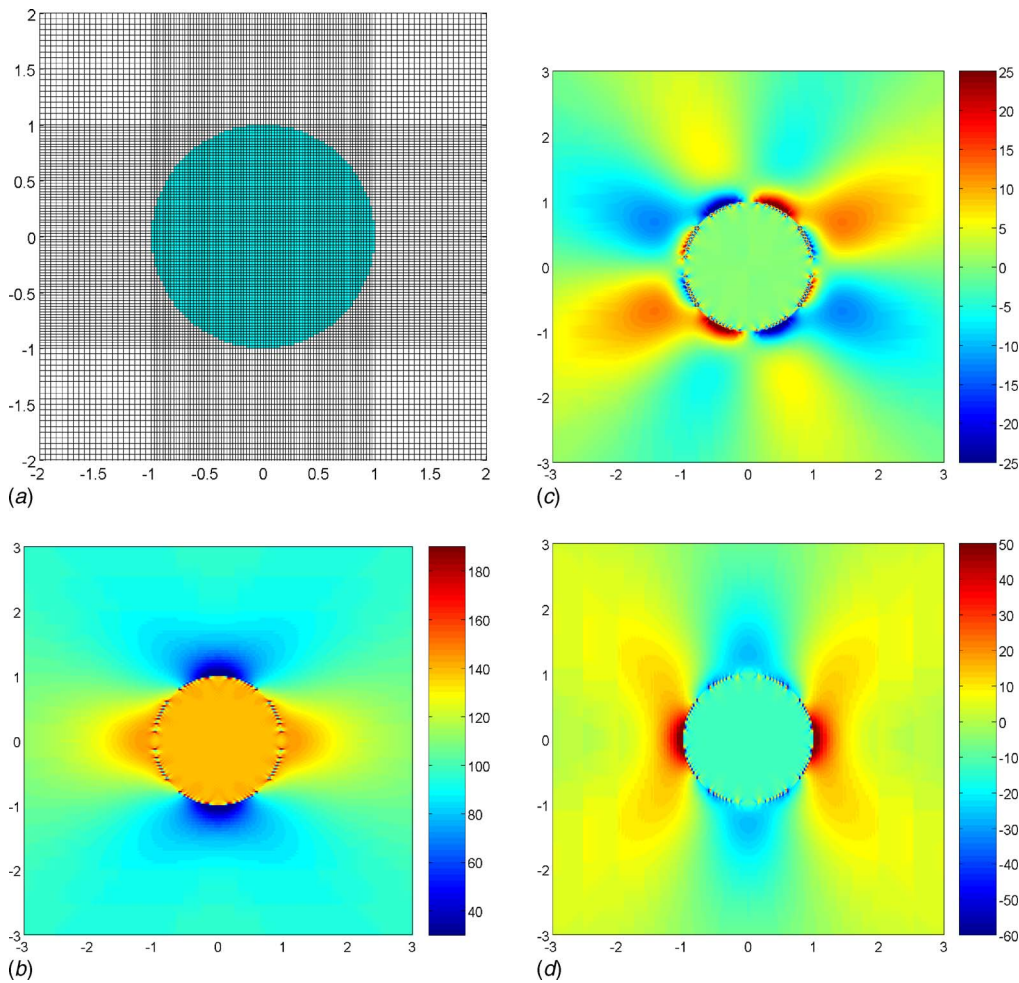


Fig. 9 Very fine discretization of the 2×2 region containing centered fiber taken from the $4 \times (75 \times 75) = 4 \times 5625$ rectangular subcell mesh used in the standard finite-volume analysis, and the resulting stress distributions in the 3×3 region. (a) 4×5625 discretization, (b) $\sigma_{xx}(x, y)$ distribution, (c) $\sigma_{xy}(x, y)$ distribution, (d) $\sigma_{yy}(x, y)$ distribution.

ecution time differences in favor of the finite-volume theory. Clearly, this requires further investigation in light of the finite-volume theory's computational promise.

4 Discussion

The parametric finite-volume predictions of the stress fields for the Eshelby problem presented in Sec. 3 exhibit slight discontinuities in some stress components in the matrix phase away from the fiber/matrix interface even when a fine mesh, see Fig. 10(b), is employed. These discontinuities have directional dependence. Specifically, small discontinuities are observed in the σ_{xx} stress component along the y direction, while the σ_{yy} and σ_{xy} stress components exhibit discontinuities along the x and θ directions, respectively. In contrast, the standard finite-volume theory produces smoother results with increasing mesh refinement except in the close vicinity of the material discontinuity (fiber/matrix interface).

The finite-volume theory is based on the satisfaction of interfacial continuity conditions in a surface-average sense, and therefore discontinuities may be present in those stress components that are not traction components. In the standard finite-volume theory, the discontinuities in the nontraction stress components tend to become very small with increasing mesh refinement away from the fiber/matrix interface, upon comparing Fig. 7(a) and Fig. 9(b) for the σ_{xx} stress component, and Fig. 8(a) and Fig. 9(c) for the σ_{xy} stress component. In this version of the finite-volume

theory, subcell discretization parallels external boundaries of the analyzed structure, and the external loading is applied in the same coordinate system. Traction components are satisfied in a surface-averaged sense across subcell faces perpendicular to each row and column of subcells spanning the entire structure. Therefore in the parametric version, the local subcell orientation must affect stress distributions referred to coordinate systems that are not aligned with subcell faces along whose surfaces the traction components are satisfied in a surface-average sense. Thus one cause of the observed small discontinuities may be the radially-oriented subcell mesh with individual subcells rotated relative to the global Cartesian coordinate system, which are not aligned with the external load. Another cause of these discontinuities may be the approximation of the Jacobian of the coordinate transformation employed in the subcell mapping by its volume average. We leave the investigation of these separate effects for the future.

5 Summary and Conclusions

The parametric formulation of the finite-volume theory for functionally graded materials developed in Part I of this contribution was verified through comparison with exact analytical and finite-element solutions for transient and steady-state thermal and mechanical response of homogeneous and graded thick-walled cylinders, as well as an heterogeneous medium in the form of a circular fiber embedded in a large square matrix. These comparisons have demonstrated that the parametric version can be used to

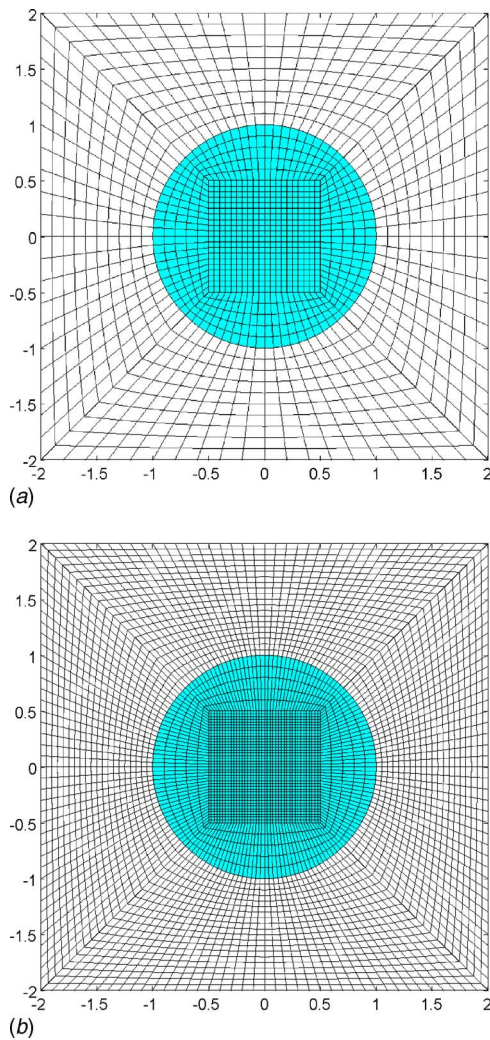


Fig. 10 Discretizations of the 2×2 region containing centered fiber used in the finite-element (a) 4×1200 Q8-element mesh and parametric finite-volume (b) 4×2600 subcell discretization analyses

model quite successfully homogeneous or layered structural components with curved boundaries by mapping a reference square subcell onto a quadrilateral subcell resident in the actual material/structural component. Microstructural details with curved boundaries previously approximated using rectangular subcells are also modeled more efficiently. In particular, the stress concentrations produced by rectangular subcell discretization of a curved material interface were eliminated by the use of quadrilateral subcells which mimic well circular or curved boundaries.

Comparison of the parametric finite-volume theory's predictions with the corresponding finite-element results demonstrated that results of comparable accuracy can be generated with substantially smaller execution times. This is due to the significantly smaller overhead required to calculate local stiffness matrices using the finite-volume theory intrinsically rooted in the method's direct averaging approach of satisfying the equilibrium equations in the large in the course of local stiffness or conductivity matrix calculation.

While the parametric finite-volume theory is a significant step in the evolution of this particular approach for the analysis of materials/structural components with heterogeneous microstructures, a number of questions remain to be addressed as discussed in Sec. 4. Nonetheless, the present development now places this theory solidly among other approaches that have been employed

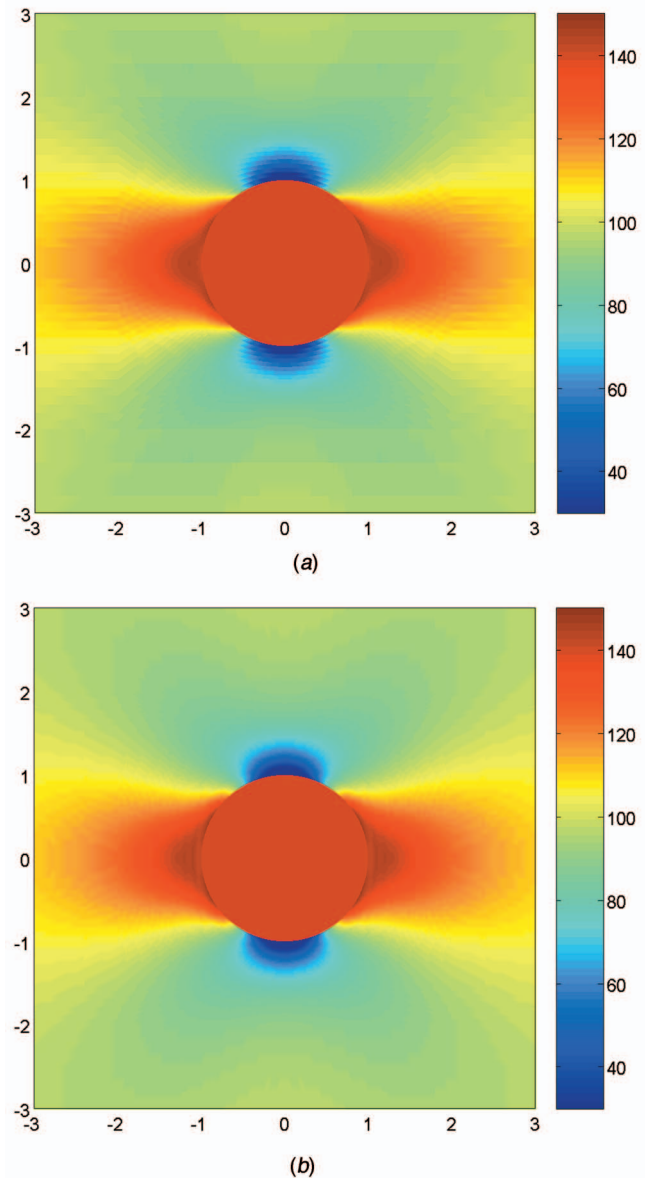


Fig. 11 Local $\sigma_{xx}(x,y)$ stress distributions in the 3×3 region containing centered fiber. Comparison of the parametric finite-volume results based on the 4×2600 subcell discretization with the finite-element results based on the 4×1200 Q8-element mesh. (a) Parametric finite-volume theory, (b) finite-element method.

in the analysis of graded materials vis-à-vis accuracy of local stress fields and flexibility of modeling actual heterogeneous microstructures, such as the finite-element methods with homogeneous and graded elements, cf. Pagano and Yuan [14], Lipton [15], and Kim and Paulino [16], and boundary-element methods, Sutradhar et al. [5]. The results presented herein suggest that with further development the finite-volume theory has the potential to become a premier computational tool for efficient and accurate analysis of both local and global responses of heterogeneous materials.

Acknowledgment

M.A.A.C. and S.P.C.M. would like to gratefully acknowledge the financial support of the Brazilian federal agencies CNPq/PADCT and CAPES. M.-J.P. acknowledges the support of the Engineered Materials Concepts, LLC.

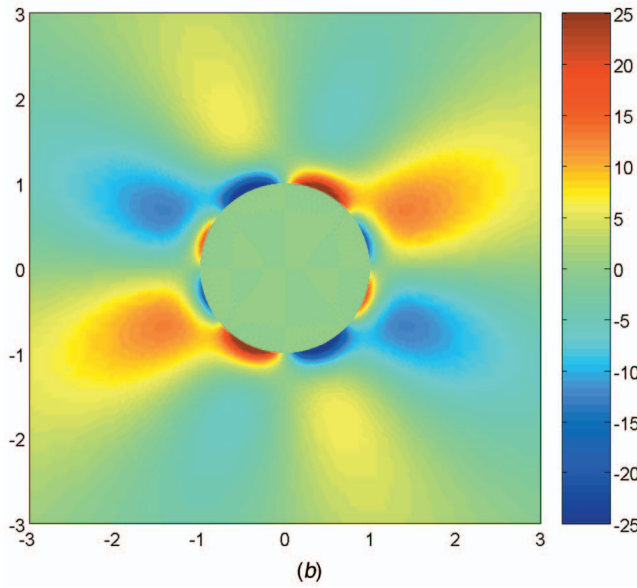
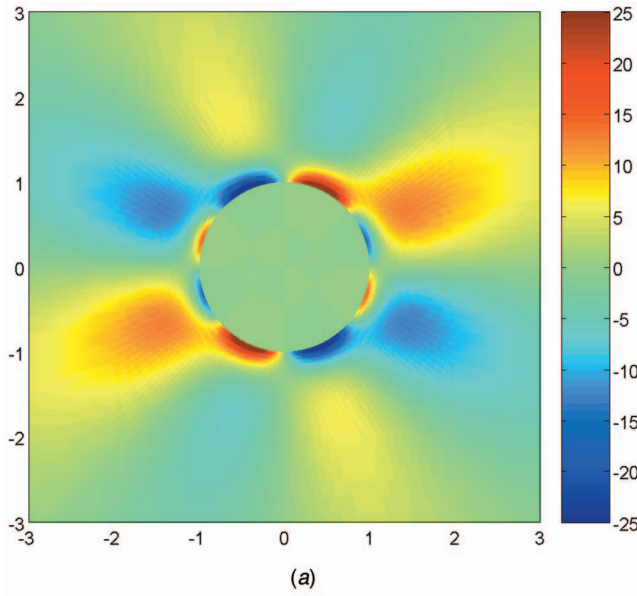


Fig. 12 Local $\sigma_{xy}(x,y)$ stress distributions in the 3×3 region containing centered fiber. Comparison of the parametric finite-volume results based on the 4×2600 subcell discretization with the finite-element results based on the 4×1200 Q8-element mesh. (a) Parametric finite-volume theory, (b) finite-element method.

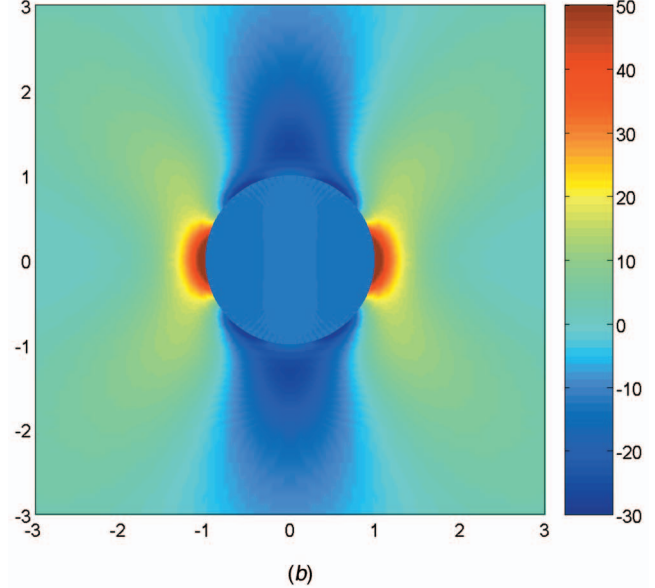
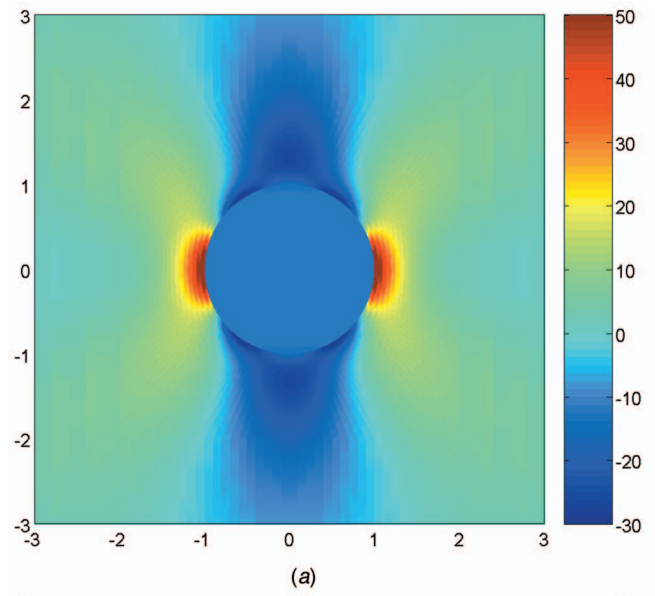


Fig. 13 Local $\sigma_{yy}(x,y)$ stress distributions in the 3×3 region containing the centered fiber. Comparison of the parametric finite-volume results based on the 4×2600 subcell discretization with the finite-element results based on the 4×1200 Q8-element mesh. (a) Parametric finite-volume theory, (b) finite-element method.

Appendix

For the graded cylinder under steady-state loading analyzed in Sec. 2.1, the coefficients A_1 and A_2 appearing in the expression for the temperature field in Eq. (6) are given by

$$A_1 = \frac{h_i h_o (T_o - T_i)}{h_o (k_i r_i^{-m_3-1} + h_i r_i^{-m_3}/m_3) + h_i (k_o r_o^{-m_3-1} - h_o r_o^{-m_3}/m_3)}$$

$$A_2 = \frac{h_o (k_i r_i^{-m_3-1} + h_i r_i^{-m_3}/m_3) T_o - h_o (k_o r_o^{-m_3-1} - h_o r_o^{-m_3}/m_3) T_o}{h_o (k_i r_i^{-m_3-1} + h_i r_i^{-m_3}/m_3) + h_i (k_o r_o^{-m_3-1} - h_o r_o^{-m_3}/m_3)}$$

(A1)

Similarly, the coefficients B_1 , B_2 , D_1 , and D_2 appearing in the expression for the radial displacement in Eq. (7) are given by

$$B_1 = \frac{d_4 d_5 - d_2 d_6}{d_1 d_4 - d_2 d_3}$$

$$B_2 = \frac{d_1 d_6 - d_3 d_5}{d_1 d_4 - d_2 d_3} \quad (A2)$$

where

$$d_1 = [(1 - \nu_o) \eta_1 + \nu_o] r_i^{\eta_1 + m_1 - 1}$$

$$d_2 = [(1 - \nu_o) \eta_2 + \nu_o] r_i^{\eta_2 + m_1 - 1}$$

$$d_3 = [(1 - \nu_o) \eta_1 + \nu_o] r_o^{\eta_1 + m_1 - 1}$$

$$d_4 = [(1 - \nu_o) \eta_2 + \nu_o] r_o^{\eta_2 + m_1 - 1}$$

$$\begin{aligned}
d_5 = & -\frac{(1+\nu_o)(1-2\nu_o)}{E_o} p_i - \{[(1-\nu_o)m_2 + 1]D_1 \\
& - (1+\nu_o)\alpha_o A_2\} r_i^{m_1+m_2} - \{[(1-\nu_o)(m_2-m_3) + 1]D_2 \\
& + (1+\nu_o)\alpha_o A_1/m_3\} r_i^{m_1+m_2-m_3} \\
d_6 = & -\frac{(1+\nu_o)(1-2\nu_o)}{E_o} p_o - \{[(1-\nu_o)m_2 + 1]D_1 \\
& - (1+\nu_o)\alpha_o A_2\} r_o^{m_1+m_2} - \{[(1-\nu_o)(m_2-m_3) + 1]D_2 \\
& + (1+\nu_o)\alpha_o A_1/m_3\} r_o^{m_1+m_2-m_3}
\end{aligned} \tag{A3}$$

and

$$\begin{aligned}
D_1 = & \left[\frac{(1+\nu_o)(m_1+m_2)\alpha_o}{(m_2+1)m_2 + (m_2+1)(m_1+1) + \nu_o m_1/(1-\nu_o) - 1} \right] \frac{A_2}{(1-\nu_o)} \\
D_2 = & \left[\frac{(1+\nu_o)(m_3-m_2-m_1)\alpha_o}{(m_2-m_3+1)(m_2-m_3) + (m_2-m_3+1)(m_1+1) + \nu_o m_1/(1-\nu_o) - 1} \right] \frac{A_1}{(1-\nu_o)m_3}
\end{aligned} \tag{A4}$$

References

- [1] Bansal, Y., and Pindera, M.-J., 2003, "Efficient Reformulation of the Thermoelastic Higher-Order Theory for FGMs," *J. Therm. Stresses*, **26**(11/12), pp. 1055–1092; see also NASA CR 2002-211909, November 2002.
- [2] Zhong, Y., Bansal, Y., and Pindera, M.-J., 2004, "Efficient Reformulation of the Thermal Higher-Order Theory for FGM's With Variable Thermal Conductivity," *Int. J. Comput. Eng. Sci.*, **5**(4), pp. 795–831; see also NASA CR 2002-211910, November 2002.
- [3] Pindera, M.-J., Bansal, Y., and Zhong, Y., 2004, "Finite-Volume Direct Averaging Theory for Functionally Graded Materials (FVDAT-FGM)," NASA Disclosure of Invention and New Technology Form 1679, September, 14, 2004.
- [4] Cavalcante, M. A. A., 2006, "Modeling of the Transient Thermo-Mechanical Behavior of Composite Material Structures by the Finite-Volume Theory," M.S. thesis, Civil Engineering Department, Federal University of Alagoas, Maceio, Alagoas, Brazil.
- [5] Sutradhar, A., Paulino, G. H., and Gray, L. J., 2002, "Transient Heat Conduction in Homogeneous and Non-Homogeneous Materials by the Laplace Transform Galerkin Boundary Element Method," *Eng. Anal. Boundary Elem.*, **26**, pp. 119–132.
- [6] Paulino, G. H., 2002, "Fracture of Functionally Graded Materials," *Eng. Fract. Mech.*, **69**(14–16), pp. 1519–1812.
- [7] Eshelby, J. D., 1957, "The Determination of the Elastic Field of an Ellipsoidal Inclusion, and Related Problems," *Proc. R. Soc. London, Ser. A*, **241**, pp. 376–396.
- [8] Jabbari, M., Sohrabpour, S., and Eslami, M. R., 2002, "Mechanical and Thermal Stresses in a Functionally Graded Hollow Cylinder Due to Radially Symmetric Loads," *Int. J. Pressure Vessels Piping*, **79**, pp. 493–497.
- [9] Arnold, S. M., Bednarczyk, B. A., and Aboudi, J., 2004, "Analysis of Internally Cooled Structures Using a Higher-Order Theory," *Comput. Struct.*, **26**, pp. 659–688.
- [10] Aboudi, J., Pindera, M.-J., and Arnold, S. M., 1999, "Higher-Order Theory for Functionally Graded Materials," *Composites, Part B*, **30**(8), pp. 777–832.
- [11] Timoshenko, S. P., and Goodier, J. N., 1951, *Theory of Elasticity*, McGraw-Hill, New York.
- [12] Carslaw, H. S., and Jaeger, J. C., 1959, *Conduction of Heat in Solids*, Clarendon, Oxford.
- [13] Dugdale, D. S., and Ruiz, C., 1971, *Elasticity for Engineers*, McGraw-Hill, London.
- [14] Pagano, N. J., and Yuan, F. G., 2000, "The Significance of Effective Modulus Theory (Homogenization) in Composite Laminate Mechanics," *Compos. Sci. Technol.*, **60**, pp. 2471–2488.
- [15] Lipton, R. P., 2003, "Assessment of the Local Stress State Through Macroscopic Variables," *Philos. Trans. R. Soc. London, Ser. A*, **361**, pp. 921–946.
- [16] Kim, J. H., and Paulino, G. H., 2002, "Isoparametric Graded Finite Elements for Nonhomogeneous Isotropic and Orthotropic Materials," *J. Appl. Mech.*, **69**(4), pp. 502–514.

Bo-Wun Huang¹

Professor
Department of Mechanical Engineering,
Cheng Shiu University,
840 Cheng Ching Road,
Niaosung,
Kaohsiung, Taiwan
e-mail: huangbw@csu.edu.tw

Jao-Hwa Kuang

Professor
Department of Mechanical and Electromechanical
Engineering,
National Sun Yat-Sen University,
Kaohsiung 80424, Taiwan
e-mail: kuang@mail.nsysu.edu.tw

The Parametric Resonance Instability in a Drilling Process

This study investigates dynamic instability in a high-speed drilling process. A pretwisted beam is used to simulate the drill. The time-dependent nature of the thrust force and the drilling depth is considered in the equation of motion of the drill. A moving Winkler-type elastic foundation assumption is applied to the drill tip to approximate the time-varying boundary conditions in the drilling process. Galerkin's method is used to formulate the characteristic equation in a discrete form. The variation of the instability regions of the drill system is solved and analyzed by employing the multiple-scales perturbation method. The numerical results indicate that the unstable regions suddenly enlarge and shift toward a lower frequency when the drill first contacts the work piece. The effects of the rotational speed, pretwisted angle, and thrust force of the drill on the variation of the dynamic instability in high-speed drilling are also studied and are found to be highly influential. [DOI: 10.1115/1.2722768]

Keywords: drilling process, pre-twisted beam, instability

Introduction

In manufacturing, drilling operations are performed on an extensive variety of machine tools, including drilling machines, milling machines, machining centers, etc. A precisely drilled hole enhances the quality of a product. The accuracy of the hole is determined largely by the particular drilling process employed [1,2]. In general, the majority of hole location errors, and reaming and fracturing of the drill, occur at the moment when the drill first enters the work piece. To improve the performance and capability of the drilling operation, it is necessary to understand the variation of the dynamic characteristics of the drilling system during the course of the drilling process. Accordingly, this study investigates the dynamic instability of a drill caused by fluctuations in the rotational speed of the drill and the applied axial thrust force.

During a drilling process, the rotational speed rarely remains constant; small fluctuations in the rotational speed are always observed, particularly in high-speed drilling processes. At a certain critical drilling speed, these speed fluctuations may cause the system to enter a dynamically unstable condition. Several studies [3–6] have investigated the instability of various pretwisted beams. However, the equations of motion of these systems are generally assumed to be time independent. Various researchers have studied the natural frequency and transverse vibrations of a drill [7–10]. Simplified pretwisted beams have been commonly employed to model a drill [6,8,11,12]. The literature contains several studies relating to the buckling load and natural frequencies of a drill bit [13–15]. Various mathematical models of the complex drill bit have been proposed to estimate the natural frequencies and cutting behaviors of a drill. The effects of complex geometry or cutting chips on the cutting and dynamic properties of the drill bit have been extensively studied [16–19]. Some investigators [12,20,21] have addressed the stability of a drill and reported that the rotational speed, pretwisted angle, and axial force may alter its dynamic stability.

It has been widely reported that the dynamic characteristics of a drill are dependent on the drilling depth and the axial thrust force. In other words, the dynamic characteristics of the system, e.g., its

natural frequencies and instability regions, vary with time during the drilling process. Therefore, the dynamic results calculated from the fixed natural frequencies and normal modes become meaningless since the structure and the boundary conditions vary during the course of drilling. Recently, a number of studies [22–26] have explored the interaction between the structure dynamic responses with moving masses, forces, and moving boundaries. The current study uses a pretwisted beam with a time-dependent axial force and a time-varying elastic boundary to approximate the drill system in a drilling process. The effects of axial load and rotational speed fluctuations on the instability distributions are analyzed and discussed.

Formulating the Equations of Motion

A little difference was found between the complex and rectangular cross sections on dynamic drill characteristics as in Ref. [27]. This study simulates the drill using the simple cantilever pretwisted beam with rectangular cross section shown in Fig. 1(a). The drill is of length L , thickness t , breadth b , and rotates at a speed of Ω . The deflection components $v(r,t)$ and $u(r,t)$ denote the two transverse flexible deflections of the drill.

The equation of motion of the rotating drill is given by

$$\frac{\partial^2}{\partial r^2} \left(EI_{yy} \frac{\partial^2 u}{\partial r^2} + EI_{xy} \frac{\partial^2 v}{\partial r^2} \right) - \frac{\partial}{\partial r} \left(\bar{I}_{yy} \frac{\partial^3 u}{\partial t^2 \partial r} + \bar{I}_{xy} \frac{\partial^3 v}{\partial t^2 \partial r} \right) + \mu \frac{\partial^2 v}{\partial r^2} - 2\mu\Omega \frac{\partial v}{\partial t} - \mu u \Omega^2 + \mu v \frac{\partial \Omega}{\partial t} = 0 \quad (1a)$$

$$\frac{\partial^2}{\partial r^2} \left(EI_{xx} \frac{\partial^2 v}{\partial r^2} + EI_{xy} \frac{\partial^2 u}{\partial r^2} \right) - \frac{\partial}{\partial r} \left(\bar{I}_{xx} \frac{\partial^3 v}{\partial t^2 \partial r} + \bar{I}_{xy} \frac{\partial^3 u}{\partial t^2 \partial r} \right) + \mu \frac{\partial^2 v}{\partial r^2} + 2\mu\Omega \frac{\partial u}{\partial t} - \mu v \Omega^2 + \mu u \frac{\partial \Omega}{\partial t} = 0 \quad (1b)$$

where E and μ are the Young's modulus and the mass per unit length of the drill respectively; and I_{xx} , I_{yy} , and I_{xy} are the moments of the cross-sectional area. The drill is considered to be pretwisted with a uniform twist angle of β .

The area moments of inertia at position r can be derived as

$$I_{xx}(r) = I_{XX} \cos^2 \left(\frac{r}{L} \beta \right) + I_{YY} \sin^2 \left(\frac{r}{L} \beta \right) \quad (2a)$$

¹Corresponding author.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received June 15, 2005; final manuscript received January 4, 2007. Review conducted by Marc P. Mignolet. Paper presented at the 2005 ASME International Mechanical Engineering Congress (IMECE2005), Orlando, FL, USA, November 5–11, 2005.

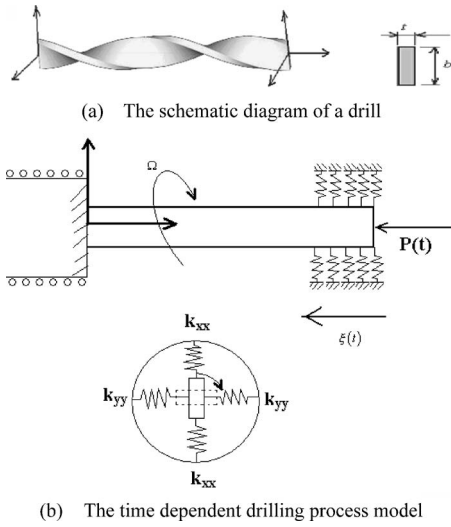


Fig. 1 Schematic of drill considered in current drilling process

$$I_{yy}(r) = I_{xx} \sin^2\left(\frac{r}{L}\beta\right) + I_{yy} \cos^2\left(\frac{r}{L}\beta\right) \quad (2b)$$

$$I_{xy}(r) = (I_{yy} - I_{xx}) \sin\left(\frac{r}{L}\beta\right) \cos\left(\frac{r}{L}\beta\right) \quad (2c)$$

where

$$I_{xx} = \frac{bt^3}{12} \quad (3a)$$

$$I_{yy} = \frac{b^3t}{12} \quad (3b)$$

and $\bar{I}_{xx}$, $\bar{I}_{xy}$, $\bar{I}_{yy}$ are the moments of mass.

The boundary conditions of the drill are given by

$$u = v = \partial u / \partial r = \partial v / \partial r = 0, \quad \text{at } r = 0 \quad (4a)$$

$$\partial^2 u / \partial r^2 = \partial^2 v / \partial r^2 = \partial^3 u / \partial r^3 = \partial^3 v / \partial r^3 = 0, \quad \text{at } r = L \quad (4b)$$

To investigate the time-dependent dynamic characteristics of the drill in a high-speed drilling process, a time-dependent boundary and axial thrust force are considered in simulations of the drilling process at different drilling depths. Since the thrust force remains constant as soon as drilling starts, only a moving drill depth is employed in the current formulation. A unit step function $u^*(t-t^*)$ is employed to indicate the moment at which drilling commences. Two moving Winkler-type elastic foundations [28] are used to simulate the time-dependent drilling depth, as shown in Fig. 1(b). The linear axial velocity motion of the foundations is considered. Finally, two moving stiffnesses $k_{xx}(t)$ and $k_{yy}(t)$, are used to represent the boundary stiffness.

Considering a time-dependent axial force and drilling depth, the equation of motion of the drill in the drilling process can be rewritten as

$$\begin{aligned} & \frac{\partial^2}{\partial r^2} \left(EI_{yy} \frac{\partial^2 u}{\partial r^2} + EI_{xy} \frac{\partial^2 v}{\partial r^2} \right) - \frac{\partial}{\partial r} \left(\bar{I}_{yy} \frac{\partial^3 u}{\partial t^2 \partial r} + \bar{I}_{xy} \frac{\partial^3 v}{\partial t^2 \partial r} \right) \\ & + Pu^*(t-t^*) \frac{\partial^2 u}{\partial r^2} + 2k_{xx}u^*(r-r^*)u + \mu \frac{\partial^2 u}{\partial t^2} \\ & - 2\mu\Omega \frac{\partial v}{\partial t} - \mu\Omega^2 u - \mu v \Omega^2 = 0 \end{aligned} \quad (5a)$$

$$\begin{aligned} & \frac{\partial^2}{\partial r^2} \left(EI_{xx} \frac{\partial^2 v}{\partial r^2} + EI_{xy} \frac{\partial^2 u}{\partial r^2} \right) - \frac{\partial}{\partial r} \left(\bar{I}_{xx} \frac{\partial^3 v}{\partial t^2 \partial r} + \bar{I}_{xy} \frac{\partial^3 u}{\partial t^2 \partial r} \right) \\ & + Pu^*(t-t^*) \frac{\partial^2 v}{\partial r^2} + 2k_{yy}u^*(r-r^*)v + \mu \frac{\partial^2 v}{\partial t^2} \\ & + 2\mu\Omega \frac{\partial u}{\partial t} - \mu v \Omega^2 + \mu u \frac{\partial \Omega}{\partial t} = 0 \end{aligned} \quad (5b)$$

where $u^*(t)$ denotes a unit step function and $P(t) = Pu^*(t-t^*)$ indicates that the axial thrust load commences at time $t=t^*$. The time-varying drill length $\delta^*(t)$ in the drilling process is given by

$$\delta^*(t) = L - f \times (t-t^*)u^*(t-t^*) \quad (6)$$

where $f \times (t-t^*)u^*(t-t^*)$ is the drilling depth and f is the feed velocity. For convenience, a nondimensional drilling depth is defined as follows

$$\xi(t) = \frac{f \times (t-t^*)u^*(t-t^*)}{L} \quad (7)$$

Taking rotational speed fluctuations into consideration, the rotational speed of the drill is assumed to be

$$\Omega(t) = \Omega_0 + f_\Omega(t) \quad (8)$$

where $f_\Omega(t)$ is a small perturbation speed superimposed upon the constant speed Ω_0 . Then $\partial \Omega / \partial t = \dot{f}_\Omega(t)$, and the equation of motion of the drill during the drilling process can be rewritten in the following form

$$\begin{aligned} & \frac{\partial^2}{\partial r^2} \left(EI_{yy} \frac{\partial^2 u}{\partial r^2} + EI_{xy} \frac{\partial^2 v}{\partial r^2} \right) - \frac{\partial}{\partial r} \left(\bar{I}_{yy} \frac{\partial^3 u}{\partial t^2 \partial r} + \bar{I}_{xy} \frac{\partial^3 v}{\partial t^2 \partial r} \right) \\ & + Pu^*(t-t^*) \frac{\partial^2 u}{\partial r^2} + 2k_{xx}u^*(r-\delta^*)u \\ & + \mu \frac{\partial^2 u}{\partial t^2} - 2\mu\Omega_0 \frac{\partial v}{\partial t} - \mu\Omega_0^2 u \\ & = 2\mu f_\Omega \frac{\partial v}{\partial t} + 2\mu\Omega_0 f_\Omega u + \mu f_\Omega^2 u + \mu \dot{f}_\Omega v \end{aligned} \quad (9a)$$

$$\begin{aligned} & \frac{\partial^2}{\partial r^2} \left(EI_{xx} \frac{\partial^2 v}{\partial r^2} + EI_{xy} \frac{\partial^2 u}{\partial r^2} \right) - \frac{\partial}{\partial r} \left(\bar{I}_{xx} \frac{\partial^3 v}{\partial t^2 \partial r} + \bar{I}_{xy} \frac{\partial^3 u}{\partial t^2 \partial r} \right) \\ & + Pu^*(t-t^*) \frac{\partial^2 v}{\partial r^2} + 2k_{yy}u^*(r-\delta^*)v + \mu \frac{\partial^2 v}{\partial t^2} \\ & + 2\mu\Omega_0 \frac{\partial u}{\partial t} - \mu v \Omega_0^2 \\ & = -2\mu f_\Omega \frac{\partial u}{\partial t} + 2\mu v \Omega_0 f_\Omega + \mu v f_\Omega^2 v - \mu \dot{f}_\Omega u \end{aligned} \quad (9b)$$

The solutions for the above equations are assumed to be

$$u(r,t) = \sum_{i=1}^m p_i(t) \phi_i(r) \quad (10a)$$

$$v(r,t) = \sum_{i=1}^m q_i(t) \phi_i(r) \quad (10b)$$

where $\phi_i(r)$ are the comparison functions of Eqs. (9a) and (9b), and $p_i(t)$, $q_i(t)$ are the corresponding weighting coefficients to be determined. This study uses five exact solutions of the uniform cantilever beam in the transverse direction, i.e., $\phi_i(r)$ for $i = 1, 2, \dots, 5$, to discretize the equations of motion of the drill.

Using Galerkin's method, the above equations of motion can be derived in the following matrix form

$$\begin{aligned}
[M] \begin{Bmatrix} \ddot{p} \\ \ddot{q} \end{Bmatrix} + 2\Omega_0[G] \begin{Bmatrix} \dot{p} \\ \dot{q} \end{Bmatrix} + \{[K]_a + Pu^*(t-t^*)[K]_b + \Omega_0^2[K]_c \\ + [K]_d\} \begin{Bmatrix} p \\ q \end{Bmatrix} = -2f_\Omega[G] \begin{Bmatrix} \dot{p} \\ \dot{q} \end{Bmatrix} - \dot{f}_\Omega[G] \begin{Bmatrix} p \\ q \end{Bmatrix} - (2\Omega_0 f_\Omega + f_\Omega^2) \\ \times [K]_c \begin{Bmatrix} p \\ q \end{Bmatrix}
\end{aligned} \quad (11)$$

where

$$[M] = \begin{bmatrix} [M^{(1)}] + [M^{(2)}] & [M^{(4)}] \\ [M^{(4)}] & [M^{(1)}] + [M^{(3)}] \end{bmatrix} \quad (12a)$$

$$[G] = \begin{bmatrix} [0] & -[M^{(1)}] \\ [M^{(1)}] & [0] \end{bmatrix} \quad (12b)$$

$$[K]_a = \begin{bmatrix} [K^{(1)}] & [K^{(4)}] \\ [K^{(4)}] & [K^{(2)}] \end{bmatrix} \quad (12c)$$

$$[K]_b = \begin{bmatrix} -[K^{(3)}] & [0] \\ [0] & -[K^{(3)}] \end{bmatrix} \quad (12d)$$

$$[K]_c = \begin{bmatrix} -[M^{(1)}] & [0] \\ [0] & -[M^{(1)}] \end{bmatrix} \quad (12e)$$

$$[K]_d = \begin{bmatrix} (2k_{xx})[K^{(5)}] & [0] \\ [0] & (2k_{yy})[K^{(5)}] \end{bmatrix} \quad (12f)$$

These matrices can be illustrated in detail as follows

$$M_{ij}^{(1)} = \int_0^1 \mu \phi_i(\bar{r}) \phi_j(\bar{r}) d\bar{r} \quad (13a)$$

$$M_{ij}^{(2)} = \frac{\bar{I}_{yy}}{L^2} \int_0^1 \frac{d\phi_i(\bar{r})}{d\bar{r}} \frac{d\phi_j(\bar{r})}{d\bar{r}} d\bar{r} \quad (13b)$$

$$M_{ij}^{(3)} = \frac{\bar{I}_{xx}}{L^2} \int_0^1 \frac{d\phi_i(\bar{r})}{d\bar{r}} \frac{d\phi_j(\bar{r})}{d\bar{r}} d\bar{r} \quad (13c)$$

$$M_{ij}^{(4)} = \frac{\bar{I}_{xy}}{L^2} \int_0^1 \frac{d\phi_i(\bar{r})}{d\bar{r}} \frac{d\phi_j(\bar{r})}{d\bar{r}} d\bar{r} \quad (13d)$$

$$K_{ij}^{(1)} = \frac{EI_{yy}}{L^4} \int_0^1 \frac{d^2\phi_i(\bar{r})}{d\bar{r}^2} \frac{d^2\phi_j(\bar{r})}{d\bar{r}^2} d\bar{r} \quad (13e)$$

$$K_{ij}^{(2)} = \frac{EI_{xx}}{L^4} \int_0^1 \frac{d^2\phi_i(\bar{r})}{d\bar{r}^2} \frac{d^2\phi_j(\bar{r})}{d\bar{r}^2} d\bar{r} \quad (13f)$$

$$K_{ij}^{(3)} = \frac{1}{L^2} \int_0^1 \frac{d\phi_i(\bar{r})}{d\bar{r}} \frac{d\phi_j(\bar{r})}{d\bar{r}} d\bar{r} \quad (13g)$$

$$K_{ij}^{(4)} = \frac{EI_{xy}}{L^4} \int_0^1 \frac{d^2\phi_i(\bar{r})}{d\bar{r}^2} \frac{d^2\phi_j(\bar{r})}{d\bar{r}^2} d\bar{r} \quad (13h)$$

$$K_{ij}^{(5)} = \int_0^1 \phi_i(\bar{r}) \phi_j(\bar{r}) u^*(\bar{r} - \bar{\delta}^*) d\bar{r} \quad (13i)$$

where

$$\bar{r} = \frac{r}{L} \quad (14a)$$

$$\bar{\delta}^* = \frac{\delta^*}{L} \quad (14b)$$

For simplicity, Eq. (11) can be rewritten as

$$\begin{aligned}
[A]\{\dot{V}\} + [B]\{V\} = -2\frac{f_\Omega}{\Omega_0}[C]\{V\} - \frac{\dot{f}_\Omega}{\Omega_0}[D]\{V\} - \left(2\frac{f_\Omega}{\Omega_0} + \frac{f_\Omega^2}{\Omega_0^2}\right)[E] \\ \times \{V\}
\end{aligned} \quad (15)$$

where

$$[A] = \begin{bmatrix} [M] & [0] \\ [0] & [K]_a + Pu^*(t-t^*)[K]_b + \Omega_0^2[K]_c + [K]_d \end{bmatrix} \quad (16a)$$

$$[B] = \begin{bmatrix} 2\Omega_0[G] & \{[K]_a + Pu^*(t-t^*)[K]_b + \Omega^2[K]_c + [K]_d\} \\ -\{[K]_a + Pu^*(t-t^*)[K]_b + \Omega^2[K]_c + [K]_d\} & [0] \end{bmatrix} \quad (16b)$$

$$[C] = \begin{bmatrix} \Omega_0[G] & [0] \\ [0] & [0] \end{bmatrix} \quad (16c)$$

$$[D] = \begin{bmatrix} [0] & \Omega_0[G] \\ [0] & [0] \end{bmatrix} \quad (16d)$$

$$[E] = \begin{bmatrix} [0] & \Omega_0^2[K]_c \\ [0] & [0] \end{bmatrix} \quad (16e)$$

and

$$\{V\} = \left\{ \begin{Bmatrix} \dot{p} \\ \dot{q} \\ p \\ q \end{Bmatrix} \right\} \quad (16f)$$

The corresponding eigenvalue problem of Eq. (15), i.e., the eigenvalue problem of the conservative gyroscopic system with constant drilling speed $f_\Omega = \dot{f}_\Omega = 0$, is given by

$$[A]\{\dot{X}\} + [B]\{X\} = 0 \quad (17)$$

After modal calculation, the normalized modal matrix $[\Phi^*]$ can be solved and the above eigenvalue problem can be normalized as

$$[I]\{\dot{X}\} + [\Lambda]\{X\} = 0 \quad (18)$$

The modal matrix of Eq. (17), i.e., $[\Phi^*]$, is formed by the real and imaginary parts of the eigenvectors. The modal matrix $[\Phi^*]$ is normalized as

$$[\Phi^*]^T[A][\Phi^*] = [I] \quad (19a)$$

$$[\Phi^*]^T[B][\Phi^*] = [\Lambda] = \begin{bmatrix} \begin{bmatrix} 0 & -\omega_1 \\ \omega_1 & 0 \end{bmatrix} & [0] & \cdot & \cdot & \cdot & [0] \\ [0] & \begin{bmatrix} 0 & -\omega_2 \\ \omega_2 & 0 \end{bmatrix} & \cdot & \cdot & \cdot & [0] \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ [0] & [0] & \cdot & \cdot & \cdot & \begin{bmatrix} 0 & -\omega_{2m} \\ \omega_{2m} & 0 \end{bmatrix} \end{bmatrix} \quad (19b)$$

where $\omega_1, \omega_2, \dots, \omega_{2m}$ are the corresponding natural frequencies of Eq. (18). Substituting the linear transformation, i.e., $\{V\} = [\Phi^*] \times \{X\}$, into Eq. (16) and pre-multiplying this equation by $[\Phi^*]$ yields

$$[I]\{\dot{X}\} + [\Lambda]\{X\} = -2\frac{f_\Omega}{\Omega_0}[S]\{X\} - \frac{\dot{f}_\Omega}{\Omega_0}[Q]\{X\} - \left(2\frac{f_\Omega}{\Omega_0} + \frac{f_\Omega^2}{\Omega_0^2}\right)[Z]\{X\} \quad (20)$$

where

$$[S] = [\Phi^*]^T \begin{bmatrix} \Omega_0[G] & [0] \\ [0] & [0] \end{bmatrix} [\Phi^*] \quad (21a)$$

$$[Q] = [\Phi^*]^T \begin{bmatrix} [0] & \Omega_0[G] \\ [0] & [0] \end{bmatrix} [\Phi^*] \quad (21b)$$

$$[Z] = [\Phi^*]^T \begin{bmatrix} [0] & \Omega_0^2[K]_c \\ [0] & [0] \end{bmatrix} [\Phi^*] \quad (21c)$$

To match the form of matrix $[\Lambda]$, the matrices on the right-hand side are partitioned into $R \times R$ pieces of 2×2 block, where $R = 2m$. Equation (19) can be uncoupled and expressed in component form as

$$\begin{aligned} \dot{s}_n - \omega_n \eta_n = & -2\frac{f_\Omega}{\Omega_0} \left(\sum_{r=1}^R S_{nr}^{11} \zeta_r + \sum_{r=1}^R S_{nr}^{12} \eta_r \right) - \frac{\dot{f}_\Omega}{\Omega_0} \left(\sum_{r=1}^R Q_{nr}^{11} \zeta_r \right. \\ & \left. + \sum_{r=1}^R Q_{nr}^{12} \eta_r \right) - \left(2\frac{f_\Omega}{\Omega_0} + \frac{f_\Omega^2}{\Omega_0^2} \right) \left(\sum_{r=1}^R Z_{nr}^{11} \zeta_r + \sum_{r=1}^R Z_{nr}^{12} \eta_r \right) \end{aligned} \quad (22a)$$

$$\begin{aligned} \dot{\eta}_n + \omega_n s_n = & -2\frac{f_\Omega}{\Omega_0} \left(\sum_{r=1}^R S_{nr}^{21} \zeta_r + \sum_{r=1}^R S_{nr}^{22} \eta_r \right) - \frac{\dot{f}_\Omega}{\Omega_0} \left(\sum_{r=1}^R Q_{nr}^{21} \zeta_r \right. \\ & \left. + \sum_{r=1}^R Q_{nr}^{22} \eta_r \right) - \left(2\frac{f_\Omega}{\Omega_0} + \frac{f_\Omega^2}{\Omega_0^2} \right) \left(\sum_{r=1}^R Z_{nr}^{21} \zeta_r + \sum_{r=1}^R Z_{nr}^{22} \eta_r \right) \end{aligned} \quad (22b)$$

where S_{nl}^{ij} , Q_{nl}^{ij} , and Z_{nl}^{ij} are the i - j th entries of the n - l th block matrices of $[S]$, $[Q]$, and $[Z]$, respectively, and $\{X\} = [s_1, \eta_1, s_2, \eta_2, \dots, s_R, \eta_R]^T$.

In this study, the perturbation velocity, $f_\Omega(t)$, is assumed to be small and periodic. It can therefore be expressed as a Fourier series, i.e., $f_\Omega(t) = \sum_{j=-Q}^Q F_j e^{i\omega_j t}$, where the parameter ω_j is the perturbation frequency. As noted, the perturbation term $f_\Omega(t)$ is very small in comparison to the constant speed, Ω_0 . Therefore, the module F_j should also be much smaller than the constant speed,

Ω_0 . Defining a small parameter term ϵ as $|F_M/\Omega_0|$, where $|F_M|$ is the maximum magnitude of components F_j for $j=1, 2, 3, \dots, Q$, Eqs. (22a) and (22b) can be rearranged as

$$\begin{aligned} \dot{s}_n - \bar{\omega}_n \eta_n = & -2\epsilon \bar{f} \left(\sum_{r=1}^R S_{nr}^{11} \zeta_r + \sum_{r=1}^R S_{nr}^{12} \eta_r \right) \\ & - i \sum_{-M}^M \bar{\omega}_j \epsilon \bar{f}_j e^{i\bar{\omega}_j t} \left(\sum_{r=1}^R Q_{nr}^{11} \zeta_r + \sum_{r=1}^R Q_{nr}^{12} \eta_r \right) - (2\epsilon \bar{f} + \epsilon^2 \bar{f}^2) \\ & \times \left(\sum_{r=1}^R Z_{nr}^{11} \zeta_r + \sum_{r=1}^R Z_{nr}^{12} \eta_r \right) \end{aligned} \quad (23a)$$

$$\begin{aligned} \dot{\eta}_n - \bar{\omega}_n s_n = & -2\epsilon \bar{f} \left(\sum_{r=1}^R S_{nr}^{21} \zeta_r + \sum_{r=1}^R S_{nr}^{22} \eta_r \right) \\ & - i \sum_{-M}^M \bar{\omega}_j \epsilon \bar{f}_j e^{i\bar{\omega}_j t} \left(\sum_{r=1}^R Q_{nr}^{21} \zeta_r + \sum_{r=1}^R Q_{nr}^{22} \eta_r \right) \\ & - (2\epsilon \bar{f} + \epsilon^2 \bar{f}^2) \left(\sum_{r=1}^R Z_{nr}^{21} \zeta_r + \sum_{r=1}^R Z_{nr}^{22} \eta_r \right) \end{aligned} \quad (23b)$$

where $\bar{f} = f_\Omega/|F_M|$, $\bar{f}_j = F_j/|F_M|$, and $\bar{\omega} = \omega/\sqrt{EI_{XX}/\rho AL^4}$.

By employing the multiple-scales perturbation method, the instability regions of the drilling system can be solved and analyzed. Due to the complexity of the present investigation and the negligible difference between the results of the first- and second-order approximations, the second-order expansion is not carried out [23,24,26]. Following calculation, the unstable regions of the drill system can be expressed as follows:

- (i) Frequency $\bar{\omega}_j$ away from $\bar{\omega}_p \pm \bar{\omega}_q$. The system is always stable in this state.
- (ii) Frequency $\bar{\omega}_j$ near $\bar{\omega}_p + \bar{\omega}_q$. Transition curves are used to locate the stability boundaries. In this case, the transition curves can be derived as follows

$$\bar{\omega}_j = \bar{\omega}_p + \bar{\omega}_q \pm 2\epsilon \sqrt{\Lambda_{pq} \bar{\Lambda}_{qp}} + O(\epsilon^2) \quad (24a)$$

where

$$\begin{aligned} \Lambda_{pq} = & \sum_{j=1}^Q \hat{F}_j \left(-iS_{pq}^{21} - S_{pq}^{22} + \frac{1}{2} \bar{\omega}_j Q_{pq}^{21} - \frac{1}{2} i \bar{\omega}_j Q_{pq}^{22} - iZ_{pq}^{21} \right. \\ & \left. - Z_{pq}^{22} \right) - \sum_{j=1}^Q \hat{F}_j \left(-S_{pq}^{11} + iS_{pq}^{12} - i\frac{1}{2} \bar{\omega}_j Q_{pq}^{11} - \frac{1}{2} \bar{\omega}_j Q_{pq}^{12} \right. \\ & \left. - Z_{pq}^{11} + iZ_{pq}^{12} \right) \end{aligned} \quad (24b)$$

Table 1 The difference in the first natural frequency of the pretwisted beam solved by employing finite element model [9,11] and the approximation model used in this work ($L=100$ cm, $t=0.5$ cm, $b=8$ cm)

Pretwisted angle (deg)	Ref. [9]	Ref. [11]	This study
0	1.00	1.00	1.000
45	1.01	1.01	1.007
90	1.02	1.03	1.028

$$\Lambda_{qp} = \sum_{j=1}^Q \hat{F}_j \left(-iS_{qp}^{21} - S_{qp}^{22} + \frac{1}{2}\bar{\omega}_j Q_{qp}^{21} - \frac{1}{2}i\bar{\omega}_j Q_{qp}^{22} - iZ_{qp}^{21} - Z_{qp}^{22} \right) - \sum_{j=1}^Q \hat{F}_j \left(-S_{qp}^{11} + iS_{qp}^{12} - i\frac{1}{2}\bar{\omega}_j Q_{qp}^{11} - \frac{1}{2}\bar{\omega}_j Q_{qp}^{12} - Z_{qp}^{11} + iZ_{qp}^{12} \right) \quad (24c)$$

(iii) Frequency $\bar{\omega}_j$ near $\bar{\omega}_p - \bar{\omega}_q$. Similarly, the transition curves of the combination resonances of the different type can be solved as

$$\bar{\omega}_j = \bar{\omega}_p - \bar{\omega}_q \pm 2\epsilon \sqrt{-\Lambda_{pq}\bar{\Lambda}_{qp}} + O(\epsilon^2) \quad (25)$$

Results and Discussion

This study investigated the dynamic instability variation in a high-speed drilling process. The geometrical parameters of the drill were considered to be: $(t/L)=0.005$, $(b/L)=0.01$, and $\beta=31.416$ rad/m. In accordance with the data provided in Ref. [29], the thrust load was specified as $P=1500$ N, the feed speed as $f=0.002$ m/s, and the lateral stiffness as $k_{xx}=k_{yy}=1 \times 10^8$ N/m². For convenience, the following dimensionless drilling speed was employed

$$\bar{\Omega}_0 = \Omega_0 / \sqrt{EI_{xx}/\rho A L^4} \quad (26)$$

Table 1 lists the difference in the first natural frequency calculated by employing the finite element method [9,11] and the method with the approximated comparison functions as mentioned in this study. Results indicated that the eigensolutions of a cantilever beam can be used to formulate the equations of motion of a twisted beam with good accuracy. Figure 2 shows the varia-

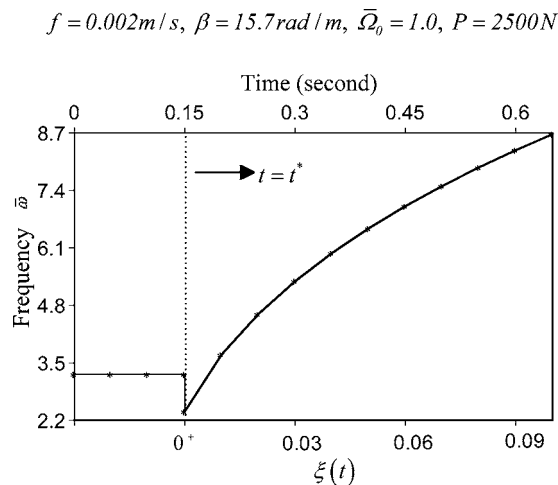
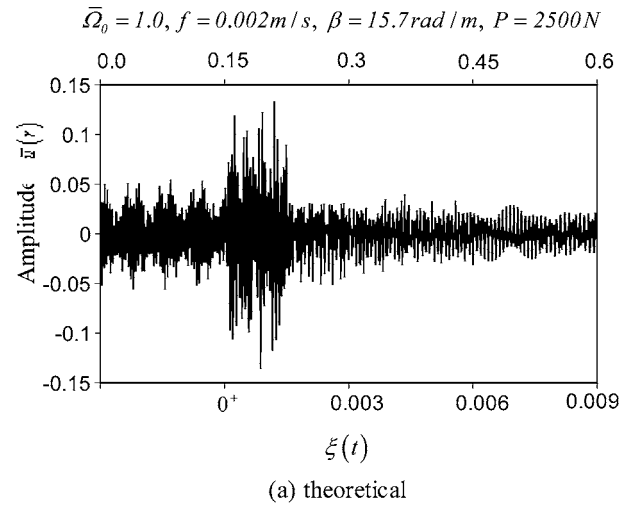


Fig. 2 Variation of lowest natural frequency of drill in drilling process



$\Omega = 560$ rpm, Diameter = 10 mm, $f = 0.002$ m/s, workpiece=steel

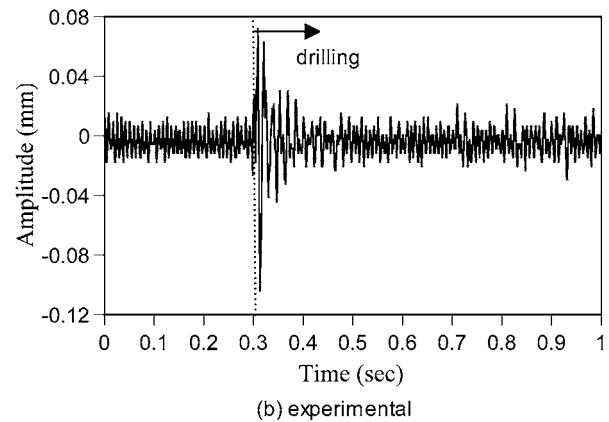


Fig. 3 Dynamic response of drill in drilling process

tion of the fundamental natural frequencies of the drill during the drilling process. The drill contacts the work piece at time $t=t^*=0.15$ s. The results show that a constant fundamental natural frequency is obtained before the drill makes contact with the work piece. The fundamental natural frequency then drops abruptly as the drill tip touches the work piece at time $t=t^*=0.15$ s. During the drilling process, i.e., $t>t^*$, the fundamental natural frequency of the drill increases gradually with increasing drilling depth, $\xi(t)$. In general, the stiffness of the drill increases as the length of the drill remaining outside of the work piece, i.e., $\delta^*(t)=L-f \times (t-t^*)u^*(t-t^*)$, decreases. The simulated and measured dynamic responses of the drill in the drilling process are illustrated in Figs. 3(a) and 3(b). The nondimensional amplitude $\bar{u}(r)$ is displayed in Fig. 3(a). A noncontact displacement sensor, KAMAN KD 2300 2S, is attached on the spindle to measure the vibration amplitude in the drilling operation. A large transverse vibration occurs at time $t=t^*=0.15$ s. The time domain responses of the drill in Figs. 3(a) and 3(b) indicate that a serve vibration is observed as the drill touches the work piece. The vibration is depressed as the drill drills into the work piece.

As noted earlier, rotational speed of the drill at normal operating conditions never remains absolutely constant. Theoretically, at a certain rotational speed, the speed fluctuations may cause the system to become dynamically unstable. In this paper, "the certain rotational speed" indicates the speed that will introduce a parametric resonance. The reason to introduce a perturbation with this speed is to derive the necessary condition for an unstable system. This perturbation assumption has been widely used [5,6,21] to

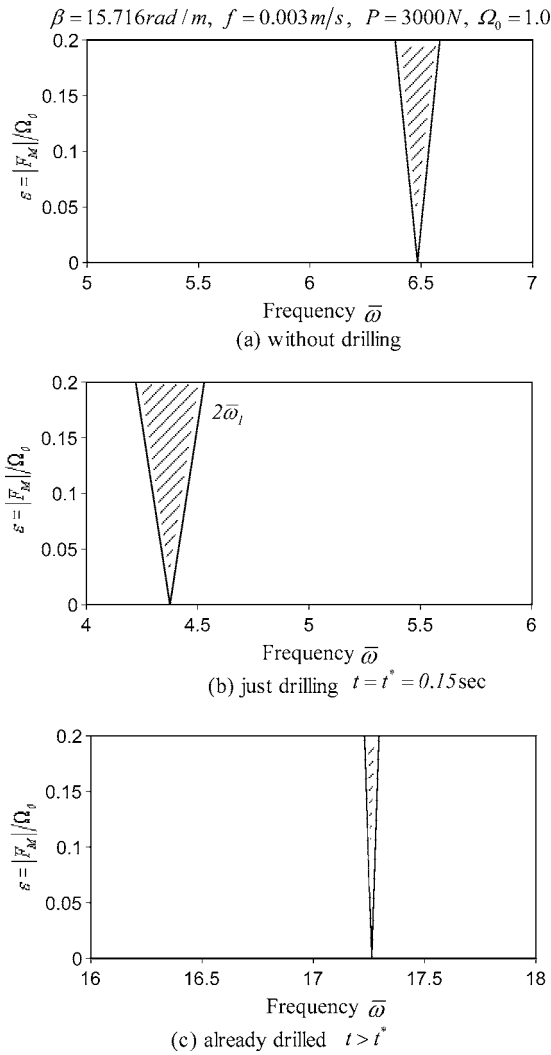


Fig. 4 Variation in instability zone of drill during drilling process

derive the instability criterion. For simplicity, this study assumes the perturbation speed to be $\bar{f}(t) = 2 \cos \bar{\omega}t$. Figure 4(a) shows that the lowest unstable region is located near $2\bar{\omega}_1 = 6.483$ prior to drilling. As shown in Fig. 4(b), a wider unstable region occurs at time $t = t^* = 0.15 \text{ s}$ when the drill first makes contact with the work piece, and the unstable region shifts from $2\bar{\omega}_1 = 6.483$ to $2\bar{\omega}_1 = 4.375$. Once the drill tip has entered the work piece, Fig. 4(c) shows that the lowest unstable region of the drill shifts gradually to a higher frequency, i.e., $2\bar{\omega}_1 = 17.262$. The prediction of the instability of a rotating drill before drilling into a work piece, as shown in Fig. 4(a), is similar to the prediction of the stability of a rotating pretwisted beam [12]. In this study, the instability of a drill in the drilling process, as shown in Fig. 4(c), is simulated as the rotating beam with a spring support at the end. Numerical results indicate that the unstable zones of the drill will be depressed and shift toward higher frequencies as the drill drilling into the work piece.

Another key factor affecting the distribution of the instability regions is the applied axial thrust load. Figure 5 shows the variation in the instability regions for different axial thrust load with different feed velocity. The lowest unstable region, i.e., located near $2\bar{\omega}_1$, for the drill system with a dimensionless perturbation speed of $\epsilon = 0.2$ is plotted. The results indicate that the unstable region becomes larger and shifts to a lower frequency at the moment when the drill first touches the work piece. As the drilling

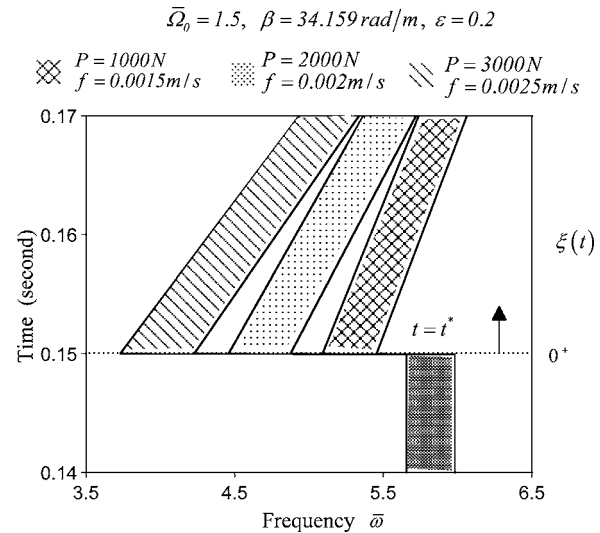


Fig. 5 Variation in the instability regions for different axial thrust load with different feed velocity

depth increases, the unstable region shifts gradually to a higher frequency. The results also show that the fundamental unstable region expands and shifts to a lower frequency when the magnitude of the applied thrust load is increased.

This study also examined the effect of the pretwisted angle of the drill on the variation of dynamic instability in the drilling system. The variations of the fundamental unstable region when drilling is performed with pretwisted angles of $\beta = 15.708$ and $\beta = 62.832$ are plotted in Figs. 6(a) and 6(b), respectively. As reported by previous studies [5,6,9,10], the current results reveal

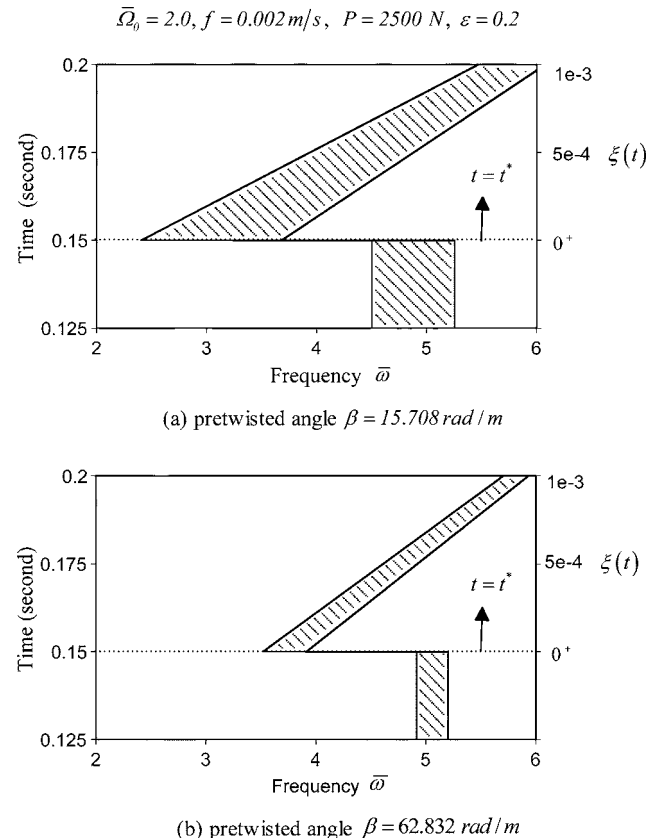


Fig. 6 Variation in instability zone of drill with different pretwisted angles

$$\beta = 31.4 \text{ rad/m}, f = 0.002 \text{ m/s}, P = 2500 \text{ N}, \varepsilon = 0.2$$

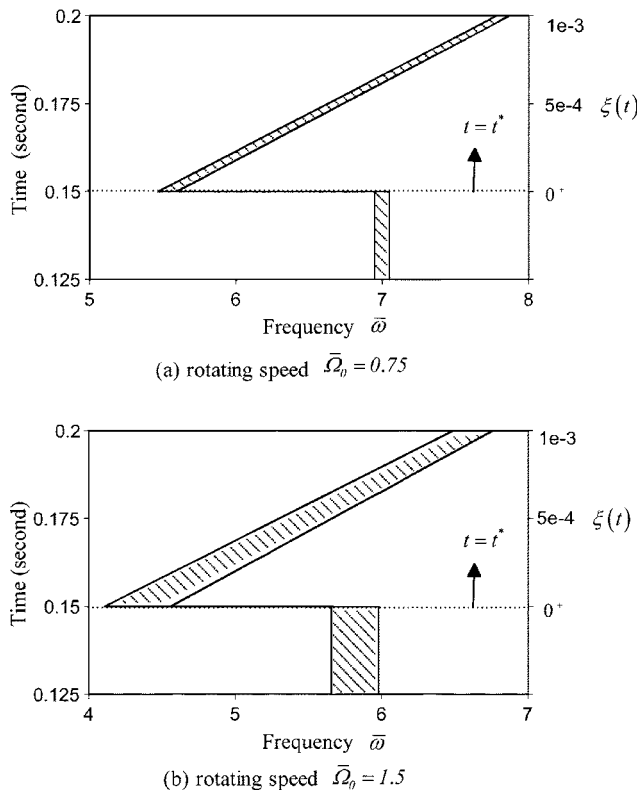


Fig. 7 Variation in instability zone of drill with different drilling speeds

that the two lower lateral natural frequencies of the drill become virtually identical as the pretwisted angle of the drill increases. The asymmetric property of the drill is diminished as the pretwisted angle is increased. The fundamental unstable region which exists near $2\bar{\omega}_1$ during the drilling process becomes smaller as the asymmetric property of the drill is reduced. Hence, the fundamental unstable region becomes narrower as the pretwisted angle increases.

Figure 7 illustrates the variation of the fundamental instability region of the drill at drilling speeds of $\bar{\Omega}_0 = 0.75$ and 1.50 . It is observed that the fundamental unstable region of the drill shifts and becomes larger as the drilling speed is increased.

Conclusions

This study has formulated and investigated the variation of the parametric resonance instability in a drilling process. The simulated and measured dynamic responses during the drilling process confirm the feasibility and accuracy of the proposed model. The major conclusions to be drawn from the current analysis and numerical results are summarized as follows:

1. The dynamic response of a drill in a drilling process is time dependent. Fluctuations of the time-varying drilling depth, drilling speed, and axial thrust load may change the natural frequencies of the drill and the distribution of the instability regions.
2. The fundamental unstable region of the drill shifts and becomes significantly larger at the moment that the drill first makes contact with the work piece. Additionally, a variation in the dynamic response occurs as the drilling depth increases.
3. The applied thrust force, drilling speed, and pretwisted angle all influence the distribution of the dynamic instability regions during the drilling process.

Acknowledgment

The current authors gratefully acknowledge the financial support provided to this study by the National Science Council, Taiwan, Republic of China under Grant No. NSC 93-2212-E-230-006.

References

- [1] Katz, Z., and Poustie, A., 2001, "On the Hole Quality and Drill Wandering Relationship," *Int. J. Adv. Manuf. Technol.*, **17**(4), pp. 233–237.
- [2] Gupta, K., Ozdoganlar, O. B., Kapoor, S. G., and DeVor, R. E., 2003, "Modeling and Prediction of Hole Profile in Drilling, Part 1: Modeling Drill Dynamics in the Presence of Drill Alignment Errors," *ASME J. Manuf. Sci. Eng.*, **125**(1), pp. 6–13.
- [3] Hsu, C. S., 1961, "On a Restricted Class of Coupled Hill's Equations and Some Applications," *ASME J. Appl. Mech.*, **28**, pp. 551–557.
- [4] Nayfeh, A. H., and Mook, D. T., 1979, *Nonlinear Oscillation*, Wiley, New York.
- [5] Young, T. H., 1991, "Dynamic Response of A Pretwisted, Tapered Beam With Non-constant Rotating Speed," *J. Sound Vib.*, **150**(3), pp. 435–446.
- [6] Liao, C. L., and Huang, B. W., 1995, "Parametric Resonance of a Spinning Pretwisted Beam With Time-Dependent Spinning Rate," *J. Sound Vib.*, **180**(1), pp. 47–65.
- [7] Rosard, D. D., 1952, "Natural Frequency of Twisted Cantilever Beams," *ASME Paper No. 52-A-15*.
- [8] Jarrett, G. W., and Warner, P. C., 1953, "The Vibration of Rotating Tapered Twisted Beam," *ASME J. Appl. Mech.*, **20**, pp. 381–389.
- [9] Tekinalp, O., and Ulsoy, A. G., 1989, "Modeling and Finite Element Analysis of Drill Bit Vibration," *ASME J. Vib., Acoust., Stress, Reliab. Des.*, **111**, pp. 148–155.
- [10] Tekinalp, O., and Ulsoy, A. G., 1990, "Effect of Geometric and Process Parameters in Drill Transverse Vibration," *ASME J. Eng. Ind.*, **112**, pp. 189–194.
- [11] Liao, C. L., and Dang, Y. H., 1992, "Structural Characteristics of Spinning Pretwisted Orthotropic Beams," *Comput. Struct.*, **45**(4), pp. 715–731.
- [12] Liao, C. L., and Huang, B. W., 1995, "Parametric Instability of a Pretwisted Beam Under Periodic Axial Force," *Int. J. Mech. Sci.*, **37**(4), pp. 423–439.
- [13] Magrab, E. B., and Gilsinn, D. E., 1984, "Buckling Loads and Natural Frequencies of Drill Bits and Fluted Cutters," *ASME J. Eng. Ind.*, **106**, pp. 196–204.
- [14] Ulsoy, A. G., 1983, "A Lumped Parameter Model for the Transverse Vibration of Drill Bits," *Proceedings of Control of Manufacturing Processes and Robotic Systems*, ASME, New York, pp. 15–25.
- [15] Fuji, H., Marui, E., and Ema, S., 1988, "Whirling Vibration in Drilling Part 3: Vibration Analysis in Drilling Workpiece With a Pilot Hole," *ASME J. Eng. Ind.*, **110**, pp. 315–321.
- [16] Rincon, D. M., and Ulsoy, A. G., 1995, "Complex Geometry, Rotary Inertia and Gyroscopic Moment Effects on Drill Vibrations," *J. Sound Vib.*, **188**(5), pp. 701–715.
- [17] Ren, K., and Ni, J., 1999, "Analyses of Drill Flute and Cutting Angles," *Int. J. Adv. Manuf. Technol.*, **15**(8), pp. 546–553.
- [18] Spanos, P. D., Chevallier, A. M., and Politis, N. P., 2002, "Nonlinear Stochastic Drill-String Vibrations," *ASME J. Vib. Acoust.*, **124**(4), pp. 512–518.
- [19] Hsieh, J. F., and Lin, P. D., 2002, "Mathematical Model of Multiflute Drill Point," *Int. J. Mach. Tools Manuf.*, **42**(10), pp. 1181–1193.
- [20] Iwastsubo, T., Sugiyama, Y., and Ishihara, K., 1972, "Stability and Stationary Vibrations of Columns Under Periodic Loads," *J. Sound Vib.*, **23**, pp. 245–257.
- [21] Lee, H. P., 1994, "Buckling and Dynamic Stability of Spinning Pre-twisted Beams Under Compressive Axial Loads," *Int. J. Mech. Sci.*, **36**(11), pp. 1011–1026.
- [22] Lee, H. P., 1996, "Dynamic Response of a Beam With a Moving Mass," *J. Sound Vib.*, **192**(2), pp. 289–294.
- [23] Ram, Y. M., and Caldwell, J., 1996, "Free Vibration of a String With Moving Boundary Conditions by the Method of Distorted Images," *J. Sound Vib.*, **194**(1), pp. 35–47.
- [24] Fung, R. F., Huang, J. S., and Chu, J. J., 1998, "Dynamic Stability of an Axially Travelling String/Slider Coupling System With Moving Boundary," *J. Sound Vib.*, **211**(4), pp. 689–701.
- [25] Law, S. S., and Zhu, X. Q., 2000, "Study on Different Beam Models in Moving Force Identification," *J. Sound Vib.*, **234**(4), pp. 661–679.
- [26] Wu, J. S., and Chen, D. W., 2000, "Dynamic Analysis of a Uniform Cantilever Beam Carrying a Number of Elastically Mounted Point Masses With Dampers," *J. Sound Vib.*, **229**(3), pp. 549–578.
- [27] Huang, B. W., 2000, "Vibration of a High Speed Drill," *Proceedings of National Science Council Research Projects*, Kaohsiung, Taiwan, December, 2000, Paper No. NSC89-2212-E-230-006 (in Chinese).
- [28] Lee, H. P., 1996, "Dynamic Stability of a Tapered Cantilever Beam on an Elastic Foundation Subjected to a Follower Force," *Int. J. Solids Struct.*, **33**(10), pp. 1409–1424.
- [29] Huang, B. W., 2003, "Dynamic Characteristics of a Drill in Drilling Process," *IMEchE, IMechE Conf. Trans.*, **217**, pp. 161–167.

Normality Structures With Thermodynamic Equilibrium Points

Q. Yang¹

e-mail: yangq@tsinghua.edu.cn

R. K. Wang

L. J. Xue

State Key Laboratory of Hydrosience
and Engineering,
Tsinghua University,
Beijing 100084, P. R. China

Enriched by the nonlinear Onsager reciprocal relations and thermodynamic equilibrium points (Onsager, Phys. Rev., 37, pp. 405–406; 38, pp. 2265–2279), an extended normality structure by Rice (1971, J. Mech. Phys. Solids, 19, pp. 433–455) is established in this paper as a unified nonlinear thermodynamic theory of solids. It is revealed that the normality structure stems from the microscale irrotational thermodynamic fluxes. Within the extended normality structure, this paper focuses on the microscale thermodynamic mechanisms and significance of the convexity of flow potentials and yield surfaces. It is shown that the flow potential is convex if the conjugate force increment cannot not oppose the increment of the rates of local internal variables. For the Rice fluxes, the convexity condition reduces to the local rates being monotonic increasing functions with respect to their conjugate forces. The convexity of the flow potential provides the thermodynamic system a capability against the disturbance of the thermodynamic equilibrium point. It is proposed for time-independent behavior that the set of plastically admissible stresses determined by yield conditions corresponds to the set of thermodynamic equilibrium points. Based on that viewpoint, the intrinsic dissipation inequality is just the thermodynamic counterpart of the principle of maximum plastic dissipation and requires the convexity of the yield surfaces. [DOI: 10.1115/1.2722772]

Keywords: irreversible thermodynamics, equilibrium point, Onsager fluxes, normality, convexity

1 Introduction

Linear irreversible thermodynamics based on the Onsager [1,2] reciprocal relations where thermodynamic fluxes and forces are assumed to be linearly dependent and related by a symmetric phenomenological coefficient matrix, have provided access to both the understanding and the analysis of a wide range of physical phenomena, see, e.g., De Groot and Mazur [3]. The linear theory is restricted to linear thermodynamic processes or else the deviation from equilibrium have to be sufficiently small so that a linear approximation is valid. However, many processes in continuum mechanics are nonlinear, especially for inelastic behaviors of solids. The nonlinear generalization of the linear theory includes the normality structure of Rice [4,5], the orthogonality condition of Ziegler [6], and the so-called “nonlinear Onsager reciprocal relations” of Edelen [7].

As an internal variable approach, the normality structure proposed by Rice [4,5] has been an appealing constitutive framework for solids undergoing nonlinear irreversible thermodynamic processes. Rice’s kinetic rate laws of microscale (e.g., at the dislocation level) internal variables, with each rate being stress dependent only via its conjugate thermodynamic force, are cornerstones of the normality structure and represent a wide class of inelastic behaviors. This type of kinetic rate laws are termed the “Rice fluxes” in this paper.

The Rice fluxes are only a sufficient condition leading to the normality structure. As indicated by Yang et al. [8] and briefed in Sec. 3, the necessary and sufficient condition is the nonlinear Onsager reciprocal relations. The general thermodynamic fluxes satisfying the nonlinear Onsager reciprocal relations is termed the “Onsager fluxes” by Edelen [9] which embody the Rice fluxes as

a special case. It is revealed that the normality structure stems from the microscale irrotational thermodynamic fluxes. In Sec. 4, the concept of thermodynamic equilibrium points is introduced into the normality structure following the work of Edelen [9]. Enriched by the nonlinear Onsager reciprocal relations and thermodynamic equilibrium points, an extended normality structure is established in this paper as a unified nonlinear thermodynamic theory of solids. It is suggested in Sec. 5 that the independence of the elementary processes at the microscale should be taken as the more fundamental postulate of the original normality structure of Rice [4], rather than the usual statement that the each internal variable rate is stress dependent only via its conjugate thermodynamic force.

Within the extended normality structure, this paper focuses on the microscale thermodynamic mechanisms and significance of the convexity of flow potentials and yield surfaces. For the Rice fluxes, this issue have been addressed by Rice [5] and Yang et al. [8] for some special cases, as shown in Sec. 6.

It is shown that the flow potential is convex if the conjugate force increment cannot not oppose the increment of the rates of local internal variables. For the Rice fluxes, the convexity condition reduces to the local rates being monotonic increasing functions with respect to their conjugate forces. The convexity of the flow potential provide the thermodynamic system a capability against the disturbance of the thermodynamic equilibrium point.

It is proposed for time-independent behavior that the set of plastically admissible stresses determined by yield conditions corresponds to the set of thermodynamic equilibrium points. Based on that viewpoint, Rice’s kinetic rate laws for time-independent behavior is the requirement that the intrinsic dissipation inequality should hold locally. Furthermore, the global intrinsic dissipation inequality is the thermodynamic counterpart of the principle of maximum plastic dissipation and requires the convexity of the yield surfaces.

The orthogonality condition of Ziegler [6] has also been a fruitful thermodynamic approach, especially in the aspect of soil mod-

¹Corresponding author.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received July 26, 2006; final manuscript received January 2, 2007. Review conducted by: Yonggang Huang.

els [10–12]. An overall discussion on the orthogonality condition from a modern thermodynamic point of view has been provided by Rajagopal et al. [13]. The relationship between the normality structure and the orthogonality condition has been addressed by Yang et al. [8,14]. As shown in Sec. 6, the homogeneity of both the Rice fluxes and Onsager fluxes directly leads to the orthogonality condition; thus, the thermodynamic processes described by the orthogonality condition are the subset of the processes described by the extended normality structure.

2 Normality Structure

The normality structure proposed by Rice [4,5] is a discrete internal-variable theory for finite deformation. This approach views inelastic deformation of a given sample of material of the type considered under macroscopically homogeneous strain and temperature as a sequence of constrained equilibrium states: the state of the material sample at any given time in the deformation history can be fully characterized by the corresponding values of strain and temperature and a collection of internal variables that represent the extent of microstructural rearrangement within the sample. Consider a material sample of size V , which is measured in an unloaded reference state and at a reference temperature ϑ_0 . Introduce the specific free energy ϕ and its Legendre transform ψ with respect to strain²

$$\phi = \phi(\boldsymbol{\varepsilon}, \vartheta, H), \quad \psi = \psi(\boldsymbol{\sigma}, \vartheta, H) = \boldsymbol{\varepsilon} : \frac{\partial \phi}{\partial \boldsymbol{\varepsilon}} - \phi \quad (1)$$

where ϑ denotes temperature; $\boldsymbol{\varepsilon}$ denotes any strain tensor, objective and symmetric, that measures deformation from an arbitrary reference state; $\boldsymbol{\sigma}$ denotes the symmetric conjugate stress such that $\boldsymbol{\sigma} : d\boldsymbol{\varepsilon}$ is the work per unit volume of the adopted reference state in any virtual deformation $d\boldsymbol{\varepsilon}$; H denotes symbolically the current pattern of microstructural rearrangement of constituent elements of the materials. At fixed H , variations of $\boldsymbol{\sigma}$ and ϑ necessarily induce a purely elastic response and lead to the stress-strain relations,

$$\boldsymbol{\sigma} = \frac{\partial \phi(\boldsymbol{\varepsilon}, \vartheta, H)}{\partial \boldsymbol{\varepsilon}}, \quad \boldsymbol{\varepsilon} = \frac{\partial \psi(\boldsymbol{\sigma}, \vartheta, H)}{\partial \boldsymbol{\sigma}} \quad (2)$$

If $\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}(\boldsymbol{\sigma}, \vartheta, H)$, the inelastic part of a strain increment is defined as

$$d^p \boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}(\boldsymbol{\sigma}, \vartheta, H + dH) - \boldsymbol{\varepsilon}(\boldsymbol{\sigma}, \vartheta, H) \quad (3)$$

and the plastic strain associated with the current plastic state H and the reference temperature ϑ_0 is

$$\boldsymbol{\varepsilon}^p = \boldsymbol{\varepsilon}^p(\mathbf{0}, \vartheta_0, H) \quad (4)$$

Consider two neighboring patterns of microstructural rearrangement denoted by $H, H + dH$. It is assumed that a set of incremental scalar internal variables $d\xi_1, d\xi_2, \dots, d\xi_n$ characterize the specific local rearrangements, which are represented collectively by dH , at sites throughout the material sample. The $d\xi$'s and dH are related by

$$\frac{1}{V} f_\alpha d\xi_\alpha = -d^p \phi = d^p \psi \quad (5)$$

where³

$$d^p \phi = \phi(\boldsymbol{\varepsilon}, \vartheta, H + dH) - \phi(\boldsymbol{\varepsilon}, \vartheta, H) \\ d^p \psi = \psi(\boldsymbol{\sigma}, \vartheta, H + dH) - \psi(\boldsymbol{\sigma}, \vartheta, H) \quad (6)$$

The inelastic part of a strain increment is, due to Eqs. (2), (5), and (6),

²The Legendre transforms actually involve a maximization on the strain.

³In this paper, Einstein's summation convention is adopted for repeated indexes. However, if an index range is listed like α in Eq. (12), the index is considered as a free index without the summation convention.

$$d^p \boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}(\boldsymbol{\sigma}, \vartheta, H + dH) - \boldsymbol{\varepsilon}(\boldsymbol{\sigma}, \vartheta, H) = \frac{\partial(d^p \psi)}{\partial \boldsymbol{\sigma}} = \frac{1}{V} \frac{\partial f_\alpha}{\partial \boldsymbol{\sigma}} d\xi_\alpha \quad (7)$$

which is termed *Rice's flow rule* in this paper. Rice's flow rule is the fundamental relationship and starting point of the normality structure. Equation (5) also defines the thermodynamic forces $f_1, f_2, \dots, f_n$ (collectively $\mathbf{f}$) conjugate to the variables,

$$\mathbf{f} = \mathbf{f}(\boldsymbol{\sigma}, \vartheta, H) \quad \text{or} \quad \mathbf{f} = \mathbf{f}(\boldsymbol{\varepsilon}, \vartheta, H) \quad (8)$$

The corresponding set of total internal variables

$$\boldsymbol{\xi} = \{\xi_1, \xi_2, \dots, \xi_n\} \quad (9)$$

generally are not state variables in the sense that thermodynamic state functions are not direct functions of $\boldsymbol{\xi}$, but instead depend on the path history of $\boldsymbol{\xi}$. In cases when $\boldsymbol{\xi}$ is one set of explicit state variables, the conjugate forces can be determined as

$$f_\alpha = V \frac{\partial \psi}{\partial \xi_\alpha} = -V \frac{\partial \phi}{\partial \xi_\alpha}, \quad \phi = \phi(\boldsymbol{\varepsilon}, \vartheta, \boldsymbol{\xi}), \quad \psi = \psi(\boldsymbol{\sigma}, \vartheta, \boldsymbol{\xi}) \quad (10)$$

Following the second law of thermodynamics, the entropy production function should be always non-negative

$$\sigma = \frac{1}{V} f_\alpha \dot{\xi}_\alpha \geq 0 \quad (11)$$

A unifying normality structure emerges in macroscopic constitutive relations when the following class of kinetic relations is adopted: at any given temperature and pattern of internal rearrangement within the material sample, the rate at which any specific structural rearrangement occurs is fully determined by the thermodynamics force associated with that rearrangement. That is,

$$\dot{\xi}_\alpha = \dot{\xi}_\alpha(f_\alpha, \vartheta, H), \quad (\alpha = 1, 2, \dots, n) \quad (12)$$

The thermodynamic fluxes satisfying Eq. (12) is termed the *Rice fluxes* in this paper. Therefore, the kinetic rate laws lead to a flow potential Q ,

$$dQ = \frac{1}{V} \dot{\xi}_\alpha df_\alpha, \quad Q = Q(\mathbf{f}, \vartheta, H) = \frac{1}{V} \int_0^{\mathbf{f}} \dot{\xi}_\alpha df_\alpha \quad (13)$$

where the integration is carried out at fixed ϑ and H , and defines a direct function Q of $\mathbf{f}$ since each term in the integrand is a total differential. Then the kinetic rate laws can be recast as

$$\dot{\xi}_\alpha = V \frac{\partial Q}{\partial f_\alpha} \quad (14)$$

Therefore, the following normality structure holds, noting $\mathbf{f} = \mathbf{f}(\boldsymbol{\sigma}, \vartheta, H)$,

$$\dot{\boldsymbol{\varepsilon}}^p = \frac{d^p \boldsymbol{\varepsilon}}{dt} = \frac{\partial Q}{\partial \boldsymbol{\sigma}}, \quad Q = Q(\boldsymbol{\sigma}, \vartheta, H) = Q(\mathbf{f}, \vartheta, H) \quad (15)$$

where t denotes time, since, due to the first equation in Eq. (13)

$$\frac{\partial Q}{\partial \boldsymbol{\sigma}} = \frac{1}{V} \frac{\partial f_\alpha}{\partial \boldsymbol{\sigma}} \dot{\xi}_\alpha \quad (16)$$

Time-independent inelastic behavior (as in the classical elastic-plastic idealization for metals) was also addressed by Rice [4] who proposed the time-independent counterpart of the kinetic rate laws (12),

$$\dot{\xi}_\alpha = \begin{cases} > 0 & \text{if } f_\alpha = f_\alpha^U \\ = 0 & \text{if } f_\alpha^L < f_\alpha < f_\alpha^U \\ < 0 & \text{if } f_\alpha = f_\alpha^L \end{cases} \quad (\alpha = 1, 2, \dots, n) \quad (17)$$

where f_α^L and f_α^U are functions of ϑ and H , which mark the lower and upper limits to the range of forces that are incapable of inducing a corresponding structural rearrangement. Evidently, Eq. (17) is consistent with Eq. (12) in the sense that $\dot{\xi}_\alpha$ is uniquely determined by f_α . It is implied by Eq. (17) that the yield condition

of an internal variable ξ_α is just its conjugate force f_α . Therefore, Eq. (17) leads to an associated flow rule, in view of Eq. (7).

It should be also noted that local inertia terms, temperature gradients, etc., within the material sample (say, as a consequence of local fields due to dislocation motion) have not been considered.

3 Nonlinear Onsager Reciprocal Relations and Onsager Fluxes

The Rice fluxes are only a sufficient condition leading to the normality structure. It is shown in this section that the necessary and sufficient condition is the nonlinear Onsager reciprocal relations. The general thermodynamic fluxes satisfying the nonlinear Onsager reciprocal relations are termed the “Onsager fluxes” by Edelen [9], which embody the Rice fluxes as a special case. As compared to the Rice fluxes (12), the following general fluxes are taken as the starting point in this section:

$$\dot{\xi}_\alpha = \dot{\xi}_\alpha(\mathbf{f}, \vartheta, H), \quad (\alpha = 1, 2, \dots, n) \quad (18)$$

Evidently, the Rice fluxes are just a specific case of the general fluxes. For the general fluxes, the necessary and sufficient condition leading to the exact differential for $f_\alpha \dot{\xi}_\alpha$ is given by

$$\frac{\partial \dot{\xi}_\alpha}{\partial f_\beta} = \frac{\partial \dot{\xi}_\beta}{\partial f_\alpha}, \quad (\alpha, \beta = 1, 2, \dots, n) \quad (19)$$

The condition (19) is the so-called nonlinear Onsager reciprocal relations by Edelen [7]. The general fluxes $\dot{\xi}$ satisfying the nonlinear Onsager reciprocal relations are termed the “Onsager fluxes” by Edelen [9]. If Eq. (18) gives a linear relation, i.e., $\dot{\xi}_\alpha = L_{\alpha\beta} f_\beta$, the well-established reciprocal relations of [1,2], $L_{\alpha\beta} = L_{\beta\alpha}$, is reproduced from the nonlinear Onsager reciprocal relations (19). Edelen [7] has shown that the nonlinear Onsager reciprocal relations (19) result from the requirement that the entropy production rate σ be a C^1 , non-negative, convex function of $\mathbf{f}$ with an absolute minimum at the equilibrium point.

The Onsager fluxes embody the Rice fluxes as a special case since the nonlinear Onsager reciprocal relations are fulfilled automatically for the Rice fluxes. Evidently, all deduced results in Sec. 2 of the normality structure still hold based on the Onsager fluxes. Therefore, we furnish the normality structure with a rigorous mathematical and physical basis and enlarge significantly the basis of Rice’s thermodynamic framework.

3.1 Physical Significance and Basis. Some equivalent expressions of the nonlinear Onsager reciprocal relations (19) are given as follows. The fluxes $\dot{\xi}$ at fixed ϑ and H constitutes a n -dimensional vector field in the space of its affinity $\mathbf{f}$ space. Evidently, the nonlinear Onsager reciprocal relations (19) are equivalent to the irrotational condition at every point in the vector field

$$\nabla_{\mathbf{f}} \times \dot{\xi} = \mathbf{0} \quad (20)$$

where $\nabla_{\mathbf{f}} \times$ denote the curl operator in $\mathbf{f}$ space. Therefore, from a fluid point of view, the Onsager fluxes $\dot{\xi}$ represent an irrotational or potential flow in its affinity space. The irrotational condition (20) can be reformulated into an equivalent integral form based on Stokes’ theorem

$$\oint \dot{\xi} \cdot d\mathbf{f} = \oint \dot{\xi}_\alpha df_\alpha = 0 \quad (21)$$

at fixed ϑ and H and for any oriented simple closed curve in $\mathbf{f}$ space.

The irrotational condition (20) or (21) furnishes fundamentally the physical basis of the normality structure. However, the basis seems not as stable and solid because Brown [15] has noted that:

The point of view adopted here is that the patterns of plastic

flow observed in nature are unpredictable because they often contain a rotational component; hence, there are many solutions for the continuity equation satisfied by the plastic displacement. Furthermore, and more fundamentally, the onset of plasticity marks the end of homogeneous strain: plastic displacement is of necessity localized on slip planes; thus, the outcome of an atomistic theory can be only plastic displacement and its spatial distribution.

The issue will be further addressed from a multiscale point of view in Sec. 3.2.

3.2 Multiscale Formulation and Significance. The local internal variable set ξ generally contains numerous elements. One set of practical constitutive equations for engineering materials usually employ a few of macroscopic internal variables. In the normality structure by Rice [4], one set of much reduced average internal variables $\zeta = \{\zeta_1, \zeta_2, \dots, \zeta_m\}$ can be introduced as the average measurements of ξ

$$\zeta_\mu = \zeta_\mu(\xi_1, \xi_2, \dots, \xi_n; V) \quad (\mu = 1, 2, \dots, m \ll n) \quad (22)$$

where V indicates averaging over the volume, which yields the incremental relationship

$$\delta \zeta_\mu = \frac{\partial \zeta_\mu}{\partial \xi_\alpha} \delta \xi_\alpha \quad (23a)$$

or

$$\dot{\zeta}_\mu = \frac{\partial \zeta_\mu}{\partial \xi_\alpha} \dot{\xi}_\alpha \quad (23b)$$

The thermodynamic forces acting on the averaging variables ζ are $g_1, g_2, \dots, g_m$ (collectively $\mathbf{g}$). In order for the averaging variables ζ to be able to describe the thermodynamic system characterized by ξ , the following equality has to hold for any arbitrary $\delta \xi$:

$$g_\mu \delta \zeta_\mu = \frac{1}{V} f_\alpha \delta \xi_\alpha \quad (24)$$

It is revealed by Yang et al. [16] that the variational equation (24) can be further formulated into an extremum principle, a *principle of maximum equivalent dissipation*. Substituting Eq. (23a) into Eq. (24) leads to

$$f_\alpha = V g_\mu \frac{\partial \zeta_\mu(\xi, V)}{\partial \xi_\alpha} = f_\alpha(\mathbf{g}, \vartheta, H) \quad (25)$$

which further leads to

$$\frac{\partial \zeta_\mu(\xi, V)}{\partial \xi_\alpha} = \frac{1}{V} \frac{\partial f_\alpha(\mathbf{g}, \vartheta, H)}{\partial g_\mu} \quad (26)$$

Therefore, the relationship between thermodynamic fluxes and affinities at microscale and macroscale is established, based on Eq. (23b) and Eq. (25)

$$\dot{\zeta}_\mu = \frac{\partial \zeta_\mu}{\partial \xi_\alpha} \dot{\xi}_\alpha = \frac{1}{V} \dot{\xi}_\alpha \frac{\partial f_\alpha(\mathbf{g}, \vartheta, H)}{\partial g_\mu} \quad (27)$$

Substituting Eq. (14) into (27), the normality structure at the macroscale emerges

$$\dot{\zeta}_\mu = \frac{\partial Q(\mathbf{g}, \vartheta, H)}{\partial g_\mu} \quad (28)$$

Multiplying (27) by dg_μ , one obtains

$$\dot{\zeta}_\mu dg_\mu = \frac{1}{V} \dot{\xi}_\alpha \frac{\partial f_\alpha(\mathbf{g}, \vartheta, H)}{\partial g_\mu} dg_\mu = \frac{1}{V} \dot{\xi}_\alpha df_\alpha \quad (29)$$

at fixed ϑ and H . The relationship (29) is deduced by Yang et al. [16]. In view of Eq. (29), the irrotational condition (21) requires that

$$\oint \dot{\xi} \cdot d\mathbf{g} = \oint \dot{\xi}_\mu dg_\mu = 0 \quad (30)$$

at fixed ϑ and H and for any oriented simple closed curve in $\mathbf{g}$ space. The microscale irrotational condition (21) always ensures its macroscale counterpart (30), but the converse statement does not hold, since there is no one-to-one correspondence between ξ and ζ . Therefore, even if the microscale irrotational condition (21) does not hold as indicated by Brown [15], the macroscale irrotational condition (30) still can be resorted to some extent.

4 Thermodynamic Equilibrium Points

In this section, the concept of thermodynamic equilibrium points is introduced into the normality structure following the work of Edelen [9]. As remarked by Edelen [9], a thermodynamic system described by autonomous ordinary differential equations is said to be thermodynamically admissible if and only if the system has a equilibrium point which is asymptotically stable. The conjugate force $\mathbf{f}$ space is a n -dimensional Euclidean space $\mathbb{R}^n$, and the equilibrium point is a point in the space

$$\mathbf{f}^0 = \mathbf{f}^0(\vartheta, H) = \{f_1^0, f_2^0, \dots, f_n^0\} \quad (31)$$

The thermodynamic system to be asymptotically stable about $\mathbf{f}^0$ is such that the system will wind up at $\mathbf{f}^0$ as t tends to infinity if the system starts out at $t=0$ at any point in some finite neighborhood of $\mathbf{f}^0$ (see [3,9]). With the equilibrium point, the intrinsic dissipation inequality (11) should be modified as (see e.g., [3])

$$\sigma = \frac{1}{\vartheta V} (f_\alpha - f_\alpha^0) \dot{\xi}_\alpha \geq 0 \quad \text{or} \quad \sigma = \frac{1}{\vartheta V} (\mathbf{f} - \mathbf{f}^0) \cdot \dot{\xi} \geq 0 \quad (32)$$

which leads to

$$\sigma = 0 \quad \text{at} \quad \mathbf{f} = \mathbf{f}^0 \quad (33)$$

Furthermore, since σ is the entropy production of a physical system, it has an absolute minimum at $\mathbf{f}^0$, and hence, we must have

$$\frac{\partial \sigma}{\partial \mathbf{f}} = \mathbf{0} \quad \text{at} \quad \mathbf{f} = \mathbf{f}^0 \Rightarrow \dot{\xi} = \mathbf{0} \quad \text{at} \quad \mathbf{f} = \mathbf{f}^0 \quad (34)$$

With the equilibrium point, Eq. (5) should be modified as follows:

$$\frac{1}{V} (f_\alpha - f_\alpha^0) d\xi_\alpha = -d^p \phi = d^p \psi \quad (35)$$

which leads to, due to Eq. (32),

$$\vartheta \sigma = \frac{1}{V} (f_\alpha - f_\alpha^0) \dot{\xi}_\alpha = -\frac{d^p \phi}{dt} = \frac{d^p \psi}{dt} \quad (36)$$

The flow potential should be modified as

$$Q = Q(\mathbf{f}, \vartheta, H) = \frac{1}{V} \int_{\mathbf{f}^0}^{\mathbf{f}} \dot{\xi}_\alpha df_\alpha \quad (37)$$

Therefore, the entropy production function can be determined uniquely by the flow potential functions

$$\sigma(\mathbf{f}, \vartheta, H) = \frac{1}{\vartheta} (f_\alpha - f_\alpha^0) \frac{\partial Q(\mathbf{f}, \vartheta, H)}{\partial f_\alpha} \quad (38)$$

and conversely (see e.g., [9])

$$Q(\mathbf{f}, \vartheta, H) = \vartheta \int_0^1 \sigma[\tau \mathbf{f} + (1 - \tau) \mathbf{f}^0, \vartheta, H] \frac{d\tau}{\tau} \quad (39)$$

Evidently, the thermodynamics of Rice [4,5] is just a special case with $\mathbf{f}^0 = \mathbf{0}$. In view of Eqs. (34) and (39), we have $Q \geq 0$ and

$$Q = 0, \quad \frac{\partial Q}{\partial \mathbf{f}} = \mathbf{0} \quad \text{at} \quad \mathbf{f} = \mathbf{f}^0 \quad (40)$$

These properties are the same as those of σ . Especially, both Q and σ have their absolute minimums at $\mathbf{f}^0$. Eqn. (38) implies that

$$(f_\alpha - f_\alpha^0) \frac{\partial Q}{\partial f_\alpha} = (f_\alpha - f_\alpha^0) \frac{\partial Q}{\partial (f_\alpha - f_\alpha^0)} \geq 0 \quad (41)$$

Thus, the flow potential Q is a non-negative and monotonic increasing function along any ray originating from $\mathbf{f}^0$. In other words, the intrinsic dissipation inequality is guaranteed by the monotonicity of Q , so the monotonicity is an essential property of Q .

5 Intrinsic Dissipation for Rice Fluxes

The intrinsic dissipation inequality (11) or its generalized form (32) is required to hold globally and not locally. In other words, the intrinsic dissipation inequality is not necessary to hold for each microscale internal variable. Let us consider the time-independent kinetic rate laws (17). Since $\dot{\xi}_\alpha = 0$ for the whole range $f_\alpha^L < f_\alpha < f_\alpha^U$, the range is the set of thermodynamic equilibrium points of f_α . Then the time-independent kinetic rate laws (17) can be condensed into

$$(f_\alpha - f_\alpha^0) \dot{\xi}_\alpha \geq 0, \quad f_\alpha^L < f_\alpha < f_\alpha^U \quad (\alpha = 1, 2, \dots, n) \quad (42)$$

where $f_\alpha = f_\alpha^L$ or $f_\alpha = f_\alpha^U$. It implies that the intrinsic dissipation inequality holds locally. With the locality, the intrinsic dissipation inequality (32) is fulfilled automatically and can be recast as

$$(f_\alpha - f_\alpha^0) \dot{\xi}_\alpha \geq 0 \quad \text{or} \quad (\mathbf{f} - \mathbf{f}^0) \cdot \dot{\xi} \geq 0 \quad (43)$$

Let us consider the locality of intrinsic dissipation for time-dependent behavior, similar to Eq. (42),

$$(f_\alpha - f_\alpha^0) \dot{\xi}_\alpha \geq 0 \quad (\alpha = 1, 2, \dots, n) \quad (44)$$

where $f_1^0, f_2^0, \dots, f_n^0$ constitute a thermodynamic equilibrium point $\mathbf{f}^0$ as given by Eq. (31). The global intrinsic dissipation inequality (32) is then ensured by the locality (44). Unlike time-independent behavior, it is assumed that there exists only one current thermodynamic equilibrium point $\mathbf{f}^0(\vartheta, H)$ for a time-dependent thermodynamic system. Evidently, the locality of intrinsic dissipation (44) is ensured by the following condition:

$$\dot{\xi}_\alpha \begin{cases} > 0 & \text{if } f_\alpha > f_\alpha^0 \\ = 0 & \text{if } f_\alpha = f_\alpha^0 \\ < 0 & \text{if } f_\alpha < f_\alpha^0 \end{cases} \quad (\alpha = 1, 2, \dots, n) \quad (45)$$

which is just the time-dependent counterpart of Eq. (17).

A thermodynamic process always involves many elementary processes. For example, each internal variable ξ_α corresponds to an elementary process in the normality structure. In general, the elementary processes are coupled, and the entire process is termed *complex* by Ziegler [6], e.g., the one defined by Eq. (18). If the elementary processes are uncoupled or independent, the entire process is termed *compound* by Ziegler [6], e.g., the one defined by Eq. (12). Thus, the original normality structure of Rice [4] is only able to describe compound processes. Since most real thermodynamic processes are complex ones, the original normality structure can only provide approximate descriptions to those real processes. Furthermore, the locality of intrinsic dissipation (42) or (44) is also a natural result of the independence or uncoupling of the elementary processes. Therefore, the independence of the elementary processes should be taken as the more fundamental postulate of the original normality structure of Rice [4], rather than the usual statement that the each internal variable rate is stress dependent only via its conjugate thermodynamic force.

6 Convexity of Flow Potential Functions

If the flow potential Q is convex with respect to $\mathbf{f}$, its Hessian matrix $\mathbf{H}$ is positive definite (see e.g., [17]). The Hessian matrix of the flow potential Q in $\mathbf{f}$ space is denoted by

$$\mathbf{H}(Q, \mathbf{f}) = H_{\alpha\beta} = \frac{\partial^2 Q}{\partial f_\alpha \partial f_\beta} \quad (46)$$

which is symmetric following its definition, i.e., $H_{\alpha\beta} = H_{\beta\alpha}$ or $\mathbf{H} = \mathbf{H}^T$. Evidently, the Hessian matrices for the Rice fluxes are diagonal ones.

Yang et al. [8] have shown that the homogeneity of the kinetic rate laws can lead to the convexity of the flow potentials. If $\dot{\xi}$ are homogeneous functions of degree q in $\mathbf{f}$

$$\frac{\partial \dot{\xi}_\alpha}{\partial f_\beta} f_\beta = q \dot{\xi}_\alpha \quad (47)$$

the nonlinear phenomenological equations emerges,

$$\dot{\xi}_\alpha = L_{\alpha\beta} f_\beta, \quad L_{\alpha\beta} = \frac{1}{q} H_{\alpha\beta} \quad (48)$$

For the Rice fluxes, Eq. (47) reduces to

$$\frac{\partial \dot{\xi}_\alpha}{\partial f_\alpha} f_\alpha = q \dot{\xi}_\alpha, \quad (\alpha = 1, 2, \dots, n) \quad (49)$$

Therefore, if the homogeneity and nonlinear phenomenological equation, Eqs. (47) and (48), hold true, the convexity is directly required by the intrinsic dissipation inequality (11). As shown by Yang et al. [8,14], the homogeneity directly leads to the orthogonality condition of Ziegler [6]. Using integration by parts and the homogeneity (47), the flow potential Q defined in Eq. (13) can be recast as

$$Q(\mathbf{f}, \vartheta, H) = \frac{1}{V} \int_0^f \dot{\xi}_\alpha df_\alpha = \frac{1}{V} \dot{\xi}_\alpha f_\alpha - \frac{q}{V} \int_0^f \dot{\xi}_\alpha df_\alpha \quad (50)$$

which leads to, due to Eq. (11),

$$Q = \frac{\vartheta}{q+1} \sigma \quad (51)$$

With the linear relation (51), the normality structures at the microscale or macroscale, Eqs. (14) and (28), reduce to the orthogonality condition of Ziegler [6]. However, the homogeneity is only a sufficient condition leading to the convexity. Let us address the necessary and sufficient condition. Note that

$$d\dot{\xi}_\alpha = \frac{\partial \dot{\xi}_\alpha}{\partial f_\beta} df_\beta = H_{\alpha\beta} df_\beta \quad (52)$$

at fixed ϑ and H . If the flow potential Q is a convex function, the convexity implies that

$$df_\alpha d\dot{\xi}_\alpha > 0 \quad (53)$$

holds for any df_α . This inequality is the necessary and sufficient condition for the convexity of flow potential functions of the Onsager fluxes. Evidently, monotonic increasing kinetic rate laws can ensure the inequality for the Rice fluxes. In general, the thermodynamic admissible Rice fluxes with a convex flow potential should satisfy

$$\frac{\partial \dot{\xi}_\alpha}{\partial f_\alpha} > 0, \quad \dot{\xi}_\alpha(f_\alpha^0, \vartheta, H) = 0, \quad (\alpha = 1, 2, \dots, n) \quad (54)$$

Evidently, Eq. (45) embodies Eq. (54) as a special case. The convexity in $\mathbf{f}$ -space generally cannot lead to the convexity in stress space. Note that, in view of Eq. (15),

$$d\dot{\epsilon}^p = \frac{\partial^2 Q}{\partial \sigma^2} : d\sigma \quad (55)$$

at fixed ϑ and H . Thus, the necessary and sufficient condition for the convexity of flow potential functions in the stress space requires that

$$d\sigma : d\dot{\epsilon}^p > 0 \quad (56)$$

holds for any $d\sigma$. The condition is quite similar to Drucker's inequality [18] if replacing $\dot{\epsilon}^p$ with ϵ^p .

Rice [5] has shown that the flow potential is convex in stress space if the local rates are monotonic increasing functions and the conjugate forces are linear in σ . It is quite easy to prove the special case. The assumed monotonicity ensures the inequality (53), and in turn, the convexity in $\mathbf{f}$ -space. The assumed linearity is

$$f_\alpha = f_\alpha^0 + \mathbf{f}'_\alpha : \sigma, \quad (\alpha = 1, 2, \dots, n) \quad (57)$$

which leads to, in view of Eq. (15),

$$\dot{\epsilon}^p = \frac{1}{V} \dot{\xi}_\alpha \mathbf{f}'_\alpha \quad (58)$$

With the two relations, it is easy to show that

$$d\sigma : d\dot{\epsilon}^p = \frac{1}{V} df_\alpha d\dot{\xi}_\alpha \quad (59)$$

Therefore, the convexity in $\mathbf{f}$ -space can be transferred to the stress space for the special case.

7 Thermodynamic Significance of Convex Flow Potentials

As discussed in Sec. 4, the flow potential Q is a non-negative and monotonic increasing function along any ray originating from $\mathbf{f}^0$, in order to ensure the intrinsic dissipation inequality. In general, the entropy production function σ does not possess the monotonicity since the following direction derivative is not ensured to be always positive,

$$(f_\alpha - f_\alpha^0) \frac{\partial \sigma}{\partial f_\alpha} = \sigma + \frac{1}{\vartheta} (f_\alpha - f_\alpha^0) \frac{\partial^2 Q}{\partial f_\alpha \partial f_\beta} (f_\beta - f_\beta^0) \quad (60)$$

Evidently, the entropy production function σ is also a monotonic increasing function along any ray originating from $\mathbf{f}^0$ if the flow potential Q is a convex function.

One set of prescribed Onsager fluxes $\dot{\xi}$ with one equilibrium point $\mathbf{f}^0$, as a n -dimensional vector field in $\mathbf{f}$ -space, ensures a scalar flow potential field in the same space. In the following part, ϑ and H are always considered to be fixed, and then the flow potential can be denoted by $Q(\mathbf{f})$ without a loss of generality. The global intrinsic dissipation inequality should hold at any point in $\mathbf{f}$ -space,

$$(\mathbf{f} - \mathbf{f}^0) \cdot \dot{\xi} \geq 0, \quad \dot{\xi} = \frac{\partial Q}{\partial \mathbf{f}} \quad \forall \mathbf{f} \in \mathbb{R}^n \quad (61)$$

If the equilibrium point has a deviation, e.g., from $\mathbf{f}^0$ to $\mathbf{f}^*$, for the prescribed flux field $\dot{\xi}$, the global intrinsic dissipation inequality,

$$(\mathbf{f} - \mathbf{f}^*) \cdot \dot{\xi} \geq 0, \quad \dot{\xi} = \frac{\partial Q}{\partial \mathbf{f}} \quad (62)$$

cannot be fully ensured. In this paper, $\mathbf{f}^*$ is termed a *quasi-equilibrium point*. Can the violation points for the inequality (62) be confined within a limited area in $\mathbf{f}$ -space?

It is true for the Rice fluxes with the locality condition (45). If the local intrinsic dissipation (44) holds, it still holds for the movement from f_α^0 to f_α^* except for the range between f_α^0 and f_α^* in view of Eq. (45), i.e.,

$$(f_\alpha - f_\alpha^*) \dot{\xi}_\alpha \begin{cases} < 0 & \text{if } f_\alpha^0 < f_\alpha < f_\alpha^* \\ \geq 0 & \text{otherwise} \end{cases} \quad (\alpha = 1, 2, \dots, n) \quad (63)$$

or

$$(f_\alpha - f_\alpha^*) \dot{\xi}_\alpha \begin{cases} < 0 & \text{if } f_\alpha^* < f_\alpha < f_\alpha^0 \\ \geq 0 & \text{otherwise} \end{cases} \quad (\alpha = 1, 2, \dots, n) \quad (64)$$

Globally, the range from $\mathbf{f}^0$ to $\mathbf{f}^*$ constitutes a n -dimensional box in $\mathbf{f}$ -space. Based on the local dissipation (63) and (64), the global intrinsic dissipation inequality (62) still holds outside the box and may be violated within the box. In other words, the possible violation area is a n -dimensional box determined by $\mathbf{f}^0$ and $\mathbf{f}^*$.

It is also true to the Onsager fluxes with a convex flow potential. Evidently, the convexity of the flow potential leads to

$$(\mathbf{f} - \mathbf{f}^*) \cdot \dot{\xi} \geq 0 \quad \text{if } Q(\mathbf{f}^*) \leq Q(\mathbf{f}) \quad (65)$$

The property requires that for a specific quasi-equilibrium point $\mathbf{f}_0^*$, the corresponding possible violation area is confined by $Q(\mathbf{f}) < Q(\mathbf{f}_0^*)$ and denoted by the set K_0 . With the monotonicity of Q and $Q(\mathbf{f}^0) = 0$, the possible violation area is just a finite neighborhood of the equilibrium point. Thus, both the locality of intrinsic dissipation and convexity of the flow potential provide the thermodynamic system a capability against the disturbance of the thermodynamic equilibrium points.

The point of view adopted here for time-independent behavior is that the set of plastically admissible stresses determined by yield conditions just corresponds to the set of thermodynamic equilibrium points. The point of view has been adopted by Eq. (42). Thus, the global intrinsic dissipation inequality takes the form,

$$(\mathbf{f} - \mathbf{f}^0) \cdot \dot{\xi} \geq 0 \quad \forall \mathbf{f}^0 \in K_1 \quad (66)$$

where K_1 is determined by the yield conditions and $\mathbf{f}$ is on the boundary of K_1 or the yield surfaces. The convexity of the yield surfaces is required by the inequality (66). Evidently, the inequality (66) is just thermodynamic counterpart of the principle of maximum plastic dissipation (see, e.g., [18,19]). This thermodynamic basis can well explain why the principle can apply to both hardening and softening materials, but Drucker's inequality [20] can only apply to hardening materials.

Ziegler [6] takes the entropy production function σ as the fundamental potential function. Rice [4] or its extended version presented in this paper takes the flow potential function Q as the fundamental one. In view of Eqs. (38) and (39), the two viewpoints are equivalent to each other. In this sense, the two theories possess the same physical basis, the microscale irrotational condition (20) or (21). Various thermodynamic principles mentioned in this paper may be stated in terms of the relationship between Q and σ . The nonlinear Onsager reciprocal relations are implied if Q and σ are related by Eq. (38). The principle of maximum plastic dissipation holds if Q and σ possess the same monotonicity. The linear relation between Q and σ , Eq. (51), leads to the orthogonality condition. The orthogonality condition reduces to the Onsager reciprocal relations if $q=1$ in Eq. (51). The thermodynamic restrictions on the microscale fluxes are listed according to their generality,

$$f_\alpha \dot{\xi}_\alpha \geq 0 \quad (67a)$$

$$\oint \dot{\xi}_\alpha df_\alpha = 0 \quad (67b)$$

$$df_\alpha d\dot{\xi}_\alpha > 0 \quad (67c)$$

$$\frac{\partial \dot{\xi}_\alpha}{\partial f_\beta} f_\beta = q \dot{\xi}_\alpha \quad (67d)$$

These restrictions are localized to each internal variable for the Rice fluxes. As a classifying postulate, the convexity condition, Eq. (67c), should possess the same status as the irrotational condition and orthogonality condition. Among them, the orthogonality condition is the strictest restriction and of the least generality.

8 Concluding Remarks

Enriched by the nonlinear Onsager reciprocal relations and thermodynamic equilibrium points, an extended normality structure is established in this paper as a unified nonlinear thermodynamic theory of solids. It is revealed that the normality structure stems from the microscale irrotational thermodynamic fluxes or Onsager fluxes. The original normality structure corresponds to the thermodynamic processes whose elementary processes are independent or uncoupled. Within the extended normality structure, the convexity condition for the flow potential is that the conjugate force increment cannot not oppose the increment of the rates of local internal variables. For the Rice fluxes, the convexity condition reduces to the local rates being monotonic increasing functions with respect to their conjugate forces. The convexity of the flow potential provides the thermodynamic system a capability against the disturbance of the thermodynamic equilibrium point.

The point of view adopted here for time-independent behavior is that the thermodynamic counterpart of the set of plastically admissible stresses determined by yield conditions is just the set of thermodynamic equilibrium points. Based on the viewpoint, Rice's kinetic rate laws for time-independent behavior is the requirement that the intrinsic dissipation inequality should hold locally. Furthermore, the global intrinsic dissipation inequality is just thermodynamic counterpart of the principle of maximum plastic dissipation and requires the convexity of the yield surfaces.

Acknowledgment

The work is supported by Project No. 50579026 of National Natural Science Foundation of China and subsidized by the special funds for major state basic research projects with Grant No. 2002cb412708, and supported by the Cultivation Fund of the Key Scientific and Technical Innovation Project, Ministry of Education of China (Grant No. 704003).

References

- [1] Onsager, L., 1931, "Reciprocal Relations in Irreversible Processes—I," *Phys. Rev.*, **37**, pp. 405–406.
- [2] Onsager, L., 1931, "Reciprocal Relations in Irreversible Processes—II," *Phys. Rev.*, **38**, pp. 2265–2279.
- [3] De Groot, S. R., and Mazur, P., 1962, *Non-Equilibrium Thermodynamics*, North-Holland, Amsterdam.
- [4] Rice, J. R., 1971, "Inelastic Constitutive Relations for Solids: An Integral Variable Theory and its Application to Metal Plasticity," *J. Mech. Phys. Solids*, **19**, pp. 433–455.
- [5] Rice, J. R., 1975, "Continuum Mechanics and Thermodynamics of Plasticity in Relation to Microscale Deformation Mechanisms," *Constitutive Equations in Plasticity*, Argon, A. S., ed., MIT Press, Cambridge, MA, pp. 23–79.
- [6] Ziegler, H., 1977, *An Introduction to Thermomechanics*, North-Holland, Amsterdam.
- [7] Edelen, D. G. B., 1972, "A Nonlinear Onsager Theory of Irreversibility," *Int. J. Eng. Sci.*, **10**, pp. 481–490.
- [8] Yang, Q., Tham, L. G., and Swoboda, G., 2005, "Normality Structures With Homogeneous Kinetic Rate Laws," *ASME J. Appl. Mech.*, **72**, pp. 322–329.
- [9] Edelen, D. G. B., 1973, "Asymptotic Stability, Onsager Fluxes and Reaction Kinetics," *Int. J. Eng. Sci.*, **11**, pp. 819–839.
- [10] Collins, I. F., and Kelly, P. A., 2002, "A Thermomechanical Analysis of a Family of Soil Models," *Géotechnique*, **52**, pp. 507–518.
- [11] Collins, I. F., 2005, "Elastic/plastic Models for Soils and Sands," *Int. J. Mech. Sci.*, **47**, pp. 493–508.
- [12] Houlsby, G. T., and Puzrin, A. M., 2000, "A Thermomechanical Framework for Constitutive Models for Rate-Independent Dissipative Materials," *Int. J. Plast.*, **16**, pp. 1017–1047.
- [13] Rajagopal, K. R., and Srinivasa, A. R., 2004, "On Thermomechanical Restrictions of Continua," *Proc. R. Soc. London, Ser. A*, **460**, pp. 631–651.
- [14] Yang, Q., Chen, X., and Zhou, W. Y., 2005, "Thermodynamic Relationship

- Between Creep Crack Growth and Creep Deformation," *J. Non-Equil. Thermodyn.*, **30**, pp. 81–94.
- [15] Brown, L. M., 1997, "Transition From Laminar to Rotational Motion in Plasticity," *Philos. Trans. R. Soc. London, Ser. A*, **355**, pp. 1979–1990.
- [16] Yang, Q., Chen, X., and Zhou, W. Y., 2006, "Multiscale Thermodynamic Significance of the Scale Invariance Approach in Continuum Inelasticity," *ASME J. Eng. Mater. Technol.*, **128**, pp. 125–132.
- [17] Maugin, G. A., 1999, *The Thermodynamics of Nonlinear Irreversible Behaviors*, World Scientific, Singapore.
- [18] Hill, R., 1950, *The Mathematical Theory of Plasticity*, Clarendon Press, Oxford.
- [19] Lubliner, H., 1990, *Plasticity Theory*, Macmillan, New York.
- [20] Drucker, D. C., 1959, "A Definition of Stable Inelastic Material," *ASME J. Appl. Mech.*, **26**, pp. 101–106.

M. Tosun¹
e-mail: tosun@sakarya.edu.tr

M. A. Gungor

Department of Mathematics,
Faculty of Arts Sciences,
Sakarya University,
Sakarya, 54187 Turkey

I. Okur

Department of Physics,
Faculty of Arts Sciences,
Sakarya University,
Sakarya, 54187 Turkey

On the One-Parameter Lorentzian Spherical Motions and Euler-Savary Formula

In this paper, we have introduced one-parameter Lorentzian spherical motion. In addition to that, we have given the relations between the absolute, relative, and sliding velocities of these motions. Furthermore, the relations between fixed and moving pole curves in the Lorentzian spherical motions have also been obtained. At the end of this study, we have expressed the Euler-Savary formula for the one-parameter Lorentzian spherical motions. [DOI: 10.1115/1.2722775]

Keywords: Lorentzian geometry, Lorentzian motions, Euler-Savary formula

1 Introduction

The determination of a point or a set of points such that its velocity norm vanishes or that is a minimum has always aroused interest among kinematicians. The explanation of this is twofold: points whose velocity, or acceleration, vanishes are important, for they allow one to write simplified equations for the velocity and acceleration of any other point of the rigid body; and a point or a set of points with a minimum velocity norm locates the connecting place of a kinematics pair, in general, a helicoidal pair, that connects the rigid body to the reference body. This connection produces a motion with the same characteristics, at least up to the first derivative of the original motion of the rigid body. Indeed, the search for points of a rigid body with a minimum velocity norm has led to the description of the velocity of a rigid body in terms of infinitesimal screws, or helicoidal fields, and therefore to the definition of the instantaneous screw axis.

By taking Lorentzian plane L^2 instead of Euclidean plane E^2 , Ergin [1] has introduced one-parameter planar motion in Lorentzian plane. Furthermore, he gave the relation between the velocities, accelerations, and pole curves of these motions. In Ref. [2], Ikawa gave the Euler-Savary formula in Lorentzian plane L^2 .

In spherical kinematics, Bottema [3] found the existence of three acceleration axes, straight lines that pass through the fixed point, whose points do not possess acceleration. Furthermore, Bottema found conditions that determine whether all three acceleration axes are real or only one is real. Later, Meyer Zur Capellen and Dittrich [4] extended Bottema's results by studying the acceleration distribution in a rigid body subjected to spherical motion.

Finally, in spatial kinematics Beyer [5] completed a thorough analysis of the acceleration distribution in a rigid body. Moreover, Beyer formulated a system of linear equations in a special coordinate system, whose solution leads to the location of the acceleration center. However, Beyer did not obtain a closed-form solution even for this special coordinate system. Finally, Beyer mentioned a previous work by Schell for the analysis of the special cases associated with the determination of the acceleration center.

Bottema and Roth [6] obtained the equation for the acceleration distribution in a rigid body in a special coordinate system. By equating the acceleration to the zero vector and solving the result-

ing system, the coordinates of the acceleration center are obtained. The coordinate system used by Beyer [5] has some similarities with that employed by Bottema and Roth.

Considering one and two parameters spherical motions in Euclidean space, Muller [7] has given the relations for absolute, sliding, relative velocities, and pole curves of these motions. In addition to that, he has expressed the corresponding formula of Euler-Savary formula related to the trajectory curves of these one-parameter spherical motions.

To investigate the geometry of the motion of a line or a point in the motion of space is important in the study of space kinematics or spatial mechanisms or in physics. The geometry of such a motion of a point or a line has a number of applications in geometric modeling and model-based manufacturing of the mechanical products or in the design of robotic motions. These are specifically used to generate geometric models of shell-type objects and thick surfaces, [8–10].

The Euler-Savary theorem is a well-known theorem that is used in serious fields of study in engineering and mathematics.

Pennock and Sankaranarayanan presented a graphical technique to locate the center of curvature of the path traced by a coupler point of a planar, single-degree-of-freedom, geared seven-bar mechanism in [11]. In this paper, the center of curvature of the path traced by an arbitrary coupler point can be obtained from existing techniques, such as the Euler-Savary equation. A schematic drawing of this geared seven-bar mechanism is shown in Fig. 1.

In the other work, Pennock and Raje expressed a graphical technique to obtain the radius of curvature of the path traced by a coupler point of a planar single-degree of freedom, indeterminate eight-bar linkage commonly referred to as the double flier linkage. In this study, the radius of curvature of the path traced by the coupler point is obtained from Euler-Savary equation [12]. A schematic drawing of a general geometry double flier linkage is shown in Fig. 2.

This paper is organized as follows. In this first part, basic concepts have been given in Minkowski space $\mathbf{R}_1^3$. In the second part, one-parameter Lorentzian spherical motions are defined. In doing so, the orthonormal frames of $\{O; \vec{e}_1, \vec{e}_2, \vec{e}_3\}$ and $\{O; \vec{e}'_1, \vec{e}'_2, \vec{e}'_3\}$ are taken representing moving Lorentzian sphere S_1^2 and fixed Lorentzian sphere S_2^2 , respectively. Without making any of these privileged, we have taken another orthonormal frame $\{O; \vec{r}_1, \vec{r}_2, \vec{r}_3\}$, called relative orthonormal frame, and given the Lorentzian spherical motions with respect to this new (relative) orthonormal frame. Furthermore, the relations between absolute, relative, and sliding velocities of one-parameter Lorentzian spherical motions have been obtained. In the third part, the rela-

¹Corresponding author.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received July 13, 2005; final manuscript received January 17, 2007. Review conducted by N. Sri Namachivaya.

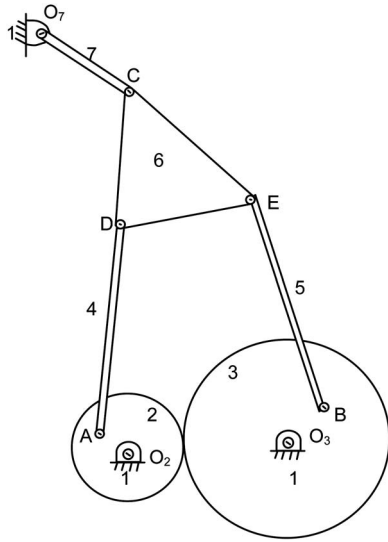


Fig. 1 The geared seven-bar mechanism

tions between the pole curves rolling on each other with respect to a spherical relative system have also been given. In the last part of this study, the Euler-Savary formula corresponding to the trajectory curves is obtained. We hope that these results will contribute to the study of space kinematics and physics applications.

2 Preliminaries

We start with preliminaries on the geometry of three-dimensional Minkowski space. Let $\mathbf{R}_1^3$ be a three-dimensional Minkowski space endowed with Lorentzian inner product g of signature $(-, +, +)$. A vector $\vec{X} = (x_1, x_2, x_3)$ of $\mathbf{R}_1^3$ is said to be timelike if $g(\vec{X}, \vec{X}) < 0$, spacelike if $g(\vec{X}, \vec{X}) > 0$ and lightlike (or null) if $g(\vec{X}, \vec{X}) = 0$. The set of all vector $\vec{X}$ such that $g(\vec{X}, \vec{X}) = 0$ is called the lightlike (or null) cone and is denoted by $\wedge$. The norm of a vector $\vec{X}$ is defined to be $\|\vec{X}\| = \sqrt{|g(\vec{X}, \vec{X})|}$. Time orientation is defined as follows: A timelike vector $\vec{X} = (x_1, x_2, x_3)$ is future pointing (respectively, past pointing) if and only if $x_1 > 0$ (respectively, $x_1 < 0$) [13]. Let $\vec{X}$ be a future pointing timelike unit vector and $\vec{Y}$ also be a future pointing timelike unit vector. If the angle between $\vec{X}$ and $\vec{Y}$ is θ , then we may have [13,14]

$$g(\vec{X}, \vec{Y}) = -\cosh \theta$$

The Lorentzian sphere and hyperbolic sphere of radius 1 in $\mathbf{R}_1^3$ are given by

$$S_1^2 = \{\vec{X} = (x_1, x_2, x_3) \in \mathbf{R}_1^3 | g(\vec{X}, \vec{X}) = 1\}$$

and

$$H_0^2 = \{\vec{X} = (x_1, x_2, x_3) \in \mathbf{R}_1^3 | g(\vec{X}, \vec{X}) = -1\}$$

respectively, [15].

H_0^2 consists of two connected components. The components of H_0^2 through $(1, 0, 0)$ and $(-1, 0, 0)$ are called the future-pointing hyperbolic unit sphere and past-pointing hyperbolic unit sphere and are denoted by H_0^{+2} and H_0^{-2} , respectively.

As in the case of Euclidean three-dimensional space, the Lorentzian cross product of $\vec{X}$ and $\vec{Y}$ is defined by

$$\vec{X} \wedge \vec{Y} = (y_2x_3 - y_3x_2, y_1x_3 - y_3x_1, y_2x_1 - y_1x_2)$$

where $\vec{X} = (x_1, x_2, x_3)$ and $\vec{Y} = (y_1, y_2, y_3)$ are the vectors of the space $\mathbf{R}_1^3$ [16].

The matrix

$$A(\theta) = \begin{bmatrix} \cosh \theta & \sinh \theta \\ \sinh \theta & \cosh \theta \end{bmatrix}$$

is called the Lorentzian rotation matrix in $\mathbf{R}_1^2$, where $\theta \in \mathbf{R}$ [14]. This matrix is similar to the rotation matrix, which is

$$\begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$$

in E^2

LEMMA 1. *Timelike vectors are transformed to timelike vectors and spacelike vectors are transformed to spacelike vectors by A. That is, A conserves the orientation [13].*

3 Lorentzian Spherical Motions and Their Velocities

Let S_1^2 and $\bar{S}_1^2$ be O -centered moving and fixed Lorentzian spheres, and related to these spheres $\{O; \vec{e}_1, \vec{e}_2, \vec{e}_3\}$ and $\{O; \vec{e}'_1, \vec{e}'_2, \vec{e}'_3\}$ be orthonormal coordinate frames moving related to each other, having the same center O , respectively. Let assume that $\{O; \vec{e}_1, \vec{e}_2, \vec{e}_3\}$ represents the moving Lorentzian sphere S_1^2 , whereas $\{O; \vec{e}'_1, \vec{e}'_2, \vec{e}'_3\}$ represents the fixed one (where base vectors $\vec{e}_2, \vec{e}_3; \vec{e}'_2, \vec{e}'_3$ are spacelike and the vectors $\vec{e}_1, \vec{e}'_1$ are timelike). Therefore,

$$g(\vec{e}_i, \vec{e}_j) = g(\vec{e}'_i, \vec{e}'_j) = \varepsilon_i \delta_{ij},$$

$$\varepsilon_i = \begin{cases} 1, & \text{are } \vec{e}_i \text{ or } \vec{e}'_i \text{ spacelike} \\ -1, & \text{are } \vec{e}_i \text{ or } \vec{e}'_i \text{ timelike} \end{cases} \quad 1 \leq i, j \leq 3$$

Adopting that none of these systems are privileged, we take another relative orthonormal frame, $\{O; \vec{r}_1, \vec{r}_2, \vec{r}_3\}$, in consideration and express the movement with respect to this relative one (where base vectors $\vec{r}_2, \vec{r}_3$ are spacelike and the vectors $\vec{r}_1$ is timelike). Therefore,

$$g(\vec{r}_i, \vec{r}_j) = \varepsilon_i \delta_{ij}, \varepsilon_i = \begin{cases} 1, & \text{is } \vec{r}_i \text{ spacelike} \\ -1, & \text{is } \vec{r}_i \text{ timelike} \end{cases}, \quad 1 \leq i, j \leq 3$$

Since each of these orthonormal frames has the same orientation, one frame is obtained by using another when rotated about O -point. Let A be an unique orthogonal Lorentzian matrix. That is, $A' = \varepsilon A^{-1} \varepsilon$, where ε is a sign matrix.

If we use the following abbreviations:

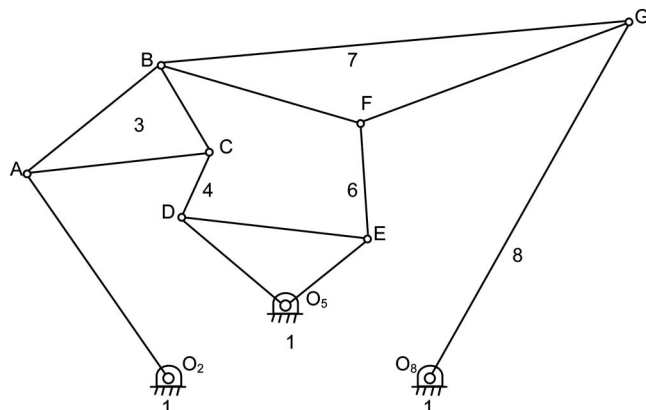


Fig. 2 Schematic diagram of the double-flier eight-bar linkage

$$E = \begin{bmatrix} \vec{e}_1 \\ \vec{e}_2 \\ \vec{e}_3 \end{bmatrix}, \quad R = \begin{bmatrix} \vec{r}_1 \\ \vec{r}_2 \\ \vec{r}_3 \end{bmatrix}, \quad E' = \begin{bmatrix} \vec{e}'_1 \\ \vec{e}'_2 \\ \vec{e}'_3 \end{bmatrix}$$

we get

$$R = AE, \quad R = A'E' \quad (1)$$

Here, the elements of the matrix A are not only continuous but all differentiable as well as we would like. Hence, one-parameter motion is determined by the matrix $A=A(t)$ and called as one-parameter Lorentzian spherical motion D_1 .

Now, let us calculate the differentials of vectors $\vec{r}_j$ with respect to S_1^2 and $\bar{S}_1^2$, respectively. If we consider Eq. (1), then differential of the relative orthonormal coordinate frame R with respect to S_1^2 and $\bar{S}_1^2$ are

$$dR = dAA^{-1}R, \quad d'R = dA'(A')^{-1}R \quad (2)$$

By choosing $dAA^{-1} = \Omega$ and $dA'(A')^{-1} = \Omega'$, Eq. (2) can be rewritten as follows:

$$dR = \Omega R, \quad d'R = \Omega' R \quad (3)$$

We can easily see that both Ω and Ω' matrices are anti-symmetric in the sense of Lorentzian, i.e., $\Omega' = -\varepsilon \Omega \varepsilon$ where Ω' is the transpose matrix of Ω and ε is sign matrix. Let assume that $\omega_{ij} (1 \leq i, j \leq 3)$ are the elements of Ω matrix. Let us denote the permutations of the indices $i, j, k = 1, 2, 3; 2, 3, 1; 3, 1, 2$, by $\omega_{ij} = \omega_k$. Then, we can easily get that

$$\Omega = \begin{bmatrix} 0 & \omega_3 & -\omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix} \quad (4)$$

In the similar way, anti-symmetric matrix Ω' in the sense of Lorentzian is obtained to be

$$\Omega' = \begin{bmatrix} 0 & \omega'_3 & -\omega'_2 \\ \omega'_3 & 0 & -\omega'_1 \\ -\omega'_2 & \omega'_1 & 0 \end{bmatrix} \quad (5)$$

Let

$$\vec{X} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

be a point in the relative frame and configure the following vector:

$$\vec{OX} = \vec{X} = X'R \quad (6)$$

If the point X is a point on the unit Lorentzian sphere, then we have

$$\|\vec{X}\|^2 = -x_1^2 + x_2^2 + x_3^2 = 1$$

Now, we compute the differentials of X with respect to Lorentzian spheres S_1^2 (moving) and $\bar{S}_1^2$ (fixed). First of all, we evaluate the differentiation of X with respect to moving Lorentzian sphere S_1^2 . If we consider Eq. (6), we obtain

$$d\vec{X} = dX'R + X'dR$$

Substituting Eq. (3) in the above equation, we have

$$d\vec{X} = (dX' + X'\Omega)R \quad (7)$$

Therefore, relative velocity of X (i.e., velocity of X with respect to Lorentzian sphere S_1^2) is $\vec{V}_r = d\vec{X}/dt$. If $\vec{V}_r = 0$, i.e., $d\vec{X} = 0$, then the point X is fixed in the moving Lorentzian sphere S_1^2 . Thus, from Eq. (7), the condition that the point X is fixed in S_1^2 is given by

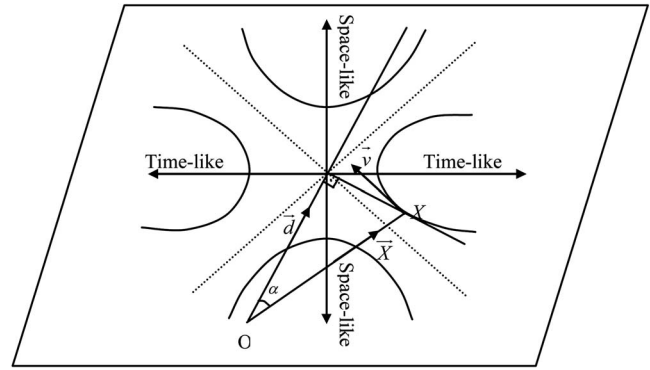


Fig. 3 Velocity vector $\vec{v}$ of point X

$$dX' = -X'\Omega \quad (8)$$

Similarly, from Eq. (3), the differential of X with respect to fixed Lorentzian sphere $\bar{S}_1^2$ is

$$d'\vec{X} = (dX' + X'\Omega')R \quad (9)$$

So, absolute velocity vector (the velocity of the point X with respect to fixed Lorentzian sphere $\bar{S}_1^2$) is $\vec{V}_a = d'\vec{X}/dt$. If $\vec{V}_a = 0$, i.e., $d'\vec{X} = 0$, the point X is fixed in the fixed Lorentzian sphere $\bar{S}_1^2$.

Hence, the condition that the point X is fixed in $\bar{S}_1^2$ is given by

$$dX' = -X'\Omega' \quad (10)$$

If the point X is fixed in moving Lorentzian sphere S_1^2 then the velocity of X with respect to $\bar{S}_1^2$ is called sliding velocity of X and denoted by $\vec{V}_f$. If Eq. (8) is substituted in (9), we get

$$\vec{V}_f = X'\Psi R \quad (11)$$

where $\Psi = \Omega' - \Omega$.

If the Pfaffian vector $\vec{\Psi}$ is taken to be

$$\vec{\Psi} = -\Psi_1\vec{r}_1 + \Psi_2\vec{r}_2 + \Psi_3\vec{r}_3, \quad \Psi_i = \omega'_i - \omega_i, \quad 1 \leq i \leq 3 \quad (12)$$

then we get

$$\vec{V}_f = \vec{\Psi} \wedge \vec{X} \quad (13)$$

Taking Eqs. (7) and (9) into account, we can easily get

$$\vec{V}_f = d'\vec{X} - d\vec{X}$$

From the above equation, we may write

$$\vec{V}_a = \vec{V}_r + \vec{V}_f$$

Therefore, we give the following theorem.

THEOREM 2. In a one-parameter Lorentzian spherical motion, absolute velocity vector of a point X is the sum of relative velocity vector and sliding velocity vector of it.

Now, to understand the meaning of the Pfaffian vector $\vec{\Psi}$ and Eq. (14), we emphasize the importance of Darboux rotation vector.

Let us consider a rotational motion about an axis. Assume that this axis passes through the origin and its direction be $\vec{d}$. We also assume that the angular velocity of this rotational motion is $\omega = \mp \|\vec{d}\|$.

Let apply this rotation motion to the point X with the position vector of $\vec{OX} = \vec{X}$ and let us define velocity vector $\vec{v}$ of this point X as follows (see Fig. 3):

$$\vec{v} = \vec{d} \wedge \vec{X}$$

The last equation implies that the vector $\vec{v}$ is orthogonal to both $\vec{X}$ and $\vec{d}$. If the angle between $\vec{d}$ and $\vec{X}$ is denoted by α and the distance of $\vec{X}$ from the rotation axis by r , then we can write, [13].

$$\|\vec{v}\| = \|\vec{d}\| \|\vec{X}\| \sinh \alpha = \mp \omega r$$

It is very clear from this equation that $\vec{v}$ is the velocity vector of the point X on the rotation about the axis $\vec{d}$ with the angular velocity of $\mp \|\vec{d}\|$. Therefore, we call $\vec{\Psi}$ Pfaffian vector as rotation vector of one-parameter Lorentzian spherical motion D_1 at the time t . Thus, we give the following theorem.

THEOREM 3. *In one-parameter Lorentzian spherical motion D_1 at the time t , for every point X there exists an infinitesimal rotational motion. In this rotational motion, Pfaffian vector plays the role of Darboux rotation vector.*

Now we add an unit vector of $\vec{p}$, which is in the direction of the rotation vector $\vec{\Psi}$. Since we have

$$\|\vec{p}\| = 1$$

then we write

$$\vec{\Psi} = \vec{p} \sqrt{-\Psi_1^2 + \Psi_2^2 + \Psi_3^2}$$

where $\Psi = \mp \|\vec{\Psi}\| = \sqrt{-\Psi_1^2 + \Psi_2^2 + \Psi_3^2}$ demonstrates the infinitesimal rotational angle, which produces the rotation in the time interval dt (the sign of Ψ depend on the direction of $\vec{p}$). The point P shown on the Lorentzian sphere ($\vec{OP} = \vec{p}$) is an instantaneous rotation pole. As the point P is characterized by that the sliding velocity is equal to zero, according to Eq. (13) if

$$\vec{\Psi} \wedge \vec{X} = 0, \quad \|\vec{X}\|^2 = 1$$

then

$$\vec{X} = \mp \vec{p}$$

THEOREM 4. *In a one-parameter Lorentzian spherical motion for any time t , there exists a couple of points P, P' for each of which the sliding velocities are zero, where P is the rotational pole S_1^2 and P' is the rotational pole $\bar{S}_1^2$. Those points remain stable on both Lorentzian spheres at any time.*

THEOREM 5. *Every point of moving Lorentzian sphere S_1^2 make a rotational motion (an instantaneous rotational motion) with angular velocity $\Psi:dt$ about the pole P (and its P' point) at every time t . Therefore, one-parameter Lorentzian spherical motion is such a rotational motion of Lorentzian sphere S_1^2 with respect to fixed Lorentzian sphere $\bar{S}_1^2$ at a time t .*

4 Canonical Relative Frames and Rolling of the Pole Curves on Each Others

Now, let us choose a special relative frame that satisfies the following:

$$\vec{p} = \vec{r}_3 \quad (14)$$

If we take $\vec{p} = \vec{r}_3$, then the vector $\vec{p}$ becomes orthogonal to $\vec{r}_1$ and $\vec{r}_2$. Therefore, since $\vec{\Psi} = \vec{p} \sqrt{-\Psi_1^2 + \Psi_2^2 + \Psi_3^2}$, from the Eq. (12) we see that $\Psi_1 = 0, \Psi_2 = 0$. Since we have $\Psi = \Omega' - \Omega$, if we consider Eqs. (4) and (5) we reach $\omega_1' = \omega_1, \omega_2' = \omega_2$. Thus, infinitesimal rotation angle of instantaneous rotation appears to be

$$\Psi = \Psi_3$$

In this case, instantaneous rotation axis is expressed as follows:

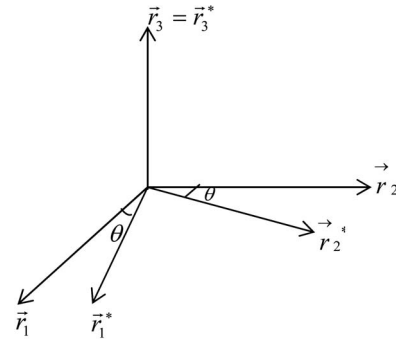


Fig. 4 Rotation about the axis $\vec{p} = \vec{r}_3$ by an angle of θ

$$\vec{\Psi} = \vec{r}_3 \Psi_3 = \vec{r}_3 (\omega_3' - \omega_3)$$

From this point on, we assume that $\Psi_3 \neq 0$. We have not given the single meaning of relative frame by using Eq. (14) because the frame obtained from the condition of $\vec{p} = \vec{r}_3$ can be rotated arbitrarily about the $\vec{r}_3$ -axis. Therefore, rotating the frames about $\vec{p} = \vec{r}_3$ -axis by an angle of θ gives us (see Fig. 4)

$$R^* = A(\theta)R \quad (15)$$

where

$$R^* = \begin{bmatrix} \vec{r}_1^* \\ \vec{r}_2^* \\ \vec{r}_3^* \end{bmatrix}, \quad R = \begin{bmatrix} \vec{r}_1 \\ \vec{r}_2 \\ \vec{r}_3 \end{bmatrix} \quad \text{and} \quad A(\theta) = \begin{bmatrix} \cosh \theta & \sinh \theta & 0 \\ \sinh \theta & \cosh \theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

This new orthonormal frame $\{O; \vec{r}_1^*, \vec{r}_2^*, \vec{r}_3^*\}$ has the following differential equations, corresponding to Eq. (3)

$$dR^* = \Omega^* R^*, \quad d'R^* = \Omega' R^* \quad (16)$$

Now we see the how we can obtain ω^* 's from ω 's, i.e., we discuss the relationship between ω 's and ω^* 's when the frame rotates by the angle of θ .

If we take into account Eq. (15), we can write

$$dR^* = dA(\theta)R + A(\theta)dR$$

Substituting Eq. (3) in the along equation, we obtain

$$dR^* = [dA(\theta) + A(\theta)\Omega]R \quad (17)$$

and using Eqs. (15) and (16) we have the following:

$$\Omega^* A(\theta) = dA(\theta) + A(\theta)\Omega \quad (18)$$

If we write Eq. (18) in matrix form we can easily see that

$$\omega_1^* = \omega_1 \cosh \theta + \omega_2 \sinh \theta$$

$$\omega_2^* = \omega_1 \sinh \theta + \omega_2 \cosh \theta$$

$$\omega_3^* = \omega_3 + d\theta$$

Thus, in this type of rotation of the frame, Pfaffian forms transform as unit vectors $\vec{r}_1$ and $\vec{r}_2$.

Now, to normalize the relative system we choose the rotation angle θ in such that

$$\omega_1^* = \omega_1 \cosh \theta + \omega_2 \sinh \theta = 0 \quad (19)$$

Eq. (19) is a conditional equation for the rotation angle θ . At this point, we suppose that the relative frame is rotated about $\vec{r}_3$ by the angle θ , which satisfies Eq. (19) and omit the asterisk. Thus, we can rewrite the Eqs. (16) and (19) for the canonical relative frame as follows. Differentiation with respect to S_1^2 is

$$\begin{bmatrix} d\vec{r}_1 \\ d\vec{r}_2 \\ d\vec{r}_3 \end{bmatrix} = \begin{bmatrix} 0 & \omega_3 & -\omega_2 \\ \omega_3 & 0 & 0 \\ -\omega_2 & 0 & 0 \end{bmatrix} \begin{bmatrix} \vec{r}_1 \\ \vec{r}_2 \\ \vec{r}_3 \end{bmatrix} \quad (20)$$

and the differentiation with respect to $\vec{S}_1^2$ is

$$\begin{bmatrix} d'\vec{r}_1 \\ d'\vec{r}_2 \\ d'\vec{r}_3 \end{bmatrix} = \begin{bmatrix} 0 & \omega'_3 & -\omega'_2 \\ \omega'_3 & 0 & 0 \\ -\omega'_2 & 0 & 0 \end{bmatrix} \begin{bmatrix} \vec{r}_1 \\ \vec{r}_2 \\ \vec{r}_3 \end{bmatrix} \quad (21)$$

$\vec{p}=\vec{r}_3$ vector draws a curve (P) on the moving sphere S_1^2 , we call this curve as moving pole curve centre of one-parameter Lorentzian movement D_1 . From Eq. (20), we have the following:

$$\frac{d\vec{r}_3}{\omega_2} = \frac{d\vec{r}_3}{ds} = -\vec{r}_1$$

This alone equation tells us that the unit tangential vector of moving pole curve (P) is $(-\vec{r}_1)$ and $\omega_2 ds$ is the arc element of (P).

In the same manner, the end point of the vector $\vec{p}=\vec{r}_3$ draws a constant pole curve (P') on the sphere $\vec{S}_1^2$. On this curve unit tangential vector at the point P is $(-\vec{r}_1)$ and arc element is $\omega_2 = ds'$ (here we took Eq. (21) into account). Thus, we can give the following theorems.

THEOREM 6. *Velocity vectors of the rotating pole (P) are the same at any time when the pole on the moving and constant sphere draw pole curves (P) and (P'), respectively.*

THEOREM 7. *In a one-parameter spherical Lorentzian movement D_1 , spherical moving pole curve (P) of S_1^2 rolls on constant pole curve (P') of $\vec{S}_1^2$ with no slide.*

THEOREM 8. *In the reverse movement of one-parameter spherical rotation motion, the spherical surfaces of S_1^2 , $\vec{S}_1^2$ and spherical pole curve (P) and (P') changes their roles.*

5 Corresponding Euler-Savary Formula for Trajectory Curves

Let point X be given on a moving Lorentzian sphere S_1^2 . This point draws a trajectory on fixed Lorentzian sphere $\vec{S}_1^2$ during the movement of one-parameter D_1 Lorentzian rotation motion. Now we search for the spherical curvature centre X' of this trajectory. Let rotation motion be given with respect to canonical relative system obeying Eq. (20). Therefore, rotation motion in time t might be thought as a small rotation of S_1^2 (moving) with respect to $\vec{S}_1^2$ (fixed) with a small rotation angle of $\Psi=\Psi_3=\omega'_3-\omega_3$ about the rotation pole P given by $\vec{OP}=\vec{p}=\vec{r}_3$. For the arc element drawn by the point X on $\vec{S}_1^2$ during D_1 motion, we can write the following (see Fig. 5):

$$ds = \sigma = \Psi \sinh \theta \quad (22)$$

The trajectory of point X on $\vec{S}_1^2$ forms a surface strip or a curvature strip on fixed Lorentzian sphere $\vec{S}_1^2$. Let us define an orthogonal frame on this trajectory of point X as follows:

- The normal vector of surface (which is normal to the surface of sphere) is given by position vector $\vec{OX}=\vec{x}$
- Unit vector $\vec{t}$ is superseded with the tangent of trajectory of point X
- Thus, unit normal vector lying inside the strip plane is

$$\vec{n} = \vec{x} \wedge \vec{t}$$

In this case, differential equations of orthogonal frame $\{\vec{x}, \vec{t}, \vec{n}\}$ are as follows:

$$d\vec{t} = -\vec{x}\sigma + \vec{n}\kappa$$

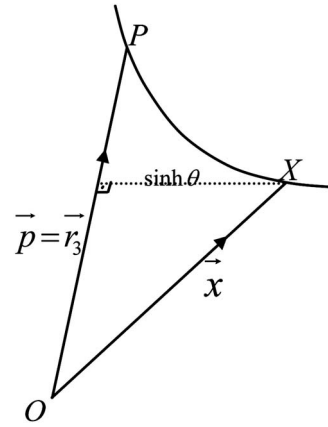


Fig. 5 Distance of point X to rotation pole P

$$d\vec{n} = \vec{t}\kappa$$

$$d\vec{x} = -\vec{t}\sigma \quad (23)$$

Here, arc element of the trajectory of point X is $\sigma=ds$ and geodesic curvature element is $\kappa=c ds$. The vectors $\vec{x}$ and $\vec{n}$ define the normal plane of the curve. This plane cuts the sphere along the trajectory normal (see Fig. 6). Thus, rotation pole P and spherical curvature centre X' of trajectory lie on this big circle. Therefore, vectors $\vec{OP}=\vec{p}$ and $\vec{OX}'=\vec{x}'$ can be given as follows (see Fig. 6):

$$\vec{p} = \vec{x} \cosh \theta + \vec{n} \sinh \theta \quad (24)$$

$$\vec{x}' = \vec{x} \cosh \rho + \vec{n} \sinh \rho \quad (25)$$

Here the spherical distance from rotation pole P to point X is θ and the spherical distance from rotation pole P to point X' is $\theta' = \theta - \rho$. Since the spherical distance of points X and X' is ρ , $\rho = \theta - \theta'$ is called as spherical curvature radius of trajectory and henceforth, by considering Eqs. (22) and (23), we write

$$\begin{aligned} d\vec{p} &= d\vec{x} \cosh \theta + \vec{x} \sinh \theta d\theta + d\vec{n} \sinh \theta + \vec{n} \cosh \theta d\theta \\ &= -\vec{t}\sigma \cosh \theta + \vec{x} \sinh \theta d\theta + \vec{t}\kappa \sinh \theta + \vec{n} \cosh \theta d\theta \\ &= \vec{t}(\kappa \sinh \theta - \sigma \cosh \theta) + (\vec{n} \cosh \theta + \vec{x} \sinh \theta) d\theta \end{aligned} \quad (26)$$

In Eq. (26) we substitute ρ into θ to find

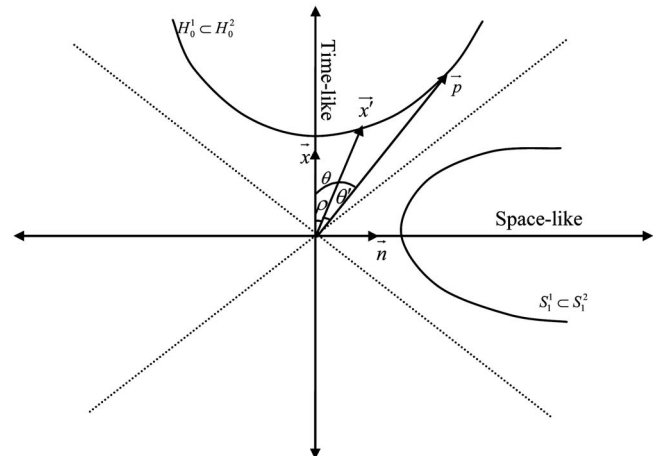


Fig. 6 Normal plane defined by vectors $\vec{x}$ and $\vec{n}$

$$d\vec{x}' = \vec{t}(\kappa \sinh \rho - \sigma \cosh \rho) + (\vec{n} \cosh \rho + \vec{x} \sinh \rho) d\rho$$

This vector shows the propagation direction of the geometric position of curvature centres, i.e., spherical evolution. That is,

$$g(\vec{t}, d\vec{x}') = \sigma \cosh \rho - \kappa \sinh \rho = 0$$

If we put the value of σ from Eq. (22), we reach

$$-\kappa \sinh \rho + \Psi \sinh \theta \cosh \rho = 0 \quad (27)$$

If we do scalar product for the both sides of Eq. (26) from the both sides, we write

$$g(\vec{t}, d\vec{p}) = \sigma \cosh \theta - \kappa \sinh \theta$$

and using Eq. (22), we reach

$$g(\vec{t}, d\vec{p}) = \Psi \sinh \theta \cosh \theta - \kappa \sinh \theta \quad (28)$$

Now we want to eliminate magnitude κ from the Eqs. (27) and (28). The value of κ evaluated from Eq. (27) is

$$\kappa = \Psi \frac{\sinh \theta \cosh \rho}{\sinh \rho}$$

Substituting this value into Eq. (28) and then rearranging gives us

$$\begin{aligned} g(\vec{t}, d\vec{p}) &= \Psi \sinh \theta \left(\cosh \theta - \frac{\sinh \theta \cosh \rho}{\sinh \rho} \right) \\ &= \Psi \sinh \theta \left(\frac{\sinh \rho \cosh \theta - \sinh \theta \cosh \rho}{\sinh \rho} \right) \\ &= \Psi \sinh \theta \frac{\sinh(\rho - \theta)}{\sinh \rho} \\ &= \Psi \frac{\sinh \theta \sinh(\theta - \rho)}{\sinh[-(\rho + \theta) - \theta]} = \frac{\Psi}{\coth \theta - \coth \theta'} \end{aligned}$$

where we adopt that $\theta' = \theta - \rho$. Now, if we denote $\vec{t}_p$ as the common tangential unit vector of pole curves (P) and (P') and ds_p as the common arc element then

$$d\vec{p} = \vec{t}_p \sigma_p$$

and

$$g(\vec{t}, \vec{t}_p)(\coth \theta - \coth \theta') = \frac{\Psi}{\sigma_p}$$

Furthermore, as

$$g(\vec{t}, \vec{t}_p) = -\sinh \alpha$$

we find the corresponding Euler-Savary formula as

$$(\coth \theta' - \coth \theta) \sinh \alpha = \frac{\Psi}{\sigma_p} \quad (29)$$

For all mutual point couples X, X' we give a meaning for Ψ/σ_p , which is the left-hand side of Eq. (29). For this aim, we consider the spherical curvature centers M and M' of curves (P) and (P'). That is, the centers M and M' have spherical distances ρ_p and ρ'_p that we call them as spherical curve radius of pole curves from

rotation pole P , respectively. Let us take M be a point of S_1^2 . In this case, the curvature center of M in the movement D_1 will be M' . This can be seen by taking the curvature circle in time t for the pole curves (P) and (P'). Therefore, curvature circle of (P), for which the spherical center is M , rotates on the curvature center of (P') for which the spherical centre is again M . In this rotation point, M draws a circle centered at M' . Thus, if we accept

$$\theta = \rho_p, \quad \theta' = \rho'_p, \quad \sinh \alpha = 1$$

in Eq. (29), then we obtain

$$\coth \rho'_p - \coth \rho_p = \frac{\Psi}{\sigma_p} \quad (30)$$

where ρ_p and ρ'_p are the spherical curvature centres for the pole curves (P) and (P'), respectively.

THEOREM 9. *In a one-parameter movement D_1 of Lorentz spherical S_1^2 with respect to the Lorentz spherical $\bar{S}_1^2$, a fixed point X in S_1^2 draws a trajectory for which its instantaneous curvature center is X' in the Lorentz sphere $\bar{S}_1^2$. In this case, we can say the following for the contradictory movement of Lorentz sphere $\bar{S}_1^2$ with respect to Lorentz sphere S_1^2 . The point X' , which is fixed in $\bar{S}_1^2$, draws a trajectory in S_1^2 for which the curvature centre is the initial point X . Mutual interrelation between the points X and X' is given by the Euler-Savary formula.*

References

- [1] Ergin, A. A., 1989, "Lorentz Düzleminde Kinematik Geometri," Doktora tezi, Ankara Üniversitesi Fen Bilimleri Enstitüsü, Ankara.
- [2] Ikawa, T., 2003, "Euler-Savary's Formula on Minkowski Geometry," *Balkan J. Geometry App.*, **8**(2), pp. 31–36.
- [3] Bottema, O., 1965, "Acceleration Axes in Spherical Kinematics," *ASME J. Eng. Ind.*, **85**, pp. 150–154.
- [4] Meyer Zur Capellen, W., and Dittrich, G., 1966, "The Instantaneous Distribution of a Spherically Moving System," *J. Mech.*, **1**, pp. 23–24.
- [5] Beyer, O., 1963, *Technische Raumkinematik*, Springer-Verlag, Berlin, pp. 207–221.
- [6] Bottema, O., and Roth, B., 1979, *Theoretical Kinematics*, North-Holland, Amsterdam, pp. 193–197.
- [7] Muller, H. R., 1963, "Kinematik Dersleri," (çeviri). Ankara Üniversitesi Fen-Fakültesi yayınları 27.
- [8] Chen, Y. J., and Ravani, B., 1987, "Offsets Surface Generation and Contouring in Computer Aided Design," *ASME J. Mech., Transm., Autom. Des.*, **109**, pp. 133–142.
- [9] Farouki, R. T., 1986, "The Approximation of Non-Degenerate Offset Surface," *Comput. Aided Geom. Des.*, **3**, pp. 15–43.
- [10] Papaioannou, S. G., and Kirtsis, D., 1985, "An Application of Bertrand Curves and Surfaces to CAD/CAM," *Comput.-Aided Des.*, **17**(8), pp. 348–352.
- [11] Pennock, G. R., and Sankaranarayanan, H., 2003, "Path Curvature of a Geared Seven-Bar Mechanism," *Mech. Mach. Theory*, **38**, pp. 1345–1361.
- [12] Pennock, G. R., and Raje, N. N., 2004, "Curvature Theory for the Double Flier Eight-Bar Linkage," *Mech. Mach. Theory*, **39**, pp. 665–679.
- [13] Birman, G. S., and Nomizu, K., 1984, "Trigonometry in Lorentzian Geometry," *Am. Math. Monthly*, **91**(9), pp. 543–549.
- [14] Birman, G. S., and Nomizu, K., 1984, "The Gauss-Bonnet Theorem for 2-Dimensional Space-Times," *Mich. Math. J.*, **31**, pp. 77–81.
- [15] O'Neill, B., 1983, *Semi Riemannian Geometry*, Academic Press, New York.
- [16] Akutagawa, K., and Nishikawa, S., 1990, "The Gauss Map and Space-Like Surfaces With Prescribed Mean Curvature in Minkowski 3-Space," *Tohoku Math. J.*, **42**, pp. 68–82.

J. P. McGarry¹

Department of Mechanical and Biomedical
Engineering and National Centre for Biomedical
Engineering Science,
National University of Ireland,
Galway, Ireland
and Department of Mechanical Engineering,
University of California,
Santa Barbara, CA 93106-5070
e-mail: pmcg@engineering.ucsb.edu

B. P. O'Donnell

P. E. McHugh

E. O'Cearbhaill

Department of Mechanical and Biomedical
Engineering and National Centre for Biomedical
Engineering Science,
National University of Ireland,
Galway, Ireland

R. M. McMeeking

Department of Mechanical Engineering,
University of California,
Santa Barbara, CA 93106-5070

Computational Examination of the Effect of Material Inhomogeneity on the Necking of Stent Struts Under Tensile Loading

This study presents a computational investigation of tensile behavior and, in particular, necking due to material inhomogeneity of cardiovascular stent struts under conditions of tensile loading. Polycrystalline strut microstructures are modelled using crystal plasticity theory. Two different idealized morphologies are considered for three-dimensional models, with cylindrical grains and with rhombic-dodecahedron grains. Results are compared to two-dimensional models with hexagonal grains. For all cases, it is found that necking initiates at a significantly higher strain than that at UTS (ultimate tensile stress). Two-dimensional models are shown to exhibit an unrealistically high dependence of necking strain on randomly generated grain orientations. Three-dimensional models with cylindrical grains yield a significantly higher necking strain than models with rhombic-dodecahedron grains. It is shown that necking is characterized by a dramatic increase in stress triaxiality at the center of the neck. Finally, the ratios of UTS to necking stress computed in this study are found to compare well to values predicted by existing bifurcation models. [DOI: 10.1115/1.2722776]

Introduction

Coronary stents are tubular metal (typically, 316L stainless steel) scaffold devices used in the treatment of heart disease. They are inserted into the coronary arteries in the heart as part of the angioplasty procedure, where arterial blockages are removed, and they remain as permanent implants for the purpose of keeping the artery open and maintaining adequate blood flow to the heart muscle. Stent structures are generated using repeating patterns of thin interconnected metal struts. For coronary applications, the struts are typically $\leq 100 \mu\text{m}$ in thickness. Stents are deployed into position in the artery by expansion using a balloon catheter. The balloon expansion results in significant plastic deformation in the stent, ensuring that it retains a permanently expanded state. The success of these micron-scale coronary stents has fueled interest in developing smaller stents for peripheral and neurovascular applications. However, for smaller stents to be designed correctly, the mechanical properties of the micron-sized struts must be understood thoroughly.

A body of work now exists in the literature on modeling the mechanical performance of stents (e.g. [1–5]). Such work includes the prediction of mechanical performance measures, such as required expansion pressures, bending flexibility, and elastic recoil on balloon deflation. Many different stent designs have been compared, and relative performance assessments have been made. However, in the bulk of the work published to date, the material has been represented using conventional isotropic elasticity- J_2 flow theory plasticity, using standard macroscopic values for material properties, and no consideration has been given to the fact that the size scale of the stent microstructure is significant com-

pared to the strut size scale ($\leq 100 \mu\text{m}$) as illustrated in Fig. 1. Only in the study of McGarry et al. [6] has the effect of microstructural features on stent deployment been investigated.

An experimental study by Murphy et al. [7] investigated the dependence of the tensile failure strain of 316L stainless-steel stent struts on strut thickness. A computational micromechanics study by Savage et al. [8] using 2D (two-dimensional) crystal plasticity models further investigated this size effect, and it was shown that localization of strains occurs more easily in thinner struts where there are few grains across the thickness, leading to early shear banding and macroscopic softening, whereas in a thick strut constraint from surrounding grains delays shear banding and macroscopic softening. However, these studies suffer from two drawbacks. First, material failure was assumed to coincide with UTS (ultimate tensile strength). It has been established that necking only occurs at UTS for very long thin strips. For shorter stubbier samples a considerable delay occurs between UTS and necking [8]. Second, the use of 2D crystal plasticity models represents a gross underestimation of the number of grains and associated plastic slip orientations in a sample.

In the current paper, these issues are addressed, focusing on the development on 3D (three-dimensional) crystal plasticity models of stent struts. Because of very large numerical problem sizes that such models generate, attention is focused on a single strut thickness, viz $75 \mu\text{m}$. Two different grain morphology and packing configurations are considered. The study has significant value in that this is the first time that 3D micromechanical modeling of this type has been applied to stent struts. The fact that stent struts are so thin is highly fortuitous from the micromechanical modeling point of view because one can reasonably aspire to develop a representative model of a complete segment of the stent device, in this case a strut. This is significant when one contrasts it with a typical situation where micromechanical modeling has been used, for example, in the study of the mechanical performance of metal-

¹Corresponding author.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received November 9, 2005; final manuscript received January 17, 2007. Review conducted by Yonggang Huang.

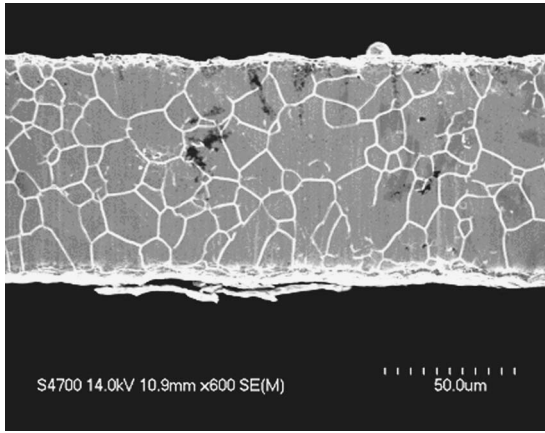


Fig. 1 SEM image of a 316L laser-cut strut, representative of that used in stents

matrix composites (e.g. [9,10]), where one can only reasonably aspire to develop a model of a small volume element of the material, and where major assumptions must be made to relate the microscale model predictions to the overall macroscale performance of the material or a component made from the material.

In the present work, the predictions of the 3D models in terms of microscale deformation and macroscale stress-strain behavior in tensile loading are examined and compared to 2D model predictions. 316L single-crystal material properties are taken from Okamoto et al. [11]. Although the study of Hoc et al. [12], using crystal plasticity and a single layer of grains, considers localisation at UTS, the current study represents the first attempt to examine the phenomenon of necking due to crystallographic orientation inhomogeneity in 3D microstructures. Such 3D finite element models of polycrystalline materials, using various versions of crystal plasticity theory, have recently appeared in the literature, (e.g., [13–15]), but such computationally demanding modeling techniques have not previously extended to study of necking. Of particular interest is the strain at which necking occurs. It has long been established due to the formulations of Cowper and Onat [16], Hutchinson and Miles [17], Hill and Hutchinson [18] and Miles [19] that bifurcation from a uniform stress state occurs post UTS (ultimate tensile strength). Strains and stresses at necking computed in this study are compared to the bifurcation stresses and strains predicted by the formulations of Miles [19] for specimens with rectangular cross sections.

It has been widely reported that triaxiality in a specimen under tensile loading is constant only until the onset of necking [20–22] with the maximum triaxiality occurring at the middle of the necked cross section [23–25]. In this work, the stress state in the vicinity of the necked region of 3D simulations is investigated.

Theory and Numerics

The simulations performed in this work are based on large strain (finite deformation) kinematics and incorporate an elastic-plastic constitutive description of the material. The ABAQUS® finite element code [26] is used for the simulations and the Patran® code is used for finite element mesh generation. The user material subroutine (UMAT) of Huang [27] is used to implement the constitutive behavior. Elasticity is considered as being linear in terms of finite deformation quantities. Plasticity is described using crystal plasticity theory, which attempts to represent the flow of dislocations along slip systems in metallic crystals in terms of continuum plastic shear strains. In particular, the rate-dependent single-crystal formulation of the theory of Pierce et al. [28] is used. Details of numerical implementation are given in

[27], and described in detail in [29]. A brief overview of aspects of the theory, which are most relevant for the application of interest here, is given below.

In crystal plasticity theory, plastic slip is assumed to obey Schmidt's law, where the rate of plastic shear strain in the direction of a particular slip system α , $\dot{\gamma}_\alpha$, is assumed to depend on the Schmidt resolved shear stress for that slip system (on the slip plane, in the direction of the slip system), τ_α . In this work, the following power-law rate-dependent relationship is used:

$$\dot{\gamma}_\alpha = \dot{a} \operatorname{sgn}(\tau_\alpha) \left\{ \left| \frac{\tau_\alpha}{g_\alpha} \right| \right\}^n \quad (1)$$

where $\dot{a}$ and n are the reference strain rate and rate sensitivity exponent, respectively. Material strain hardening is specified by the slip system strain hardness g_α . This, in turn, is given by the path-dependent integration of the following evolution equation:

$$\dot{g}_\alpha = \sum_{\beta=1}^N h_{\alpha\beta} |\dot{\gamma}_\beta| \quad (2)$$

$$g_\alpha = g_0, \quad t = 0$$

where the summation is performed over all slip systems active at a material point ($\beta: 1 \rightarrow N$). The matrix of hardening moduli $h_{\alpha\beta}$ is given by

$$h_{\alpha\beta} = h q_{\alpha\beta} \quad (3)$$

where $q_{\alpha\beta}$ is a matrix representing the relationship between self and latent hardening in the crystal and h is given by the derivative of the hardening function g with respect to the accumulated slip γ_a , defined as

$$\gamma_a = \int_0^t \sum_{\alpha=1}^N |\dot{\gamma}_\alpha| dt \quad (4)$$

The following form of the strain hardness function g [28] is used, which is known to accurately capture the shear stress-shear strain behavior of metallic crystals:

$$g(\gamma_a) = g_0 + (g_\infty - g_0) h_0 \tanh \left| \frac{h_0 \gamma_a}{g_\infty - g_0} \right| \quad (5)$$

This function involves three material strain-hardening constants, g_0 , g_∞ , and h_0 , that must be determined by fitting to experimental tensile stress-strain data. The accumulated slip is a measure of the total crystallographic plastic strain at a material point in the crystal. Taylor isotropic hardening was assumed, meaning that self and latent hardening rates were assumed to be the same, and in this context $q_{\alpha\beta}$ was a fully populated matrix of "1"s.

A rate-dependent formulation of the theory is used here due to the fact that this removes nonuniqueness of solution problems often associated with rate-independent crystal plasticity formulations [28] and also produces good numerical stability. The user-defined material subroutine (UMAT) in which the theory was originally implemented [27] is modified for the current study to automatically control the global time step size, to ensure both accuracy and efficiency of solution.

Material Properties

316L stainless steel has an face-centered-cubic (fcc) crystal structure implying 12 slip systems in three dimensions. Elasticity is assumed to be isotropic and standard published values are used: Young's modulus, $E=209$ GPa and Poisson's ratio, $\nu=0.28$. It was shown in Savage that changing from isotropic elasticity to anisotropic cubic elasticity in polycrystalline stent strut models had a negligible effect on the overall predicted material behavior. With regard to plasticity, experimentally measured tensile stress-strain curves for 316L single crystals were found in the literature [11] and these data were used to calibrate the strain hardness function

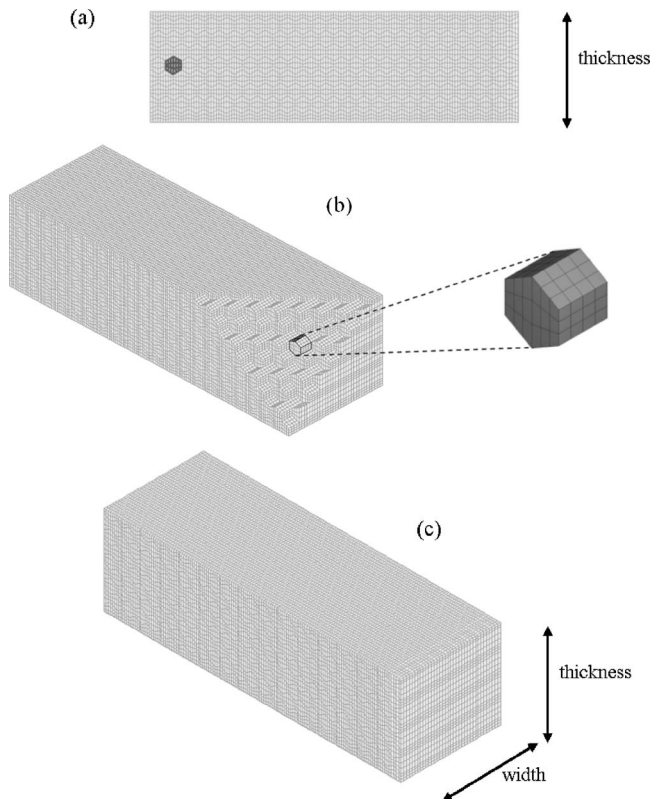


Fig. 2 (a) 2D hexagonal grain finite element model with a typical hexagonal grain highlighted. 3D cylindrical grains finite element model: (b) break-out view with a typical grain highlighted and (c) full finite element mesh. The dimensions “width” and “thickness” discussed in the text are shown.

for the crystal (Eq. (5)). The resulting values of the strain-hardening constants are $g_0=50$ MPa, $g_\infty=330$ MPa and $h_0=225$ MPa. The rate dependence of stainless steel at room temperature has been widely reported in the literature [30–33]. A rate exponent (n) of 20 is assumed for the current study because such a value not only reflects such experimentally observed rate dependence but also ensures robust convergence of the solution. A reference strain rate ($\dot{\epsilon}$) of 0.001 s^{-1} is also assumed and models are subjected to an applied nominal strain rate of this value for all simulations.

Model Development

Two 3D models were generated. In both cases, a straight segment of stent strut is modeled using assemblages of grains, and in both cases the microstructure is assumed to be ideal, meaning that regular idealized grain morphologies are used.

3D Cylindrical Grains Model. The generation of this model can be best understood by first considering a 2D idealization of the microstructure. In the plane, the microstructure is represented as an assemblage of hexagons, as illustrated in Fig. 2(a), all regular except for the partial hexagons at the boundaries. To generate a 3D microstructure, one can imagine “extruding” the hexagons in the out-of-plane direction, a distance equal to the width of the strut, to generate hexagonal cylinders. Individual grains can then be generated by cutting the cylinders at regular intervals along their lengths, as illustrated in Fig. 2(b). Image analysis in [8] of micrographs of the stent struts tested in [7] revealed an average grain area of $92 \mu\text{m}^2$. Based on this the size of the in-plane hexagons was set at $92 \mu\text{m}^2$ in the model. Assuming a strut thickness of $75 \mu\text{m}$, this translates into approximately nine grains across the thickness (seven full grains and two partial grains). The stent

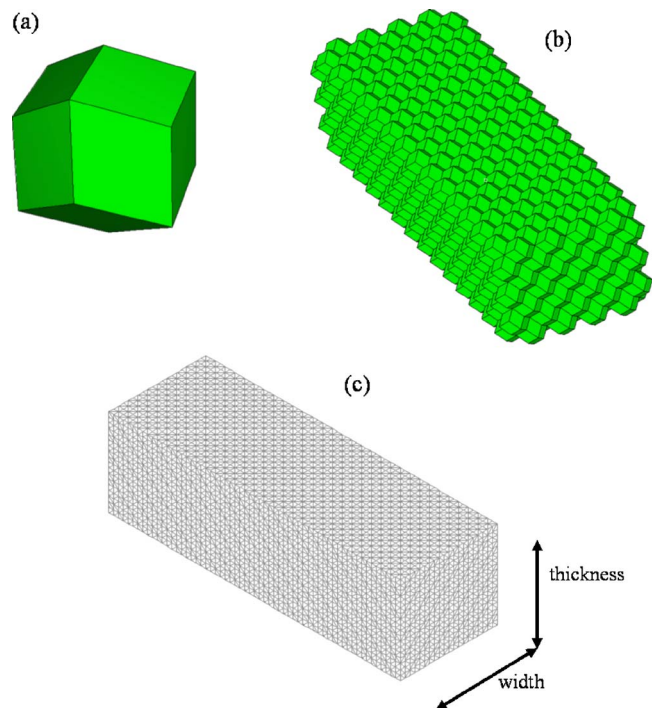


Fig. 3 3D close-packed rhombic-dodecahedron model: (a) one rhombic-dodecahedron grain, (b) arrangement of full rhombic-dodecahedron grains in the model (partial grains at surfaces and edges not shown), and (c) full finite element mesh. The dimensions “width” and “thickness” discussed in the text are shown.

struts tested in [7] had an out-of-plane width of $\sim 85 \mu\text{m}$ and a width of this order is used in the model (the exact dimension is $83.3 \mu\text{m}$). Reflecting the number of grains through the thickness, the hexagonal cylinders are divided into nine equally sized grains in their axial direction. A length-to-thickness aspect ratio of 3.32:1 is used [8], and this results in a model length of $249 \mu\text{m}$. The final model has 1980 grains. The model was meshed using eight-noded brick elements with reduced integration and hourglass control. Attempting to ensure a reasonable number of elements per grain for accuracy results in a model with 134,784 elements. Tensile stretching is simulated by constraining the left-hand end of the model in the axial direction and by applying a uniform axial displacement to the right-hand end. Rigid-body motion is prevented by applying appropriate constraints at the left-hand end. The full mesh is shown in Fig. 2(c). Random lattice orientations in 3D were assigned to the grains and runs were performed for three random orientation sets.

3D Close Packed Model. A second 3D model was generated based on a desire to have an internal structure that would have greater physical realism than the cylindrical grains model discussed above. The morphology selected is based on an assemblage of rhombic dodecahedrons; this assemblage represents a close packing of space, in the same sense as close-packed arrangement of spheres. Indeed, if one considered a close-packed arrangement of spheres and if one “straightened out” the surfaces of the spheres, so as to fill the gaps between the spheres, the spheres would become rhombic dodecahedrons. To represent the strut, the model is structured to have the same overall dimensions as for the cylindrical grains model ($75 \mu\text{m} \times 83.3 \mu\text{m} \times 249 \mu\text{m}$). A single rhombic-dodecahedron grain is shown in Fig. 3(a). To generate the strut model, full rhombic-dodecahedron grains are arranged as shown in Fig. 3(b). Partial grains are then added at the sides and edges to generate smooth surfaces and straight edges in the model. The final model consists of a total of 1875 grains (full plus par-



Fig. 4 Plots of deformed meshes for two different sets of crystal orientations at a strain of 0.95

tial). To facilitate meshing the 12-sided rhombic-dodecahedron grains, tetrahedral elements are used, in particular, 10-noded tetrahedrons with reduced integration. Attempting to ensure a reasonable number of elements per grain for accuracy results in a model with 129,600 elements. The model is constrained and loaded in exactly the same way as for the cylindrical grains model. The full mesh is shown in Fig. 3(c). Random lattice orientations in 3D were assigned to the grains and runs were performed for three random orientation sets.

2D Model. The validity of the simplification of a generalized plane strain assumption used in [6,8] is examined by comparison of the aforementioned 3D models to a 2D generalized plane strain model with a similar number of grains through the stent thickness and an out of plane depth of $0.85 \mu\text{m}$. A typical mesh comprising of 220 grains and 3744 elements is shown in Fig. 2(a). Both

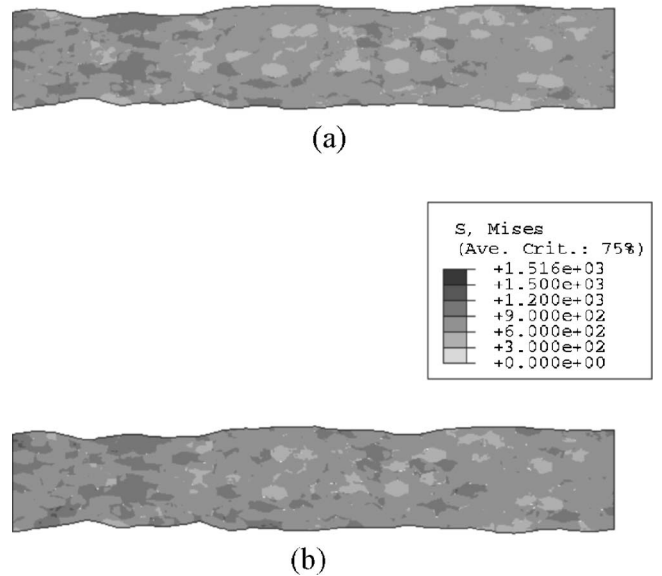


Fig. 6 Contour plots of von Mises stress at a strain of 0.5 for (a) a fine mesh with second-order elements and (b) a coarse mesh with first-order elements

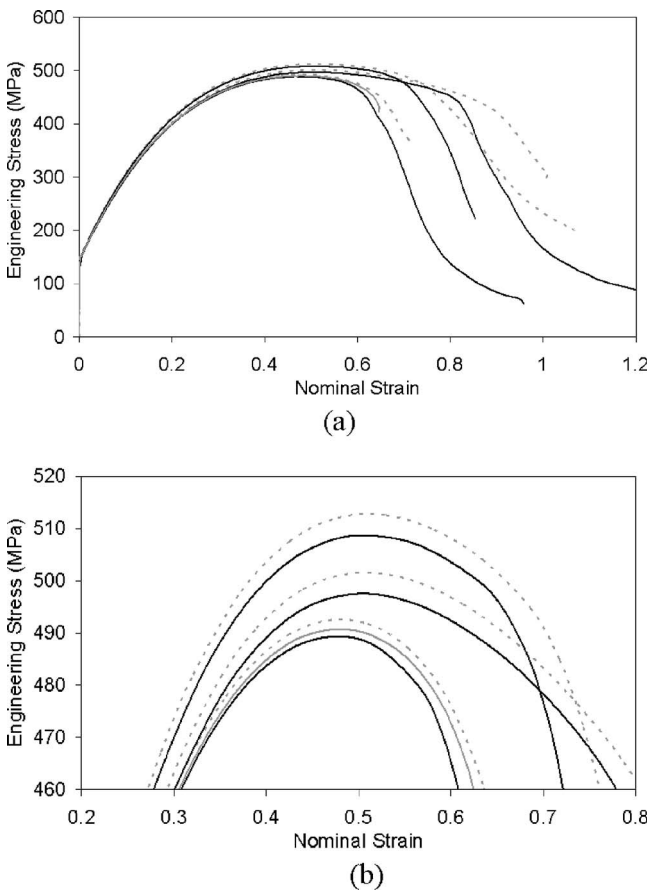


Fig. 5 Stress-strain curves for 2D models using a coarse mesh with first-order elements (broken gray curves), a fine mesh with second-order elements (solid black curves), and a fine mesh with first-order elements (solid gray curve): (a) full stress-strain curves and (b) stress-strain relationships at UTS

four-noded (first-order) and eight-noded (second-order) quadrilateral elements are used and a dense mesh consisting of 10,605 elements is also considered. In the ABAQUS® implementation, generalized plane strain is achieved by integration across the plane to yield net-zero out-of-plane forces.

It is worth noting that one could view the models as representing an evolution toward physical realism. For the 2D model, the grains are implicitly cylinders of arbitrary length in the out-of-plane direction. The cylindrical grains model is an evolution of this in that the cylinders are given a definite length (the width of the strut) and are subdivided to generate grains, with the number of grains determined from experimental data. Although the result is that this model is physically reasonable in terms of numbers of grains, etc., it has an unphysical bias in that its structure in the thickness direction is different from that in the width direction. The final model, the close-packed model, eliminates this bias. For all three model types, axial stretching is applied at a very low nominal strain rate (equal to $\dot{a}$) to eliminate any spurious rate-sensitivity effects that might arise from the use of a rate-dependent material model.

In this work, the microscopic mechanical performance of the models is inferred from comparison of deformed meshes and contour plots of von Mises stress. Although the von Mises stress has no theoretical significance in crystal plasticity formulations, it is a stress invariant that reflects the level of computed plastic deformation. The macroscopic performance of the models is quantified in terms of macroscopic engineering stress-nominal strain (s - e) curves. The effect of grain configuration and grain orientations on yield, hardening behavior, and UTS is examined, with the effect of element type and mesh density being considered, in addition, for 2D models. Material failure is then considered, with the onset of necking being determined by examination of the evolution of cross-sectional areas throughout the strut length. The stress state

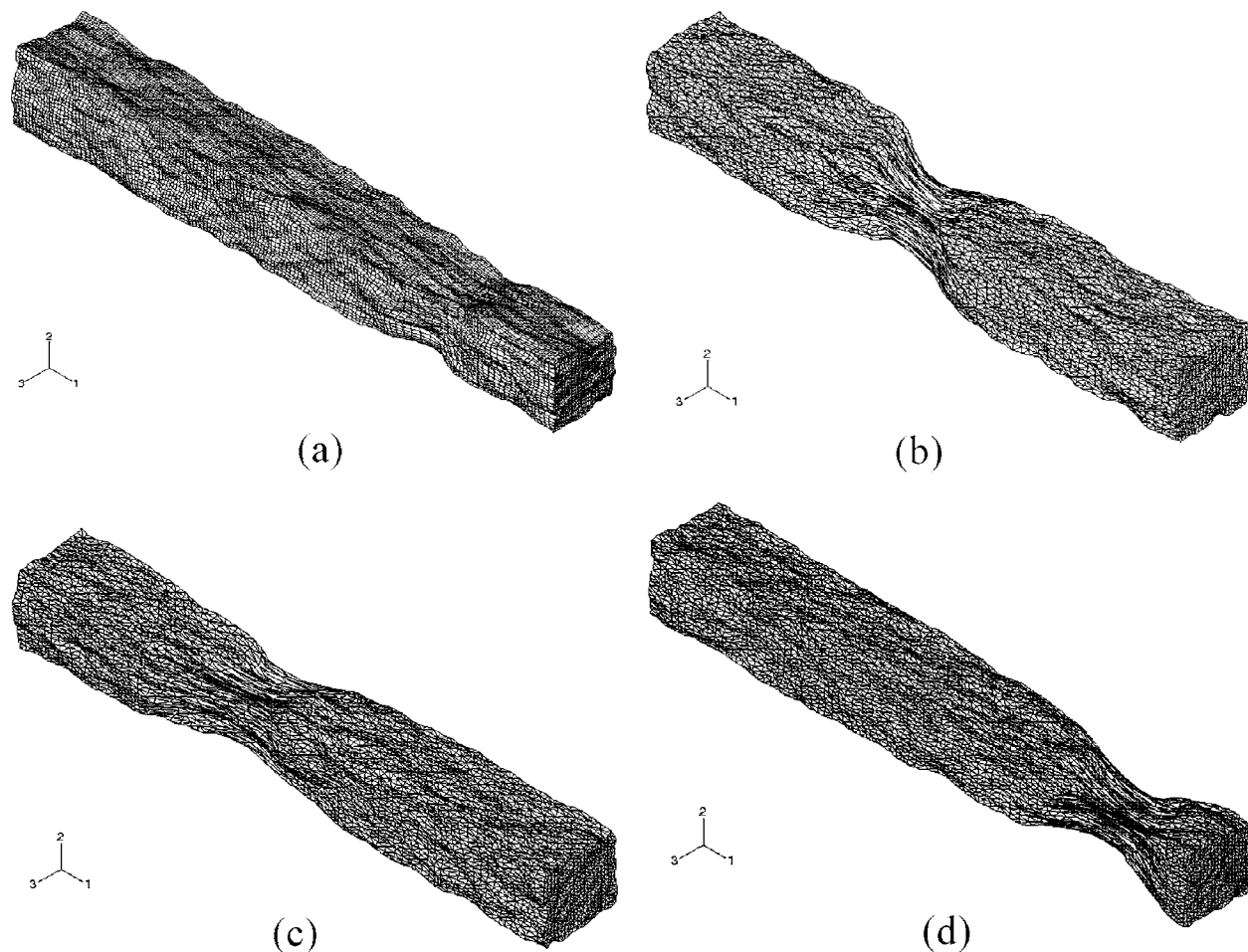


Fig. 7 Deformed mesh at a strain of 0.95 for (a) the cylindrical grain configuration, (b)–(d) the rhombic-dodecahedron crystal configuration with three different sets of random crystal orientations

in the necking region is also examined. Finally, the computed failure strains are compared to values predicted by the bifurcation modeling of Miles [19].

Results

Figures 4(a) and 4(b) show plots of deformed 2D meshes at a macroscopic nominal strain of 95% for two different sets of random crystal orientations. Clearly, the location of the development of shear bands and necking is highly dependent on the randomly assigned crystal orientations. From Fig. 5(a) is apparent that the random orientations also affect the UTS of the strut. Indeed, at a nominal strain of 0.1, a difference in engineering stress of 10 MPa (over 3%) is computed for two different sets of crystal orientations. At a nominal strain of 0.5, this difference increases to 20 MPa (4% difference). In terms of the strain at UTS, a maximum difference of 0.03 is computed.

The effect of element type and mesh density on the computed UTS is shown in Fig. 5(b). A difference in strain at UTS of 0.013 is computed between a dense mesh comprising of second-order elements and a coarse mesh comprising first-order elements for all orientations. A corresponding UTS difference of 4.3 MPa is computed for all orientation sets. Such differences can be explained by the higher levels of stress localizations computed using a dense mesh with second-order elements. Contour plots of von Mises stress at a nominal strain of 0.5 are shown for the dense mesh with second-order elements in Fig. 6(a) and for the coarse mesh with first-order elements in Fig. 6(b) for the same set of crystal orientations. In the former, a peak value of 1606 MPa is computed

compared to 1516 MPa in the latter. The use of a dense mesh with first-order elements is found to produce s - e curves closer to those computed using second-order elements (Fig. 5(b)). A peak von Mises stress of 1576 MPa is computed for the first-order dense mesh, which is quite close to the peak value shown in Fig. 6(a).

Figures 7(a) and 7(b) show plots of deformed meshes for the cylindrical crystal model and the rhombic-dodecahedron crystal model, respectively, at a macroscopic strain of 0.95. Necking is significantly more progressed in the latter case. Figures 7(c) and 7(d) show the deformed mesh for the rhombic-dodecahedron crystal model computed using two additional sets of crystal orientations demonstrating the effect of crystal orientations on the location of shear band formation and necking. A similar dependence is observed for cylindrical grains, again with necking being less pronounced for all crystal orientation sets than is the case for rhombic-dodecahedron grains at the same strain.

Figure 8 shows that struts comprised of cylindrical grains yield s - e curves that are relatively insensitive to crystal orientations until the development of necking. The same could be said of struts comprised of rhombic-dodecahedron grains up to a strain of 0.2. Indeed, the hardening behavior of both configurations is identical up to this point, after which the hardening rate of the rhombic-dodecahedron crystal configuration reduces and UTS values ranging from 481 MPa to 488 MPa at strains ranging from 0.56 to 0.57 are computed. The UTS for the cylindrical crystal configuration is 508 MPa in all cases at a strain of 0.64. Figure 9 reveals a lower level of stress localization in the case of the cylindrical

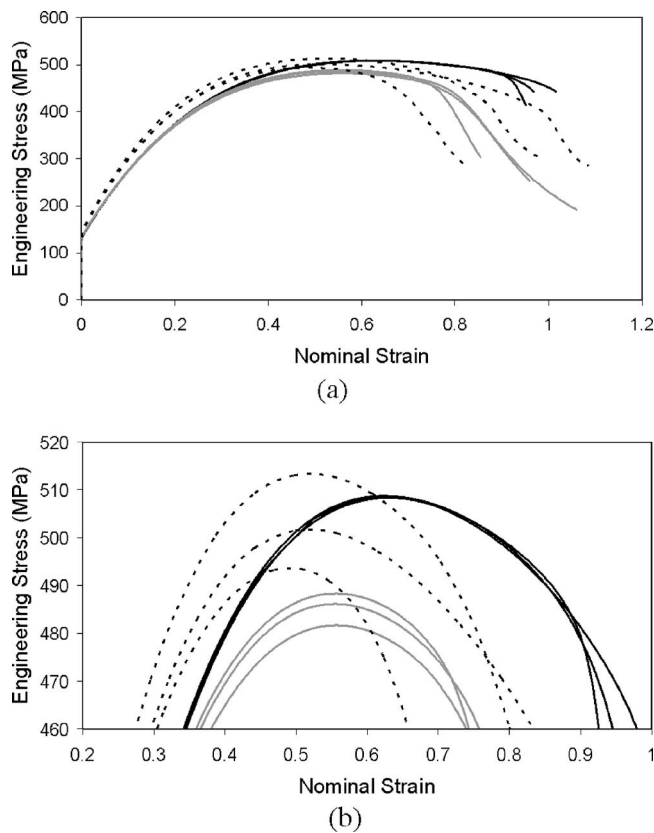


Fig. 8 Stress-strain curves for 3D models with the rhombic-dodecahedron crystal configuration (solid gray lines), cylindrical crystal configuration (solid black lines) and for 2D models with first-order elements (broken gray lines): (a) full stress-strain plot and (b) stress-strain relationships at UTS

crystal configuration than is the case for rhombic-dodecahedron crystal configuration at a strain of 0.5 (prior to UTS).

Turning attention again to the 2D simulations, it is apparent from Fig. 8 that the initial hardening rate is much higher than that computed using the 3D meshes. The random crystal orientations also have a much greater effect, with UTS values ranging from 494 MPa to 513 MPa. Strains at UTS range from 0.47 to 0.53, which is lower than all corresponding values computed using 3D analyses.

The onset of necking can be determined by considering the evolution of cross-sectional areas at ten locations evenly distributed throughout the strut length, as shown in Fig. 10(a). Cross-sectional area evolution for the 3D analysis using rhombic-dodecahedron grains shown in Fig. 7(d) is presented in Fig. 10(b). Details of significant stresses and strains for this simulation are given on the seventh row of Table 1 (3D rhom ori1). In Fig. 10(a), a point of inflection is observed for section 8 at a nominal strain of 0.74, leading to a dramatic reduction in cross-sectional area, as depicted in Fig. 7(d). All other sections exhibit a reduction in the rate of area decrease following this point with elastic unloading eventually occurring. This localization of the deformation to the vicinity of section 8 is a clear indication of necking. A similar pattern is revealed for the 2D analyses corresponding to the first row of Table 1 (2D hex ori1) as shown in Fig. 10(c). The onset of necking occurs at a nominal strain of 0.6 between sections 7 and 8. The localization of the neck can be observed with the elastic unloading of sections 7 and 8, initiating much later than all other sections outside the necking region. Figure 10(d) shows a comparison of neck cross-sectional areas for 3D analyses using rhombic-dodecahedron and cylindrical grains. The point of inflection, marking the onset of a dramatic reduction in cross-sectional

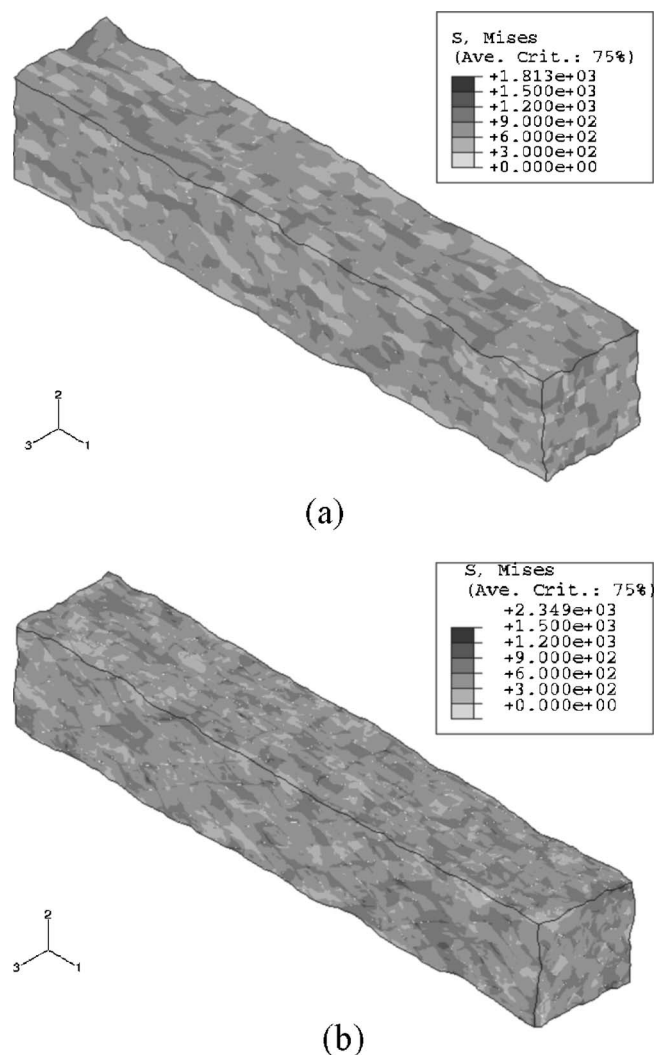


Fig. 9 Contour plots of von Mises at a strain of 0.5 for (a) cylindrical crystal configuration and (b) rhombic-dodecahedron crystal configuration

area, occurs at a much higher strain for the latter case. As can be seen in Table 1, the ratio of necking strain to strain at UTS is much higher for all analyses using cylindrical grains. This indicates that the necking strain is much more dependent on grain structure than is the strain at UTS and that strain at UTS is an unreliable indicator of necking.

True-stress–true-strain curves in Fig. 11(a) are constructed using the cross-sectional areas plotted in Fig. 10(c) for 3D analyses with cylindrical and rhombic-dodecahedron grains. As would be expected, the points of inflection occur at true strains corresponding to the nominal strains identified at the onset of necking in Fig. 10(c). As can be seen from Table 1, the ratios of true stress at necking to true stress at UTS are less sensitive to grain type than are the corresponding true strain ratios.

For the case of the rhombic-dodecahedron grain analysis a true stress–nominal strain curve is constructed. This plot is shown in Fig. 11(b) along with the corresponding engineering stresses and an alternative plot of true stress based on the well-established relationships whereby $\sigma = s(1 + e)$, where e is the nominal strain, s is the engineering stress, and σ is the true stress. This formula assumes a constant volume deformation and a homogeneous cross section. In the current analyses, compressible elastic deformation is negligible and, as can be seen from Fig. 9, the strut cross sections are relatively homogeneous at UTS. However, both true

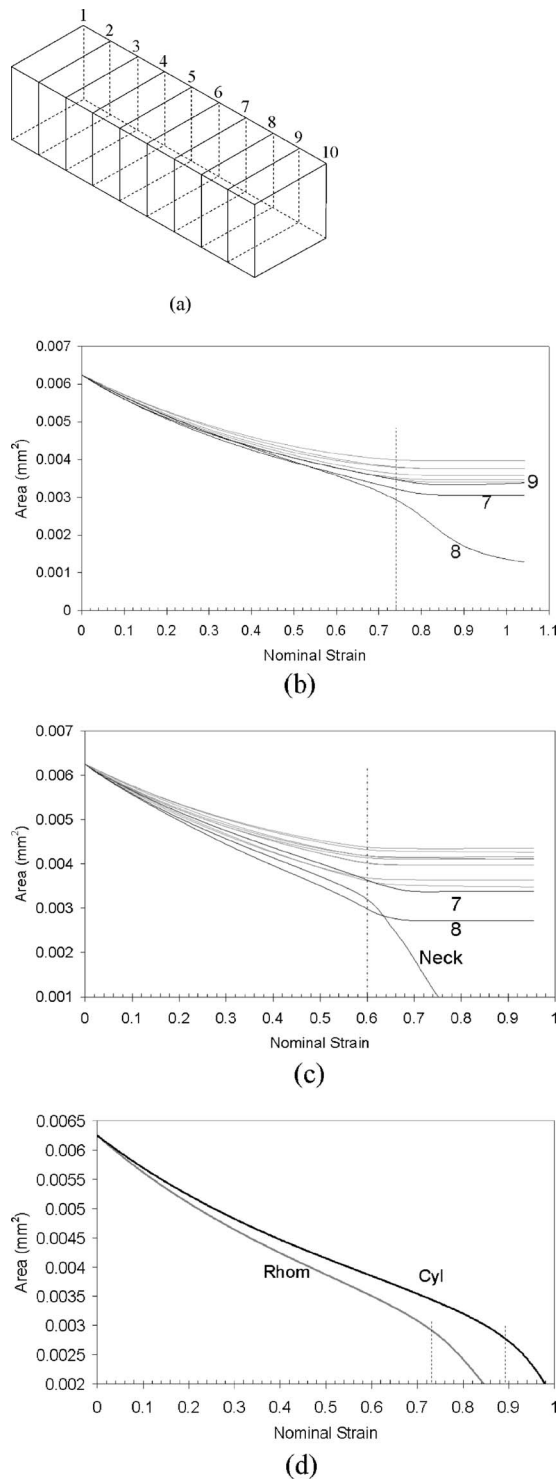


Fig. 10 (a) Cross-sectional areas at ten evenly spaced sections throughout strut length, (b) 3D analysis with rhombic-dodecahedron grains (curves for sections 7–9 labeled), and (c) 2D analysis (curves for sections 7, 8 and the location of the neck (between 7 and 8) labeled). (d) Comparison of cross-sectional area at neck for 3D analyses using rhombic-dodecahedron and cylindrical grains.

stress curves deviate significantly at strains as low as 0.2. At higher strains, the true stress curve derived from the aforementioned formula exhibits softening. Although such behavior is unphysical, it is interesting to note from Fig. 11(b) that the onset of softening (point B) occurs at the same nominal strain as the point

of inflection of the true stress curve calculated from the actual neck cross-sectional area (point A). This point can be determined from the engineering stress-nominal strain curve from the criterion given in Eq. (6) corresponding to point C in Fig. 11(b)

$$\frac{ds}{de} = \frac{-s}{1+e} \quad (6)$$

As can be seen from Table 1, the true strain at which this criterion is satisfied corresponds very closely to the necking strain for all 3D analyses: compare column 5 (ϵ) with column 7 (ϵ_{form}) in Table 1. In order to examine further the robustness of this criterion, 2D analyses, where crystal orientations have a more significant effect on mechanical behavior, are considered. A study of cross-sectional areas reveals a large variation of nominal strains at necking, as shown in Fig. 12(a), ranging from 0.6 to 0.81. It can be seen from Table 1 that the corresponding true strains at necking are very similar to the true strains that satisfy Eq. (6) for all three orientations, despite the strong dependence of the stress-strain relationships on grain orientations (compare column 5 to column 7 in Table 1). It should also be noted that the use of eight-noded second-order elements results in an increase in failure strain of <3% for all orientation sets.

The portions of the engineering stress-engineering strain curves between UTS and necking are highlighted in Fig. 12(b). This further underlines the inadequacy of UTS as a predictor of material ductility and failure.

In order to further examine the localization of deformation during necking, von Mises stress, hydrostatic stress, and the logarithmic strain component in the direction of stretching are computed for three distinct groups of elements in the interior of the strut, for the rhombic-dodecahedron grain analysis, as shown in Fig. 13, located remote from the site of necking (Fig. 13(a)), immediately adjacent to the site of necking (Fig. 13(b)), and at the site of necking (Fig. 13(c)). Plots of averaged logarithmic strain (Fig. 14(a)) reveal a distinct change in material behavior at a nominal strain of 0.74. Following this point, strains remote from the necking region undergo no further increase. In contrast, strains at the neck increase at a higher rate following this point. In the group of elements adjacent to the neck, the increase in strain becomes lower following a nominal strain of 0.85 as the region of necking becomes further localized. As regards the von Mises stress (Fig. 14(b)), a reduction of von Mises stress following a nominal strain of 0.74 in elements remote from the necking region clearly indicates the occurrence of elastic unloading. Further localization of the necking region is evident at a nominal strain of 0.85 with the von Mises stress then reducing in the elements immediately adjacent to the neck. At a nominal strain of 0.9, the von Mises stress in the neck begins to decrease slightly. In contrast, the hydrostatic stress (Fig. 14(c)) continues to increase dramatically in the same group of elements for nominal strains of >0.9. The hydrostatic stress reduction adjacent to the neck also reveals further localization of the necking region at a nominal strain of 0.85. Plots of averaged values of triaxiality, defined as (hydrostatic stress/von Mises stress), are shown for the three aforementioned regions of the struts in Fig. 14(d). An additional curve is computed for the elements along the outer edges of the necked region. Following the onset of necking, the triaxiality of the elements at the center of the necked region increases at a greater rate than that of the elements at the edges of the necked region. Interestingly, the triaxiality of the region immediately adjacent to the necked region decreases at a much greater rate than that of the region remote from necking.

It is worth noting that the 3D crystal plasticity simulations proved to be extremely demanding on computational resources; each taking on the order of 25 days on two processors of an SGI 3800 high-performance computer. The 2D crystal plasticity simulations took on the order of 24 h on the same machine using a single processor.

Table 1 e =engineering strain, s =engineering stress, ϵ =true strain, σ =true stress, ϵ_{UTS} =true strain at UTS, σ_{UTS} =true stress at UTS, ϵ_n =true strain at initiation of necking, σ_n =true stress at initiation of necking, σ_{form} =true stress at which the criterion of Eq. (6) is satisfied

		U.T.S.				necking			
		e	s	ϵ	σ	ϵ	σ	ϵ_{form}	$\epsilon_n/\epsilon_{UTS}$ σ_n/σ_{UTS}
2D hex	ori1	0.470	493.4	0.385	714.0	0.467	910.0	0.465	1.212 1.275
	ori2	0.529	512.5	0.425	715.1	0.528	925.0	0.530	1.243 1.294
	ori3	0.499	501.5	0.405	724.8	0.582	940.0	0.590	1.438 1.297
3D cyl	ori1	0.634	508.5	0.491	899.0	0.636	1066.0	0.635	1.295 1.186
	ori2	0.638	508.3	0.493	898.8	0.635	1065.0	0.635	1.287 1.185
	ori3	0.630	508.7	0.489	898.8	0.640	1072.0	0.660	1.310 1.193
3D rhom	ori1	0.564	481.4	0.447	808.2	0.553	988.0	0.546	1.236 1.222
	ori2	0.563	488.3	0.447	806.9	0.540	976.0	0.538	1.209 1.210
	ori3	0.562	486.0	0.446	808.7	0.544	991.0	0.552	1.220 1.225

In order to examine the generality of Eq. (6) for the onset of necking, homogeneous struts with geometric imperfections are considered. Plots of normalized decreases in area versus nominal strain are shown in Fig. 15(a) for a 3D strut with a 1% linear decrease in width and thickness from the strut end to the strut center point. From this plot, it is difficult to pinpoint the exact strain at which necking initiates. Clearly, some localization occurs prior to UTS, but it should also be pointed out that a significant increase in the rate of area reduction at the strut center point occurs prior to the nominal strain determined from Eq. (6). A 2D

generalized plane strain simulation with a similar linear reduction in thickness (Fig. 15(b)) suggests that such an increase in the rate of area reduction occurs at a higher nominal strain than that determined from Eq. (6). A similar trend is observed for a 2D strut with a notch at the center point (Fig. 15(c)). Further simulations involving variation in the percentage decrease in strut width and thickness and a nonlinear decrease in width and thickness over the length of the strut for both 2D and 3D geometries suggest that Eq. (6) cannot be readily generalized to necking, resulting from geometric defects in homogeneous struts. Indeed, the vastly different

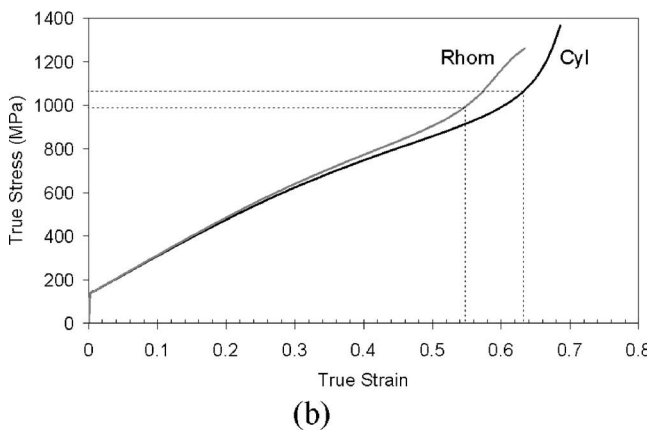
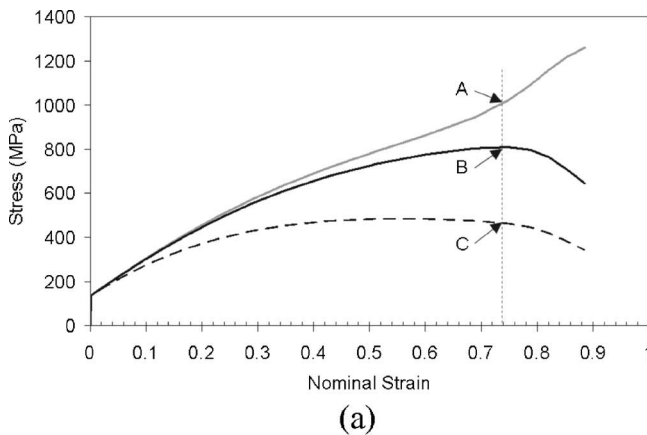


Fig. 11 (a) Comparison of engineering stress (broken line), true stress computed from $\sigma = s(1 + e)$ (black line) and true stress computed from actual neck cross-sectional area (gray line) for 3D analysis using rhombic-dodecahedron grains. (b) Comparison of true stress-true strain curves for 3D analyses using rhombic-dodecahedron and cylindrical grains.

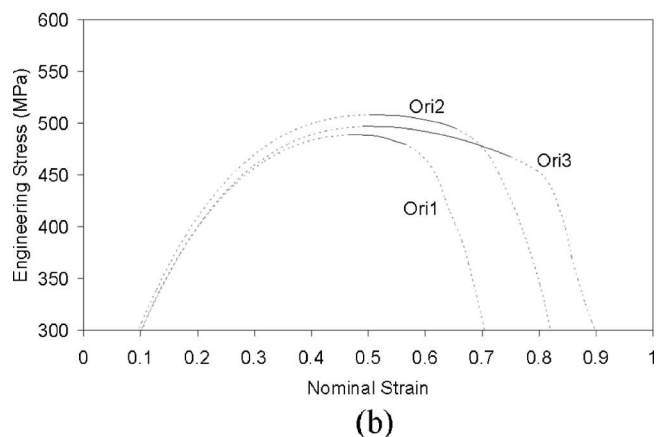
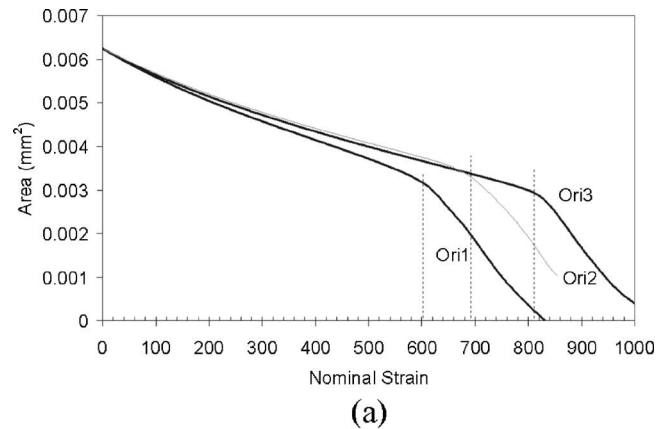


Fig. 12 (a) Cross-sectional area at the neck for the 2D analyses using three different sets of random crystal orientations and (b) sections of engineering stress-nominal strain curves, where hardening occurs post UTS shown in black, for the 2D analyses

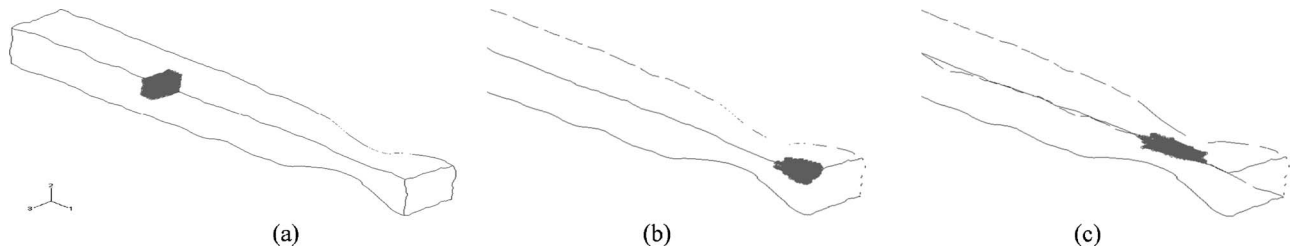


Fig. 13 3D rhombic-dodecahedron grain model shown in Fig. 7(d). Location of element groups (a) remote from neck, (b) adjacent to neck, and (c) at neck.

ratios of strain at necking to strain at UTS computed for 3D and 2D simulations merits further investigation that is beyond the scope of the current study.

In the aforementioned work of Miles [19], the ratio of bifurcation true stress to maximum true stress for a bar of rectangular cross section under uniaxial tension is given as

$$\frac{\sigma_b}{\sigma_m} = 1 + \left(1 - \frac{dE_t}{d\sigma}\right)^{-1} \left(\frac{1}{12}\alpha^2 + \frac{1}{12}\frac{\mu}{\sigma_m} \frac{c^2}{b^2}\alpha^2\right) \quad (7)$$

where E_t is the tangent modulus, c is the half width of the bar, b is the half depth of the bar, μ is the shear modulus and $\alpha = k\pi b/L$ where L is the bar length and k is an integer which corresponds to the position at which a neck associated with the bifurcation occurs. For example a value of 1 gives the earliest occurrence of a bifurcation corresponding to a neck at one end of the specimen. Similarly for true strains

$$\frac{\varepsilon_b}{\varepsilon_m} = 1 + \frac{1}{\varepsilon_m} \left(1 - \frac{dE_t}{d\sigma}\right)^{-1} \left(\frac{1}{12}\alpha^2 + \frac{1}{12}\frac{\mu}{\sigma_m} \frac{c^2}{b^2}\alpha^2\right) \quad (8)$$

whereby

$$\frac{\sigma_b - \sigma_m}{\varepsilon_b - \varepsilon_m} = \sigma_m \quad (9)$$

at the point of bifurcation. σ_b and ε_b refer to true stress and true strain at the bifurcation, and σ_m and ε_m refer to true stress and true strain at UTS. Using the dimensions, shear modulus, and average hardening properties of the specimens considered in this study, and assuming $k=1$, a ratio of true stress at bifurcation to true stress at UTS of 1.344 is calculated. Ratios of necking stress to stress at UTS for the analyses performed in the current work are shown in Table 1. Values range from 1.275 to 1.297 for 2D analyses and from 1.185 to 1.225 for 3D analyses. Similarly, Eq. (8)

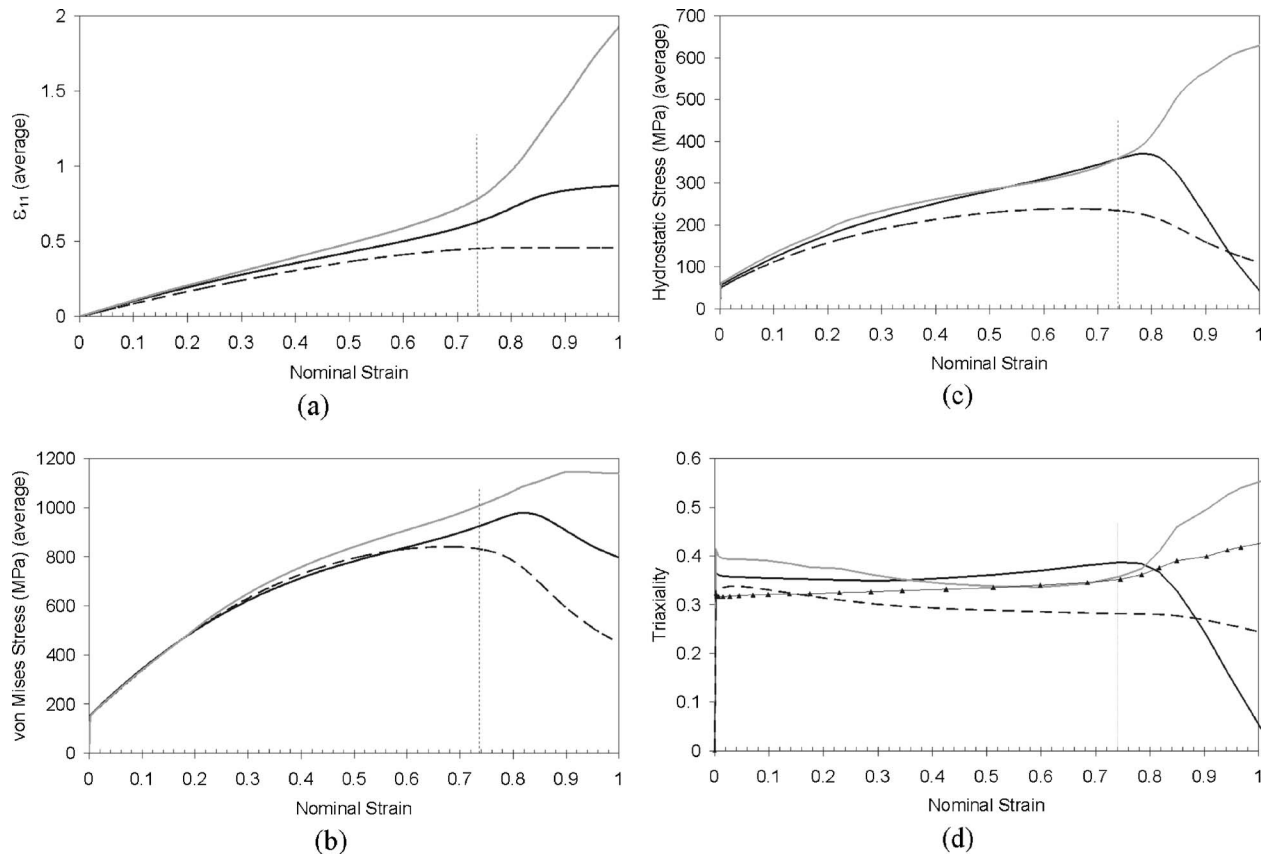


Fig. 14 Plots of average (a) logarithmic strain ε_{11} , (b) von Mises stress, (c) hydrostatic stress, and (d) triaxiality in region remote from necking (broken lines), in region adjacent to necking (black lines) and at neck (gray lines). (d) shows an additional curve for the average triaxiality of the elements on the outer edges of the necked region (black with triangles).

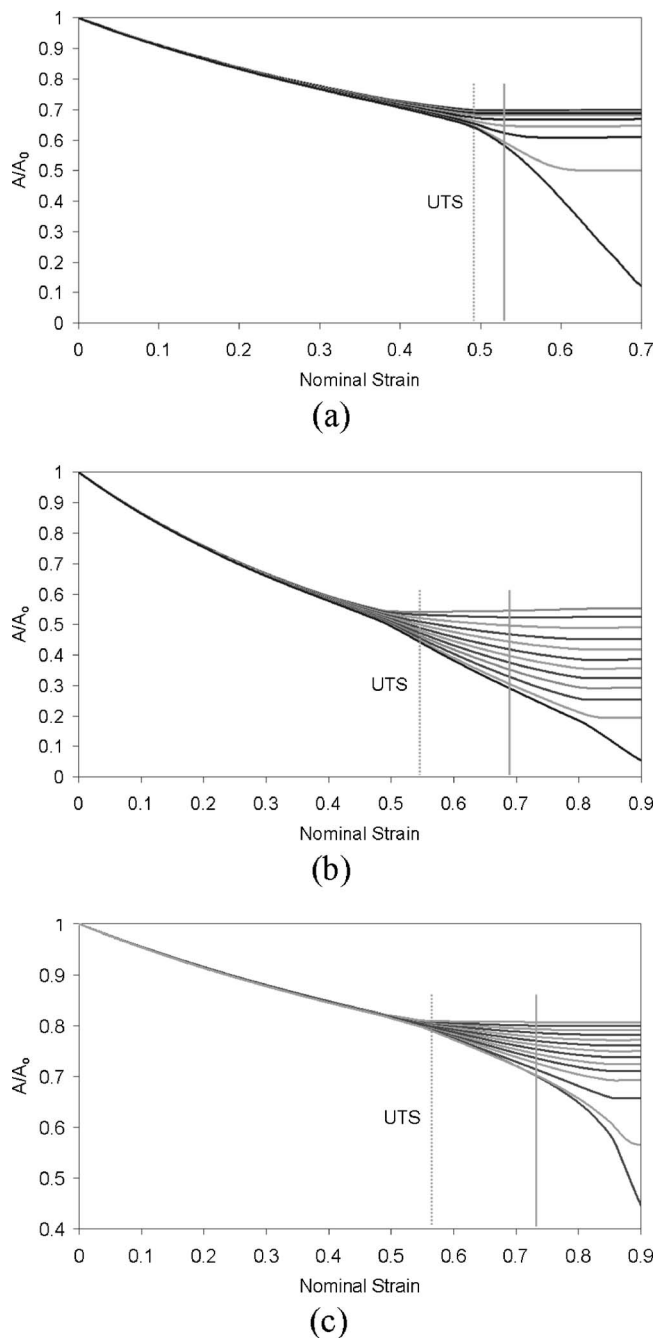


Fig. 15 Plots of section areas normalized by original section areas versus nominal strain for homogeneous struts with geometric imperfections: (a) 3D simulation for tapered strut, (b) 2D simulation for tapered strut, and (c) 2D simulation for notched strut. Unlabeled vertical line corresponds to nominal strain at which Eq. (6) is satisfied and vertical line labeled UTS corresponds to the nominal strain at UTS.

yields a ratio of true strain at bifurcation to true strain at UTS of 1.765. From Table 1, it can be seen that all computed ratios of strain at necking to strain at UTS are far lower than this value.

Discussion and Conclusions

2D analyses exhibit a high dependence on crystal orientation in terms of hardening behavior, UTS, and necking strain. 3D analyses, in particular, cylindrical crystal configurations, reveal far less dependence on crystal orientations. This can easily be explained by the fact that 2D geometries are comprised of far fewer crystals;

thus, crystals orientated favorably toward shear slip and consequent formation of shear bands play a far more dominant role in the deformation of the strut than is the case for the 3D models, where the deformation of such favorably oriented crystals is constrained by many more neighboring crystals that could be less favorably orientated for shear band formation. Second-order elements were shown to yield a lower rate of hardening and UTS than first-order elements. The computed differences are similar for all sets of crystal orientations. The hardening rate and yield stress of 3D is lower than that computed for 2D analyses, resulting from the fact that the 2D generalized plane strain is an artificially stiff formulation.

In contrast to 2D analyses, all *s-e* curves for 3D analyses are identical up to 20% strain, regardless of crystal orientation configuration. After this point, however, the rhombic-dodecahedron crystal configuration exhibits slightly more dependence on crystal orientation and a lower UTS than the cylindrical crystal configuration. It should be noted that both configurations were constructed as a 3D microstructural idealization for a given number of grains across the strut width and depth. This leads to a slightly higher number of grains in the cylindrical configuration (1980) than the rhombic-dodecahedron configuration (1875). If the lower number of crystal orientations associated with the rhombic-dodecahedron configuration were significant, the effects would be apparent in the early stages of hardening as was the case in the 2D analyses. It can be inferred from 2D analyses that element type and mesh density do not effect patterns of dependence on crystal orientations. It is therefore felt that the differences in the two 3D configurations lie in the fact that the square nature of the crystal geometry and alignment in the out-of-plane direction for the cylindrical configuration discourages the formation of shear bands and subsequent necking in the out-of-plane direction.

With regard to comparison of the computational *s-e* curves generated in this study to the polycrystalline data measured experimentally by Okamoto et al. [11], a number of observations should be made. The elastic modulus of the experimental curve is extremely low, suggesting experimental error. Low measured elastic moduli for 316L are not totally unusual; the tensile stress-strain data of Weldon et al. [34] revealed elastic moduli significantly lower than the expected value of (crice) 200 GPa, and this was not due to defects in the material. The elimination of such errors was achieved by the calibration of the experiment and the introduction of a correction factor Cuddy [35]. Even though the elastic modulus in [11] is low, the subsequent hardening rate is almost identical to that computed using the 3D geometries and lower than the hardening rates computed using 2D models.

The experimentally measured stresses are higher than those computed using the 3D models (UTS: ~540 MPa for alloy T in [11] versus ~510 MPa, maximum 3D model prediction). One possible explanation for this is the very low strain rate used in the present study. The experimentally measured strain at UTS of ~75% (for alloy T in [11]) is also higher than the computationally predicted values (~63%, maximum 3D model prediction). This is more than likely due to the fact that the number of grains in our models is far lower than is the case in the macroscopic polycrystalline specimens used experimentally by Okamoto et al. [11]. The additional grains would provide further constraint against the formation of shear bands and, as is evident from this study, increase the strain at UTS. This phenomenon of greater constraint on the development of localized deformation (and, hence, greater macro-scale ductility) with increase in strut thickness (and, hence, the number of grains in the cross section) was investigated and confirmed in the experimental study of Murphy et al. [7] and the 2D computational study of Savage et al. [8]. Given current computational constraints, bearing in mind the extremely long run times required for the simulations reported here, the 3D macroscopic modeling of larger samples, reflexive of the Okamoto et al. polycrystalline samples, is not feasible. The effect of strain rate on hardening and UTS has not yet been numerically established but

is also the subject of an ongoing investigation. It is also worth noting that a recent comprehensive experimental investigation of the mechanical performance of 316L wires and laser-cut tube struts [35] has revealed a general increase in UTS with increase in wire/strut thickness; this corroborates the observation here that the UTS values in [11] for (relatively) large polycrystalline samples are greater than the predicted values for 75 μm thick strut models. In light of all of these considerations, it could be concluded that in overall terms, the model predictions are not unreasonable when compared to the Okamoto et al. data.

The curves computed in this study do not compare favorably to the experimental curves of Murphy et al. [7] for stainless-steel struts of thickness 75 μm in terms of yield strength, hardening rate, UTS, and failure strain. For example, in [7], a yield strength, UTS and strain at UTS of ~ 400 MPa, 650 MPa, and $\sim 25\%$, respectively, were observed. Even the introduction of failure mechanisms, such as void coalescence and intergranular fracture, would not account for discrepancies in yield strength and initial hardening rate. This suggests that the single-crystal material properties used in this study, determined from the single-crystal experimental data of Okamoto et al. [11], are not relevant to the stainless-steel alloy used by Murphy et al. [7]. A study involving the determination of single-crystal properties appropriate for the stent struts tested by Murphy et al. [7] is currently in preparation.

In addition, it should be emphasized that the macroscopic mechanical properties of 316L are highly dependent on the particular alloy composition and heat treatment. This can be seen by contrasting the experimental data of [11,7], quoted above, with that reported in [34], where for an annealed 316L alloy a yield strength of ~ 690 MPa, a UTS of ~ 890 MPa and an elongation at break of $\sim 37\%$ were observed. This significant scatter in experimental data strongly supports the notion that in using such models for material property prediction, appropriate single-crystal properties for the material in question must be used.

Taking all the above computational model predictions versus experimental observation discussion into consideration, it is the authors' opinion that, once a representative 3D model morphology is in place, the specification of the single-crystal material properties is vital for the prediction of reasonable macroscale mechanical properties, such as yield strength and UTS. In fact, this may be practically all that is required for dealing with materials that are relatively defect free. The inclusion of microscale damage and fracture models, which would hugely complicate the computational analysis, may only be necessary for materials with a large defect content, or should the user wish to explicitly model material separation leading up to final rupture. Certainly, for example, the inclusion of damage models in the analysis for improvement in the comparison of the model predictions with the Okamoto et al. data would be futile as they would only serve to reduce the predicted UTS and strain at UTS below existing low values.

In summary, 3D modeling results highlight the inaccuracies inherent in 2D generalized strain crystal plasticity in terms of prediction of yield stress, hardening, and UTS. The unrealistically low number of grains in 2D models allows for the formation of shear bands at lower strains and for an excessive dependence on randomly generated crystal orientations. With regard to the crystal configuration in 3D models, cylindrical grains, due to their unphysical geometry in the out-of-plane direction, provide an excessive inhibition of shear band formation compared to the more realistic rhombic-dodecahedron configuration.

A detailed examination of cross-sectional areas throughout the strut length reveals a distinct pattern of necking characterized by a rapid localized decrease in cross-sectional area in the vicinity of the neck and elastic unloading at all other sections. Such localized deformation initiates at strains significantly higher than strains at UTS, with ratios of true strain at necking to true strain at UTS ranging from 1.21 to 1.44 for 2D analyses and from 1.21 to 1.31 for 3D analyses. Again, the wide range of values for 2D analyses can be related to the artificially low number of grains in such

models. Necking occurs at higher strains for simulations with cylindrical grains than for simulations with rhombic-dodecahedron grains as the former configuration is less conducive to the formation of shear bands. The occurrence of necking post UTS is consistent with the bifurcation studies of Cowper and Onat [16], Hutchinson and Miles [17], and Miles [19] and the numerical study of Tvergaard [36]. For the specimens considered in this study, the ratios of true stress at necking to true stress at UTS compared favorably to the ratio determined from the formulation of Miles [19]. Similar comparisons for true strain ratios do not compare as favorably. This is because the prediction of true strain in the work of Miles [19] is based on the assumption of a constant slope of the true-stress–true-strain curve following UTS. In the present detailed computational study, this assumption is shown to be inappropriate; in fact, the onset of necking is marked by a definite point of inflection in the true-stress–true-strain plots.

The quantity $s(1+e)$ is shown to deviate from the corresponding true-stress values at strains far lower than the strain at UTS, despite the relative homogeneous strut cross sections at UTS. It is also observed that the onset of necking coincides with a slope of the $s-e$ curve of $-s/(1+e)$ for all analyses. Although the rigor of this condition remains to be established, as does its generality, it suggests that the $s-e$ curve can potentially be used to pinpoint the onset of necking. However, initial computations suggest that the formula cannot be readily generalized to necking due to geometric imperfections in homogeneous rate-independent struts. An ongoing investigation is being performed to test the applicability of the formula to a wider range of inhomogeneous strut microstructures for different loading rates, rate exponents, and single-crystal strain-hardening properties. For future developments of detailed computational models of the type considered here, it may be necessary to consider developing enhanced constitutive representations of the material for large localized strain levels to better take account of evolving dislocation structures, as discussed in Hoc et al. [12].

With regard to the computed stress state at the site of necking, a dramatic increase in hydrostatic pressure at the center of the neck is computed. This is consistent with the findings of Tvergaard and Needleman [24] for round bars. The high rate of increase of triaxiality at the center of the necking region compared to the outer edges of the necking region is consistent with the reported findings of Alves and Jones [25] and Hancock and Brown [23]. It is shown by Bao [37] that the increased triaxiality at the center of the neck is a critical factor for ductile crack formation in tensile specimens. Most significant in this result is the fact that the stress patterns at the neck of notched specimens are replicated in the current analyses where necking is entirely due to material inhomogeneity and not artificially induced by geometric features [38].

Acknowledgment

The authors acknowledge support from the Programme for Research in Third Level Institutions (PRTL), administered by the Higher Education Authority (HEA), and the Irish Council for Science, Engineering and Technology (IRCSET). The authors wish to acknowledge the input of Dr. David Quinn, Medtronic Ltd. Galway. The authors also wish to thank Prof. David Hayhurst, Dr. Yueming Liang, and Dr. Alexander Spuskanyuk for several insightful discussions. The simulations in this work were performed on the SGI 3800 high-performance computer at NUI, Galway under an academic license from ABAQUS, Inc.

References

- [1] Brauer, H., Stolpmann, J., Hallmann, H., Erbel, R., and Fischer, A., 1999, "Measurement and Numerical Simulation of the Dilatation Behaviour of Coronary Stents," *Materialwiss. Werkstofftech.*, **30**, pp. 876–885.
- [2] Dumoulin, C., and Cochelin, B., 2000, "Mechanical Behaviour Modelling of Balloon-Expandable Stents," *J. Biomech.*, **3**, pp. 1461–1470.
- [3] Etave, F., Finet, G., Boivin, M., Boyer, J.-C., Rioufol, G., and Thollet, G., 2001, "Mechanical Properties of Coronary Stents Determined by Using Finite

- Element Analysis," *J. Biomech.*, **34**, pp. 1065–1075.
- [4] Tan, L. B., Webb, D. C., Korni, K., and Al-Hassani, S. T. S., 2001, "A Method for Investigating the Mechanical Preproperties of Intracoronary Stents Using Finite Element Numerical Simulation," *Int. J. Cardiol.*, **78**, pp. 51–67.
 - [5] Migliaiavacca, F., Petrini, L., Colombo, M., Auricchio, F., and Pietrabissa, R., 2002, "Mechanical Behaviour of Coronary Stents Investigated Through the Finite Element Method," *J. Biomech.*, **35**, pp. 803–811.
 - [6] McGarry, J. P., O'Donnell, B. P., McHugh, P. E., and McGarry, J. G., 2004, "Analysis of the Mechanical Performance of a Cardiovascular Stent Design Based on Micromechanical Modelling," *Comput. Mater. Sci.*, **31**, pp. 421–438.
 - [7] Murphy, B. P., Savage, P., McHugh, P. E., and Quinn, D. F., 2003, "The Stress-Strain Behaviour of Coronary Stent Struts is Size Dependent," *Ann. Biomed. Eng.*, **31**, pp. 686–691.
 - [8] Savage, P., O'Donnell, B. P., McHugh, P. E., Murphy, B. P., and Quinn, D. F., 2004, "Coronary Stent Strut Size Dependent Stress-Strain Response Investigated Using Micromechanical Finite Element Models," *Ann. Biomed. Eng.*, **32**, pp. 202–211.
 - [9] McHugh, P. E., Asaro, R. J., and Shih, C. F., 1993, "Computational Modelling of Metal Matrix Composite Materials—I. Isothermal Deformation Patterns in Ideal Microstructures," *Acta Metall. Mater.*, **41**, pp. 1461–1476.
 - [10] Bruzzi, M. S., McHugh, P. E., O'Rourke, F., and Linder, T., 2001, "Micromechanical Modelling of the Static and Cyclic Loading of an Al 2124–SiC MMC," *Int. J. Plast.*, **17**, pp. 565–599.
 - [11] Okamoto, K., Yoshinari, A., Kaneda, J., Aono, Y., and Kato, T., 2000, "Production and Property Evaluation of Single Crystal Austenitic Stainless Steels," *Mater. Trans., JIM* **41**(7), pp. 806–814.
 - [12] Hoc, T., Rey, C., and Raphanel, J. L., 2001, "Experimental and Numerical Analysis of Localization During Sequential Test for an IF-Ti Steel," *Acta Mater.*, **49**, pp. 1835–1846.
 - [13] Diard, O., Leclercq, S., Rousselier, G., and Cailletaud, G., 2005, "Evaluation of Finite Element Based Analysis of 3D Multicrystalline Aggregates Plasticity: Application to Crystal Plasticity Model Identification and the Study of Stress and Strain Fields Near Grain Boundaries," *Int. J. Plast.*, **21**, pp. 691–722.
 - [14] Becker, R., 2004, "Effects of Crystal Plasticity on Materials Loaded at High Pressures and Strain Rates," *Int. J. Plast.*, **20**, pp. 1983–2006.
 - [15] Marketz, W. T., Fischer, F. D., and Clemens, H., 2003, "Deformation Mechanisms in TiAl Intermetallics—Experiments and Modelling," *Int. J. Plast.*, **19**, pp. 281–321.
 - [16] Cowper, G. R., and Onat, E. T., 1962, "The Initiation of Necking and Buckling in Plane Plastic Flow," *Proceedings of the Fourth U.S. National Congress of Applied Mechanics*, R. M. Rosenberg, ed., ASME, New York, p. 1023.
 - [17] Hutchinson, J. W., Miles, J. P., 1974, "Bifurcation Analysis of Onset of Necking in an Elastic-Plastic Cylinder Under Uniaxial Tension," *J. Mech. Phys. Solids*, **22**, pp. 61–71.
 - [18] Hill, R., and Hutchinson, J. W., 1975, "Bifurcation Phenomena in Plane Tension Test," *J. Mech. Phys. Solids*, **23**, pp. 239–264.
 - [19] Miles, J. P., 1975, "Initiation of Necking in Rectangular Elastic-Plastic Specimens Under Uniaxial and Biaxial Tension," *J. Mech. Phys. Solids*, **23**, pp. 197–213.
 - [20] Argon, A. S., Im, J., and Needleman, A., 1975, "Distribution of Plastic Strain and Negative-Pressure in Necked Steel and Copper Bars," *Metall. Trans. A*, **6**, pp. 815–824.
 - [21] Mackenzie, A. C., Hancock, J. W., and Brown, D. K., 1977, "Influence of State of Stress on Ductile Failure Initiation in High-Strength Steels," *Eng. Fract. Mech.*, **9**, pp. 167–188.
 - [22] La Rosa, G., Mirone, G., and Risitano, A., 2001, "Effect of Stress Triaxiality Corrected Plastic Flow on Ductile Damage Evolution in the Framework of Continuum Damage Mechanics," *Eng. Fract. Mech.*, **68**, pp. 417–434.
 - [23] Hancock, J. W., and Brown, D. K., 1983, "On the Role of Strain and Stress State in Ductile Failure," *J. Mech. Phys. Solids*, **31**, pp. 1–24.
 - [24] Tvergaard, V., and Needleman, A., 1984, "Analysis of the Cup-Cone Fracture in a Round Tensile Bar," *Acta Metall.*, **32**, pp. 157–169.
 - [25] Alves, M., and Jones, N., 1999, "Influence of Hydrostatic Stress on Failure of Axisymmetric Notched Specimens," *J. Mech. Phys. Solids*, **47**, pp. 643–667.
 - [26] HKS, 1997, *ABAQUS® Theory Manual*, ABAQUS, Providence, RI.
 - [27] Huang, Y., 1991, "A User-Material Subroutine Incorporating Single Crystal Plasticity in the ABAQUS Finite Element Program," Harvard University Report No. MECH 178.
 - [28] Pierce, D., Asaro, R. J., and Needleman, A., 1983, "Material Rate Dependence and Localised Deformation in Crystalline Solids," *Acta Metall.*, **31**(12), pp. 1951–1976.
 - [29] Kysar, J., 2001, "Continuum Simulations of Directional Dependence of Crack Growth Along a Copper/Sapphire Bicrystal Interface, Part I: Experiments and Crystal Plasticity Background," *J. Mech. Phys. Solids*, **49**, pp. 1099–1128.
 - [30] Eleiche, A. M., Albertini, C., and Montagnani, M., 1985, "The Influence of Strain-Rate History on the Ambient Tensile-Strength of AISI Type-316 Stainless-Steel," *Nucl. Eng. Des.*, **88**(2), pp. 131–141.
 - [31] Krempl, E., and Kallianpur, V. V., 1984, "Some Critical Uniaxial Experiments for Viscoplasticity at Room-Temperature," *J. Mech. Phys. Solids*, **32**(4), pp. 301–314.
 - [32] Krempl, E., 1979, "An Experimental-Study of Room-Temperature Rate-Sensitivity, Creep and Relaxation of AISI Type-304 Stainless-Steel," *J. Mech. Phys. Solids*, **27**, pp. 363–375.
 - [33] Kujański, D., Kallianpur, V., and Krempl, E., 1980, "An Experimental Study of Uniaxial Creep, Cyclic Creep and Relaxation of AISI Type-304 Stainless-Steel at Room-Temperature," *J. Mech. Phys. Solids*, **28**, pp. 129–148.
 - [34] Weldon, L. M., McHugh, P. E., Carroll, W., Costello, E., and O'Bradiagh, C., 2005, "The Influence of Passivation and Electropolishing on the Performance of Medical Grade Stainless Steels in Static and Fatigue Loading," *J. Mater. Sci.: Mater. Med.*, **16**, pp. 107–117.
 - [35] Cuddy, H., 2005, "Experimental Analysis of the Influence of Size Scale on the Static Mechanical Properties of Thin Biomedical Grade Stainless Steel Wires and Struts," M.Eng.Sc thesis, National University of Ireland, Galway.
 - [36] Tvergaard, V., 1993, "Necking in Tensile Bars With Rectangular Cross-Section," *Comput. Methods Appl. Mech. Eng.*, **103**, pp. 273–290.
 - [37] Bao, Y. B., 2005, "Dependence of Ductile Crack Formation in Tensile Tests on Stress Triaxiality, Stress and Strain Ratios," *Eng. Fract. Mech.*, **72**, pp. 505–522.
 - [38] Mika, D. P., and Dawson, P. R., 2001, "Polycrystal Plasticity Modeling of Intracrystalline Boundary Textures," *Acta Mater.*, **47**, pp. 1355–1369.

Dynamic Failure Behavior of Polycrystalline Alumina Under Impact Loading

Guowen Yao

Associate Professor
School of Civil Engineering and Architecture,
Chongqing Jiaotong University,
Chongqing, 400074, China

Zhanfang Liu

Professor
Department of Engineering Mechanics,
Chongqing University,
Chongqing, 400044, China

Plate impact experiments and impact recovery experiments were performed on 92.93 wt. % aluminas using a 100 mm dia compressed-gas gun. Free surface velocity histories were traced by a velocity interferometry system for any reflector (VISAR) velocity interferometer. There is a recompression signal in free surface velocity, which shows evidence of a failure wave in impacted alumina. The failure wave velocities are 1.27 km/s and 1.46 km/s at stresses of 7.54 GPa and 8.56 GPa, respectively. It drops to 0.21 km/s after the material released. SEM analysis of recovered samples showed the transit of intergranular microcracks to transgranular microcracks with increasing shock loading. A failure wave in impacted ceramics is a continuous fracture zone, which may be associated with the damage accumulation process during the propagation of shock waves. Then a progressive fracture model was proposed to describe the failure wave formation and propagation in shocked ceramics. The governing equation of the failure wave is characterized by inelastic bulk strain with material damage and fracture. Numerical simulation of the free surface velocity was performed in good agreement with the plate impact experiments. And the longitudinal, lateral, and shear stress histories upon the arrival of the failure wave were predicted, which present the diminished shear strength and lost spall strength in the failed layer. [DOI: 10.1115/1.2722777]

1 Introduction

Since failure waves were first observed propagating in a glass rod under dynamic compression by Bless et al. [1] and in a glass plate under high-pressure impulsive loading by Rasorenov et al. [2] and Kanel et al. [3], a series of plate impact experiments, bar impact experiments, and impact recovery experiments have been performed on a range of glasses under various impact stress diagnosed through embedded manganin stress gauge and strain gauge by Rosenberg et al. [4], Bourne et al. [5,6], and Millett et al. [7], velocity interferometer system for any reflector (VISAR) and high-speed photography by Cazamias et al. [8], He et al. [9], Bless et al. [10], and Bourne et al. [11]. There are diverse physical and mechanical characterizations that identify a failure wave propagating behind the elastic precursor, and there has been a considerable amount of effort expended to explain them.

The failure fronts are generated in silicate and filled glasses at a stress near or below their Hugoniot elastic limits (HELs) and propagate from impact surface to interior at velocities in the range of 1.5–2.5 km/s. It is worth noting that a particular feature of the failure waves is that the velocity is dependent on the impact stress. The failed glass has lower acoustic impedance and sound speed than the intact material. The failed layer nearly loses complete tensile strength, and its shear strength is significantly degraded. The longitudinal stress and transverse strain remain constant across the failure front, but the transverse stress and longitudinal strain are increasing with time in the region behind the failure front. All these variations of material properties across this front provide experimental evidences for the existence of a failure wave phenomenon for glass under plate normal impact loading.

In recent years there has been some wide research made into other brittle materials. Bourne et al. [5,12,13] and Zhang et al. [14] have extended these studies to the polycrystalline ceramics alumina, silicon carbide and titanium diboride, gabbro, and 3D-C/SiC composite materials and have postulated similar

impact-induced fracture front in these brittle materials. There has also been recent discussion of the phenomenon of gradual failure behind the elastic wave in mortar by Grote et al. [15]. Rasorenov et al. [2], Clifton [16], and Grady [17] discussed the mechanism of its formation and propagation and proposed some scientific hypotheses, including shear-induced microcracking, phase transformation, and the transfer of elastic shear strain energy to dilatant strain energy summarized by Brar [18].

Thus far the study on the failure wave has focused on accumulative acquaintance with its physical and mechanical characters monitored in experiments with restriction of experimental conditions of high-pressure loading and monitoring techniques for high resolving wave profiles. And there is not a perfect model to explain the formation and propagation of the failure wave and mechanism of interaction between failure wave and material, even to simulate the basic mechanical response characters behaved in experiments satisfactorily. This is the emphasis and difficulty in study on the failure wave, and main reason why scholars have been interested in failure waves for so many years. In the work presented, we have conducted a matrix of plate impact experiments on alumina monitored by VISAR focused on the rear surface of the sample in seeking to pursue the failure wave in brittle materials other than glass. Then a progressive fracture model was proposed to analyze the failure wave formation and propagation in ceramics under planar impact loading with research into their heterogeneous mesostructures and its result in high singularity in the stress field. Then numerical simulation of the mechanical properties about the failure wave was performed to compare to alumina plate impact experiments.

2 Plate Impact Experiments

Plate impact experiments on alumina specimens were carried out on the 100 mm light gas gun at the National Defense Science and Technology Key Laboratory of China Academic Engineering Physics. Impact velocities were measured to an accuracy of 1.5% using three pairs of electric signal pins at different distances away from the impact surface. The copper flyers and targets were circular with different diameters of 94 mm and 100 mm, with their

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received February 14, 2006; final manuscript received January 7, 2007. Review conducted by Sridhar Krishnaswamy.

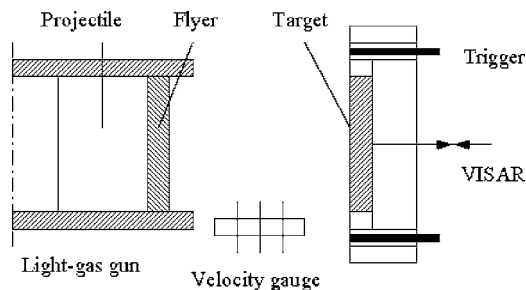


Fig. 1 Plate impact experimental schematic with VISAR

two cut faces polished in order to ensure smoothness of the impact and measurement area. Free surface velocity histories were traced using VISAR with a fringe constant 101 m/s/fringe and a measured response time 1.5 ns. The free surface of target was polished and aluminized with a layer 5000 Å in thickness to strengthen the reflection of incident laser (see Fig. 1, which shows a schematic of the experimental setup). The impact recovery experiments were also performed to study microstructures of impacted samples. Cushion rubber was filled in the target room to absorb the dynamic energy of flyer and target. The flyer and target will be embedded in rubber.

The alumina samples manufactured by No.799 Factory in Sichuan, which consist of 92.93 wt. % alumina by weight and a

small amount of silicon dioxide, calcium oxide, and lanthana analyzed by energy spectrum at National Key Laboratory in Chengdu Science and Technology University. The relevant parameters of the specimens are density $3.896 \times 10^3 \text{ kg/m}^3$, longitudinal sound speed 9.259 km/s, and a shear wave velocity 5.557 km/s, respectively. The longitudinal sound speed in copper flyer is 4.60 km/s, and the thickness of flyers and targets range from 4.0 mm to 6.1 mm. The acoustic impedance ratio of flyer and target is 1.14, then a long-enough duration pulse generates at the impact surface to avoid the unloading wave propagating into targets from flyers. A summary of experimental conditions and results are presented in Table 1.

Figure 2 shows reduced VISAR data by software from the experiments of shots 405 and 425 under shock stresses 7.54 GPa and 8.56 GPa. These profiles indicate that the alumina specimens did not spall. The distinct feature of note on the traces is the slight recompression signal pointed on top of the stress wave. This velocity jump behaves beyond the elastic behaviors because there is not a reflected tensile pulse recorded in the profiles and the time interval between the start of free surface motion and the moment of this reloading signal is less than the elastic wave reverberation time in the sample. And, as a typical brittle material, alumina does not behave like plastic deformation in the macroscopic; thus, this inelastic behavior does not characterize plasticity, though the free surface velocity profile has two-wave structure. The additional weak compression wave is associated with a reflection from a layer of material that has a dynamic impedance lower than that of the intact alumina, and this material layer bordering the interface does not pass tensile stresses. Therefore, we conclude that the shock-compressed alumina is comminuted behind this interface. This phenomenon is akin to the failure wave, which has been observed to occur in glasses under shock compression shown in Fig. 3 by He et al. [9], Kanel et al. [19], and Zhao et al. [20].

On the assumption that the moving speed of failed layer boundary is the failure wave velocity C_F , a simple evaluating equation

Table 1 Parameters of plate impact experiments

Parameters	Shot 405	Shot 425
Impact velocity (m/s)	397.8	448.8
Impact stress (GPa)	7.54	8.56
Impactor thickness (mm)	4.14	6.10
Target thickness (mm)	6.08	6.04

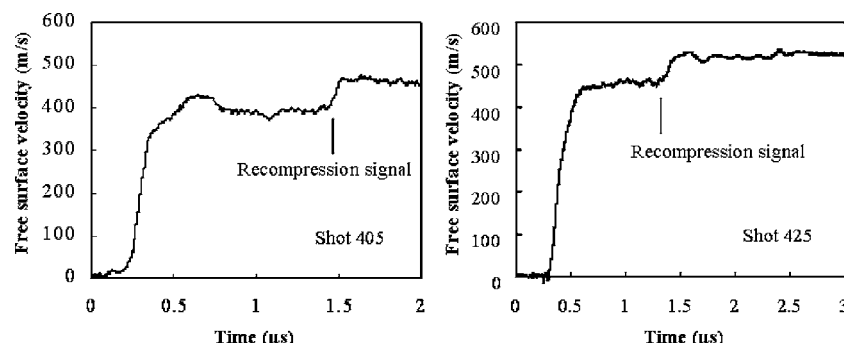


Fig. 2 Free surface velocity profiles showing small recompression for shots 405 and 425

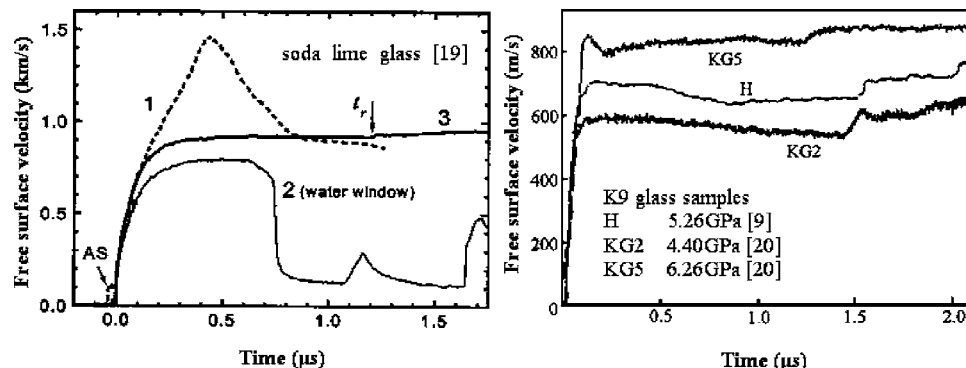


Fig. 3 Free surface velocity profiles of glass monitored by VISAR [9,19,20]

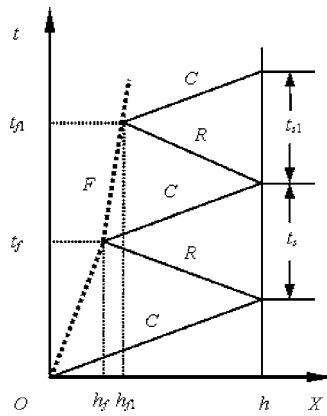


Fig. 4 Propagation and interaction of compression, rarefaction and failure waves

for C_F has been derived as the following (see Fig. 4, which shows a diagram of elastic wave and failure wave propagating). The thickness of the failed layer h_f is determined from the measured time interval of t_s through

$$h_f = h - \frac{1}{2} C_P t_s \quad (1)$$

where h is the sample thickness and C_P is the longitudinal wave speed in alumina specimen. Then the failure wave velocity C_F can be estimated by

$$C_F = \frac{h_f}{h/C_P + t_s/2} \quad (2)$$

It implies that the failure wave has propagated at a velocity of 1.27 km/s in shot 405 km/s and 1.46 km/s in shot 425 on average before the moment t_f . The free surface velocity history from VISAR measurements has shown that the failure front propagates

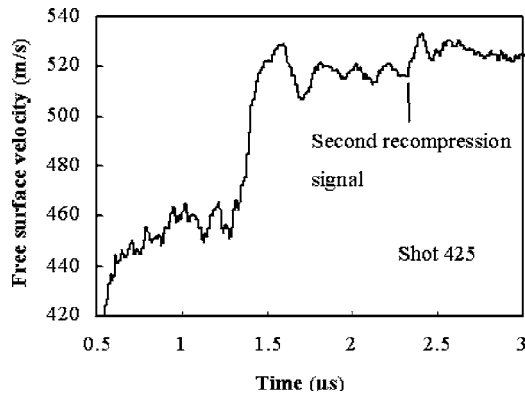


Fig. 5 Expanded region of free surface velocity profile showing second small recompression signal from shot 425

at a speed much lower than longitudinal stress wave velocity, depending on the peak shock stress.

The free surface velocity profile from shot 425 is analyzed further in expanded region, and there is another smaller recompression signal observed following the first (see Fig. 5, which shows second smaller recompression indicated by a narrow). This can be explained if the reflected rarefaction wave from rear surface is reflected again on the failure layer and then reflected on rear surface where a weak jump of velocity is produced at the same time. During the time interval t_{s1} of two recompression signals, the distance of failure layer expanded can be determined through

$$\Delta h_f = h_{f1} - h_f = h - \frac{1}{2} C_P t_{s1} - h_f \quad (3)$$

Then, the average velocity C_{F1} of failure wave propagating from the moment t_f to t_{f1} can be estimated from the measured time interval of t_{s1} by

$$C_{F1} = \frac{\Delta h_f}{(t_s - t_{s1})/2} \quad (4)$$

This implies that the failure wave has propagated at an average velocity of 210 m/s in shot 425 following unloading by the reflected rarefaction wave. This unloading slows down and even eventually arrests the failure procedure in material and results in great lowness in the failure wave propagating.

3 Microscopic Observations of Alumina Samples

To explain the failure process of shock-compressed polycrystalline ceramics in a mesoscope, initial and soft-recovered samples were scanned by S530 scanning electron microscope (SEM). Each fragment was cut in the center along a plane parallel to impact surface with 0.2 mm distance to impact surface. Figure 6(a) shows the microstructures of initial 92.93 wt. % alumina. Grains and intergranular pores distribute randomly with diameters 1–15 μm . Intergranular glassy phase is distinct in compact area. And initial porosity is 5.68%, determined by metallurgical analysis software. Pores and glasses weaken mechanical capabilities, and these heterogeneous mesostructures result in high singularity in stress distribution. Figure 6(b) shows intergranular microcracks in recovered sample after 5.76 GPa loading and Fig. 6(c) shows transgranular microcracks in recovered sample after 8.65 GPa loading. Microcracking transmits from intergranular to transgranular with increasing impact compression. Alumina grains begin to fragment with transgranular microcracks, and original pores begin to collapse. And discontinuous microcracks induce dilation after unloading.

The polycrystalline ceramics are heterogeneous in mesoscope. There are many pores, microcracks, and other defaults inducing high singularity in stress distribution. Once the local stress exceeds the threshold, the original microcracks will grow up along the pores and crystal boundaries, and new microcracks will nucleate in ceramics under shock loading. The original and nucleated microcracks grow up and expand, then excite the neighbor microcracks nucleation and expansion. Thus, the failure wave appears and propagates from impact surface to interior of specimen and at higher velocity under stronger shock loading. In essence, the fail-

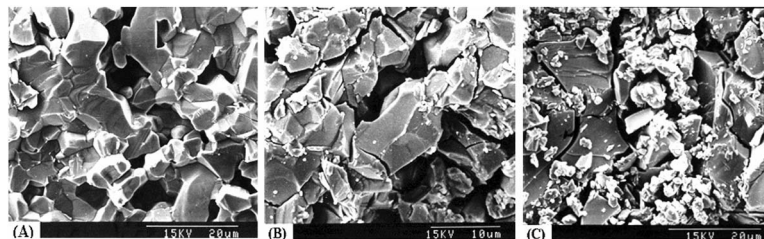


Fig. 6 SEM micrographs of (a) initial and recovered alumina samples under (b) 5.76 GPa, and (c) 8.65 GPa shock loading

ure wave is characterized by moving damage or fracture zone of material that presented by microcracking system in mesoscope, and it is also called after fracture wave by Resnyansky et al. [21].

4 Progressive Fracture Model

4.1 Governing Equation of the Failure Wave. The mesoscopic deformation of discontinuous microcrack interfaces and collapse of original pores in impacted ceramics can be equivalent to the inelastic bulk strain ε_V^{ne} , statistically, which consists of dilatant bulk strain ε_V^c from nucleation and expansion of microcracks and condensed bulk strain ε_V^p from collapse of original pores. Thus, the state variable ε_V^{ne} at the failure front can be governed by the wave equation of

$$\frac{\partial \varepsilon_V^{ne}}{\partial t} + C_F \frac{\partial \varepsilon_V^{ne}}{\partial X} = \varepsilon_{V0}^{ne} \quad (5)$$

where ε_{V0}^{ne} is source or convergence of inelastic volume, which controls the attenuation of the failure wave propagation, and $C_F(X, t) = C_0 \xi(X, t)$ is the failure wave speed, which depends on the strength of shock loading. The material parameter C_0 is the failure wave speed when the shock loading is up to the specimen's HEL and $\xi(X, t)$ denotes the strength of shock loading through

$$\xi = \frac{\tau}{\tau_{HEL}} H(\tau - \tau_{THD}) \quad (6)$$

where H is the heaviside function, τ_{THD} is the threshold of the failure wave formation, τ_{HEL} is the shear stress when the shock loading is equal to the specimen's HEL, which corresponds with transition from intergranular to transgranular microcracking in ceramics, and $\tau(X, t)$ is the maximum shear stress in mesoscope. For brittle polycrystalline ceramics, the inelastic deformation and failure response are governed by mean stress and shear stress; thus, $\tau(X, t)$ can be expressed by

$$\tau = \frac{\sqrt{3}}{2} \Phi(\alpha I_1 + \sqrt{J_2}) \quad (7)$$

where $I_1 = \sigma_{ii}$, $J_2 = 1/2 \sigma'_{ij} \sigma'_{ij}$, α is a positive material parameter, and Φ defines the stress singularity induced by heterogeneous mesostructures. In glass, the failure wave was observed with impact stress up to half the HEL [22]. Thus, Φ is equal to 0.5. Equation (7) can be rewritten under a one-dimensional (1D)-strain condition as

$$\tau = \frac{1}{2} \Phi[\sqrt{3} \alpha (\sigma_1 + 2\sigma_2) + |\sigma_1 - \sigma_2|] \quad (8)$$

where σ_1 and σ_2 are longitudinal and lateral stresses.

Equation (6) is criterion of dynamic failure for ceramics. Damage nucleates when shear stress $\tau(X, t)$ reaches the threshold τ_{THD} ; thus, there are inelastic deformation and failure in polycrystalline ceramics under shock loading below their HELs. And velocity of the failure wave increases with increasing pressure, just as expressed in Eqs. (6) and (7).

4.2 Conservation Equations of the Failure Wave. There are some basic assumptions for the failure wave that it is a discontinuous surface and has no apparent thickness. Figure 7 illustrates a failure wave and the elastic precursor. Ahead of the failure front and behind the elastic precursor, the density is ρ_1 , the stress is σ_1 , the strain is ε_1 , the particle velocity is v_1 , and the internal energy is e_1 ; behind the failure front they are ρ , σ , ε , v , and e , respectively. If one considers the failure front propagation from position A to B in time interval dt . During this time interval, the front has moved by a distance of $C_F dt$. Based on the conservation of energy in $\overline{AB}$ during dt , the equation for the conservation of energy at the failure front is drawn in Lagrangian form as

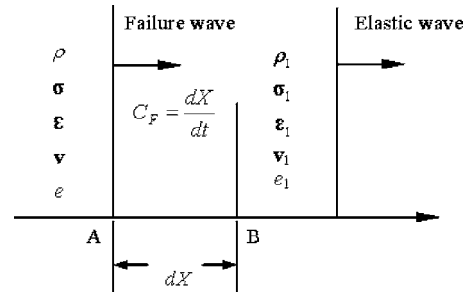


Fig. 7 Schematic of the failure wave and elastic precursor for the conservation of mass, momentum, and energy

$$J = C_F D + C_F [\sigma : d\varepsilon^e] + \frac{1}{2} \rho_1 [\mathbf{v} \cdot \mathbf{v}] \quad (9)$$

where $[\Psi]$ denotes $[\Psi - \Psi_1]$, the density of energy flow is $J = -[\sigma \cdot \mathbf{v}] \cdot \mathbf{n}$, and the ratio of energy dissipation is

$$D = e^c + e^p = \sigma : d\varepsilon^c + \sigma : d\varepsilon^p \quad (10)$$

And the equations for the conservation of mass and momentum at the failure front are set up in Lagrangian form as

$$[\mathbf{v}] = -C_F [\varepsilon] \quad (11)$$

$$[\sigma \cdot \mathbf{n}] = -\rho_1 C_F [\mathbf{v}] \quad (12)$$

4.3 Constitutive Relations of Ceramics. The inelastic dilatant strain increment due to nucleation and expansion of microcracks in brittle material is calculated from linear Drucker-Prager model as follows

$$d\varepsilon_V^c = d\varepsilon^c \frac{1}{3 - \tan \psi} \left[\sqrt{\frac{3}{J_2}} (\sigma_1 - \sigma_2) + \tan \psi \right] \quad (13)$$

where $d\varepsilon^c$ is the equivalent inelastic strain increment which is equal to $|d\varepsilon_1^c|$ in the uniaxial compression case, and ψ is the material's friction angle. Mean stress influences the yield of brittle materials by inner friction, so the bulk strain is inelastic of ceramics under mean stress. In general, the dilatant bulk strain from Eq. (13) is higher than that in experiment; thus, ψ can be assumed to be material's dilation angle. For brittle granule solid, its dilation angle is smaller than the friction angle.

When shear stress reaches the failure threshold τ_{HEL} , nucleation and expansion of intergranular microcracks transit to transgranular microcracks in ceramics and original pores begin to collapse. Then inelastic deformation is gradually governed by mean stress from shear stress, just like sands. According to the conclusion that the bulk strain and porosity of sand soil is in direct proportion to the denary logarithm of mean stress, the relation of porosity n of fragment ceramics and mean stress p can be written as

$$n = n_{HEL} - C_c \lg \frac{p}{p_{HEL}} \quad (14)$$

where n_{HEL} is the porosity of the failed layer under shock loading up to its HEL and C_c is a compressive coefficient. Then the inelastic bulk strain increment from collapse of pores $d\varepsilon_V^p$ is equal to dn . The stress state when pores collapse satisfies the failure criterion; thus, n_{HEL} can be derived by original porosity n_0 and the failure volume V_{HEL} of brittle material under shock loading up to its HEL, i.e.,

$$n_{HEL} = \frac{V_{HEL}}{V_0} n_0 \quad (15)$$

In order to show the lateral stress history of the failed layer under 1D-strain condition, the material failure ratio is defined by

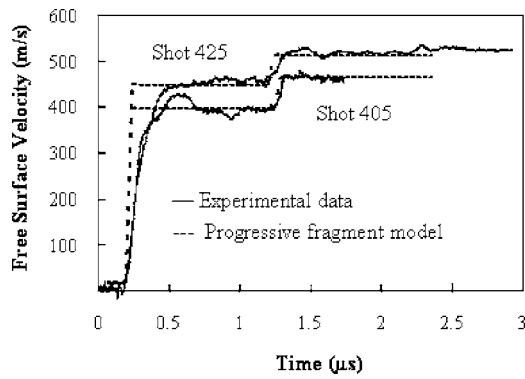


Fig. 8 Free surface velocity histories of alumina specimens under shock loading

$$\kappa = \frac{|\epsilon_{Vd}^{ne}| + |\epsilon_{Vc}^{ne}|}{|\epsilon_{Vd}^{e}| + |\epsilon_{Vd}^{ne}| + |\epsilon_{Vc}^{ne}|} \quad (16)$$

where ϵ_V^e is elastic strain behind the elastic precursor but before the failure wave. κ trends to 1, and Poisson's ratio ν trends to 0.5 with increasing of shock compression. Thus, general Poisson's ratio ν' in the failed layer is written as

$$\nu' = \nu + \left(\frac{1}{2} - \nu\right)\kappa \quad (17)$$

So, the lateral stress can be expressed as

$$\sigma_2 = \sigma_3 = \frac{\nu'}{1 - \nu'} \sigma_1 \quad (18)$$

With increasing of shock loading and fragmentation of the failed layer, lateral stress σ_2 trends to longitudinal stress σ_1 , and thus, shear stress trends to zero when material's bulk deformation from mean stress is dominant other than shear deformation.

5 Simulation of the Failure Wave

Alumina grains begin to fragment with transgranular microcracks and original pores begin to collapse in shots 405 and 425 with shock stresses 7.54 GPa and 8.56 GPa, respectively. Then inelastic deformation is gradually governed by mean stress. The material dilation angle ψ of 92.93 wt % alumina is assigned to be 40 deg, then $\alpha = \sqrt{3} \tan \psi / 9 = 0.1615$. The failure ratios of shots 405 and 425 in their failed layers are 0.320 and 0.556, respectively. Compressive coefficient C_c is defined by Eq. (15). According to metallurgical analysis on micrographs of impact recovery specimens scanned by electron microscope, the porosity is zero with most grains in fragments and pores stuffed when impact stress up to 1.5HEL. Figure 8 shows the free surface velocity histories with recompression signal of 92.93 wt % alumina in shots 405 and 425, and the progressive fracture modeling is in good agreement with experiments. Figure 9 shows the longitudinal stress, lateral stress, and shear stress histories of shots 405 and 425 at 1.16 mm into the targets, in which we see that the lateral stress increases greatly upon the arrival of the failure wave. The longitudinal stress remains constant across the failure front, therefore the shear stress decreases somewhat in the failed layer, and the material shear strength as well. The stress evolvments above agree well with that of impacted SiC ceramics traced by Bourne et al. [13].

6 Summary

Ceramics are extensively applied to national defense engineering and military science as effective armor defense with their excellent physical and mechanical capabilities, especially higher dynamic elastic threshold and acoustic velocity than metals. We performed plate impact experiments of 92.93 wt. % aluminas with

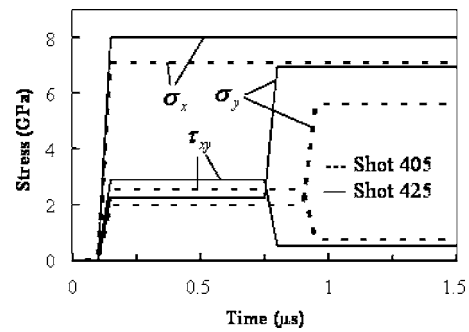


Fig. 9 Longitudinal, lateral, and shear stresses histories of specimens under shock loading simulated by progressive fracture model

100 mm diameter compressed-gas gun and the free surface velocities were traced by VISAR. There is a reloading signal observed in free surface velocity, which indicates the failure wave propagation behind the elastic precursor. The failure wave propagates at a speed much lower than longitudinal stress wave velocity, depending on the peak shock stress. And the failed layer has much lower dynamic impedance than that of the intact material. The unloading by the reflected rarefaction wave slows down and even eventually arrests the failure front propagating in alumina. SEM analysis of intact samples shows heterogeneous mesostructures, and SEM analysis of soft-recovered samples shows transit of intergranular microcracks to transgranular microcracks with increasing shock loading. The failure wave is a continuous fracture or damage front, which may be associated with nucleation and expansion of microcracks from impact surface to interior during the propagation of shock waves.

Based on the study of mesostructure heterogeneity and its induced high singularity in the stress field in ceramics, the inelastic bulk strain is decomposed into two parts of dilatant bulk strain from nucleation and expansion of discontinuous microcracks and condensed bulk strain from collapse of original pores. And the criterion of damage and failure of ceramics is expressed by shear stress in mesoscope. Then governing equation of the failure wave is built up, which characterized by the inelastic bulk strain and dynamic constitutive relations of shocked alumina as well. The material failure ratio proposed relates the longitudinal stress with lateral stress under 1D-strain condition very well. Numerical simulation of the failure wave formation and propagation in impacted ceramics is in good agreement with 92.93 wt. % alumina impact experiments. And the progressive fracture model predicts the longitudinal stress, lateral stress, and shear stress histories across the failure wave.

Acknowledgment

This study was supported by the Natural Science Foundation of China Grant No. 10572158, and NSAF Grant No. 10376046.

References

- [1] Bless, S. J., Brar, N. S., and Rosenberg, Z., 1990, "Failure of Ceramic and Glass Rods Under Dynamic Compression," *Shock Compression of Condensed Matter—1989, 6th American Physical Society Topical Conference on Shock Compression of Condensed Matter*, S. C. Schmidt et al., eds., New Mexico, pp. 939–942.
- [2] Rasorenov, S. V., Kanel, G. I., Fortov, V. E., and Abashev, M. M., 1991, "The Fracture of Glass Under High-Pressure Impulsive Loading," *High Press. Res.*, **6**, pp. 225–232.
- [3] Kanel, G. I., Rasorenov, S. V., and Fortov, V. E., 1992, "The Failure Waves and Spallations in Homogeneous Brittle Materials," *Shock Compression of Condensed Matter—1991, 7th American Physical Society Topical Conference on Shock Compression of Condensed Matter*, S. C. Schmidt et al., eds., VA, pp. 451–454.
- [4] Rosenberg, Z., Bourne, N. K., and Millett, J., 1996, "Direct Measurements of Strain in Shock-Loaded Glass Specimens," *J. Appl. Phys.*, **79**, pp. 3971–3974.
- [5] Bourne, N. K., Millett, J., and Rosenberg, Z., 1997, "On the Origin of Failure

- Waves in Glass," J. Appl. Phys., **81**, pp. 6670–6674.
- [6] Bourne, N. K., Millett, J., Rosenberg, Z., and Murray, N., 1998, "On the Shock Induced Failure of Brittle Solids," J. Mech. Phys. Solids, **46**, pp. 1887–1908.
- [7] Millett, J., Bourne, N. K., and Rosenberg, Z., 2000, "Measurements of Strain in a Shock Loaded, High-Density Glass," *Shock Compression of Condensed Matter—1999*, AIP Conf. Proc., **505**, pp. 607–610.
- [8] Cazamias, J. U., Fiske, P. S., and Bless, S. J., 2000, "Sound Speeds of Post-Failure Wave Glass," *Fundamental Issues and Applications of Shock-Wave and High-Strain-Rate Phenomena, EXPLOMET 2000*, K. P. Staudhammer et al., eds., New Mexico, pp. 173–179.
- [9] He, H. L., Jing, F. Q., Jin, X. G., and Kanel, G. I., 1998, "Microstructure Damage of Glasses Under Shock Wave Compression," J. High Press. Phys., **12**, pp. 241–249 (in Chinese).
- [10] Bless, S. J., Brar, N. S., Kanel, G. I., and Rosenberg, Z., 1992, "Failure Waves in Glass," J. Am. Ceram. Soc., **75**, pp. 1002–1004.
- [11] Bourne, N. K., Rosenberg, Z., and Field, J. E., 1995, "High-Speed Photography of Compressive Failure in Glasses," J. Appl. Phys., **78**, pp. 3736–3739.
- [12] Bourne, N. K., Rosenberg, Z., and Field, J. E., 1998, "Surface Fracture Zones in Shock-Loaded Polycrystalline Ceramics," *Shock Compression of Condensed Matter—1997*, AIP Conf. Proc., **429**, pp. 493–496.
- [13] Bourne, N. K., Millett, J., and Pickup, I., 1997, "Delayed Failure in Shocked Silicon Carbide," J. Appl. Phys., **81**(9), pp. 6019–6023.
- [14] Zhang, Q. M., Huang, F. L., and Han, L. M., 2000, "Failure Wave Motion of 3D-C/SiC Composites Subjected to Shock Compression," Chin. Sci. Bull., **45**, pp. 408–411.
- [15] Grote, D. L., Park, S. W., and Zhou, M., 2001, "Experimental Characterization of the Dynamic Failure Behavior of Mortar Under Impact Loading," J. Appl. Phys., **89**, pp. 2115–2123.
- [16] Clifton, R. J., 1993, "Analysis of Failure Waves in Glasses," Appl. Mech. Rev., **46**(12-1), pp. 540–546.
- [17] Grady, D. E., 1994, "Dynamic Failure of Brittle Solids," Sandia Technical Report No. TMDG0694.
- [18] Brar, N. S., 2000, "Failure Waves in Glass and Ceramics Under Shock Compression," *Shock Compression of Condensed Matter—1999*, AIP Conference Proceedings, M. D. Furnish et al., eds., Utah, AIP, 505, pp. 601–606.
- [19] Kanel, G. I., Bogatch, A. A., Razorenov, S. V., and Zhen, Chen, 2002, "Transformation of Shock Compression Pulses in Glass Due to the Failure Wave Phenomena," J. Appl. Phys., **92**(9), pp. 5045–5052.
- [20] Zhao, J. H., Sun, C. W., Zhao, F., Duan, Z. P., et al., 2002, "Velocity Overshoot of Rear Free-Surfaces of Glass Under Impact," *Explo. Shock Waves*, **22**(1), pp. 72–78 (in Chinese).
- [21] Resnyansky, A. D., Romensky, E. I., and Bourne, N. K., 2003, "Constitutive Modeling of Fracture Waves," J. Appl. Phys., **93**, pp. 1537–1545.
- [22] Partom, Y., 1998, "Modeling Failure Waves in Glass," *Int. J. Impact Eng.*, **21**(9), pp. 791–799.

Magnetic Force and Thermal Expansion as Failure Mechanisms of Electrothermal MEMS Actuators Under Electrostatic Discharge Testing

Jonathan D. Weiss

Sandia National Laboratories,
P.O. Box 5800,
Mail Stop 1081,
Albuquerque, NM 87185-1081
e-mail: jdweiss@sandia.gov

Like microelectronic circuits, microelectromechanical systems (MEMS) devices are susceptible to damage by electrostatic discharge (ESD). At Sandia National Laboratories, polysilicon electrothermal MEMS actuators have been subjected to ESD pulses to examine that susceptibility. Failures, in the form of cracks at points of high stress concentration, occurred that could not be explained by thermal degradation of the polysilicon caused by excessive heating, or by excessive displacement of the legs of the actuator of the same nature that occur in normal operation. One hypothesis presented in this paper is that the internal magnetic forces between the legs of the actuator, resulting from the ESD-associated high current pulses, might produce vibrations of amplitude sufficient to produce these cracks. However, a dynamic analysis based on simple beam theory indicated that such cracks are unlikely to occur, except under rather extreme conditions. On the other hand, these same current pulses also cause resistive heating of the legs and, therefore, thermally induced compression that can lead to buckling. Buckling stresses, particularly when augmented by magnetic forces, can readily explain failure. Both the magnetic and thermal analyses were performed using the human body model and the machine model of ESD. A justification for ignoring shuttle motion and eddy currents induced in the substrate during the ESD pulse is presented, as well.

[DOI: 10.1115/1.2723813]

Keywords: electrothermal actuator, micromechanics, failure analysis, magnetic forces, thermal expansion, buckling of beams

Introduction

Electrothermal actuators have been extensively studied as a means of generating rectilinear motion in microelectromechanical systems (MEMS) [1,2] because of their ability to produce higher output force and larger displacement at lower voltages than their electrostatic counterparts. They operate through the thermal expansion caused by direct resistive heating. However, like integrated circuits, they can be severely damaged by electrostatic discharge (ESD) and electrical overstress (EOS) [3]. The former event involves the passage of a very high current through the device for a time much too short for thermal conduction to be consequential, while the latter is essentially a steady-state event involving currents that are higher than would flow through the device in normal operation. Recent ESD testing at Sandia National Laboratories of MEMS-based polysilicon electrothermal actuators produced failures that appear to be unrelated to the thermal degradation of polysilicon (oxidation, surface roughening, plasticity, and, in the extreme cases, melting) that had been associated with failures during EOS testing using progressively higher current [4,5].

Obtained from Ref. [4], Fig. 1 is a scanning electron micrograph of a typical electrothermal actuator, while a line drawing of the same device (absent the guide and the gauge) is contained in the top half of Fig. 2. The bottom half of that figure is a typical cross section of a polysilicon leg. In operation, the object to be

displaced is attached to the top of the shuttle, and upward motion is guaranteed by means of a slight inclination of the legs ($\sim 1^\circ$ deg, although it looks larger in Fig. 1). The thermal expansion of these resistive elements is caused by the total time-dependent current flowing through them, $I(t)$. Symmetry of construction and the presence of a guide basically guaranteed no motion of the shuttle parallel to the legs. The ESD failures observed were brittle fractures that occurred at the points of high stress concentration where the polysilicon legs were attached to either the anchor or the shuttle. It is noted in Ref. [4] that a calculation indicated that it was impossible for the displacement of the shuttle, as would occur in normal operation, to be large enough to produce such stresses, even if the polysilicon legs were heated to their melting point ($\sim 1415^\circ\text{C}$). This result eliminates excessive displacement of the shuttle as the cause of the failure. However, no specific explanation of the failure was offered, although the authors recognized that the motion of the relatively massive shuttle was probably negligible during the very brief time of an ESD pulse, and that this constraint probably contributed to the failure. Aside from the observation of ESD failures, no detailed failure analysis was performed on the actuators to determine why some failed and others did not.

In this paper, two failure mechanisms are proposed and analyzed that require neither motion of the shuttle nor high temperatures that last long enough to produce thermal degradation of the silicon (above $\sim 700^\circ\text{C}$ for ~ 1 min, according to Ref. [5], as opposed to the microsecond timescale relevant to ESD). These are magnetic forces and thermal expansion, which will act synergistically.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received October 6, 2005; final manuscript received January 31, 2007. Review conducted by Matthew R. Begley.

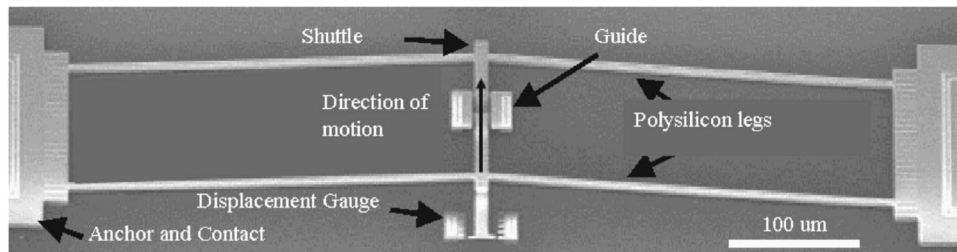


Fig. 1 Scanning electron micrograph showing a plan view the electrothermal actuator

A magnetic force is generated when current is sent through the polysilicon legs, causing them to attract one another and setting them into vibration. Being internal, this force produces no bulk motion of the system, in particular the shuttle. Can the currents be large enough during ESD testing that the vibrating beams will experience stress at the anchor points sufficient to break them? As for thermal expansion, it will cause a length change in the legs that tends to be constrained by the surrounding structures. The result is a lateral instability, or tendency to buckle, that amplifies the effects of the magnetic force. Will this amplification be large enough to fracture the legs at the anchor points?

Mechanical Uncertainties

Two important characteristics of the legs are needed for the analysis, which are not readily available. The first is the stress concentration factor at the anchor points. This factor multiplies the stress obtained from the conventional theory of flexure. It has been determined photoelastically and depends on factors like the radius of curvature of the leg at the anchor point, shoulder widths, other beam dimensions, and the dimensions of the support [6]. There may be no comprehensive theory of this effect and it may be necessary to perform a numerical analysis of stress in order to

treat it accurately in a given situation. However, based on the dimensions of the legs, an estimate of the radius of curvature, and the empirical data in Ref. [6], a stress concentration factor of 1.6 is reasonable. It could be 50% larger.

The second factor is the ultimate tensile strength of the polysilicon. A typical value for this number is around 2 GPa [7], but, as a brittle material, it depends noticeably on surface roughness, surface flaws, surface area, and grain structure. Furthermore, as noted in Ref. [7], the strength is layer dependent, with the first layer ("Poly 1") of the electrothermal actuator typically being the weakest. Its ultimate tensile strength varies from about 1–1.5 GPa. Nevertheless, 2 GPa will be used as a rough benchmark of the ultimate tensile strength.

Current Wave Forms For ESD Testing

As discussed in Ref. [8], a primary example of ESD through a sensitive device is from a human being who has been charged by, for example, walking on a carpet and is discharged by touching the device. A simulation of this event is known as the human body model (HBM), which is implemented with the circuit shown in the top half of Fig. 3. This circuit has become a standard of the ESD association. A 100 pf capacitor is slowly charged by a high-

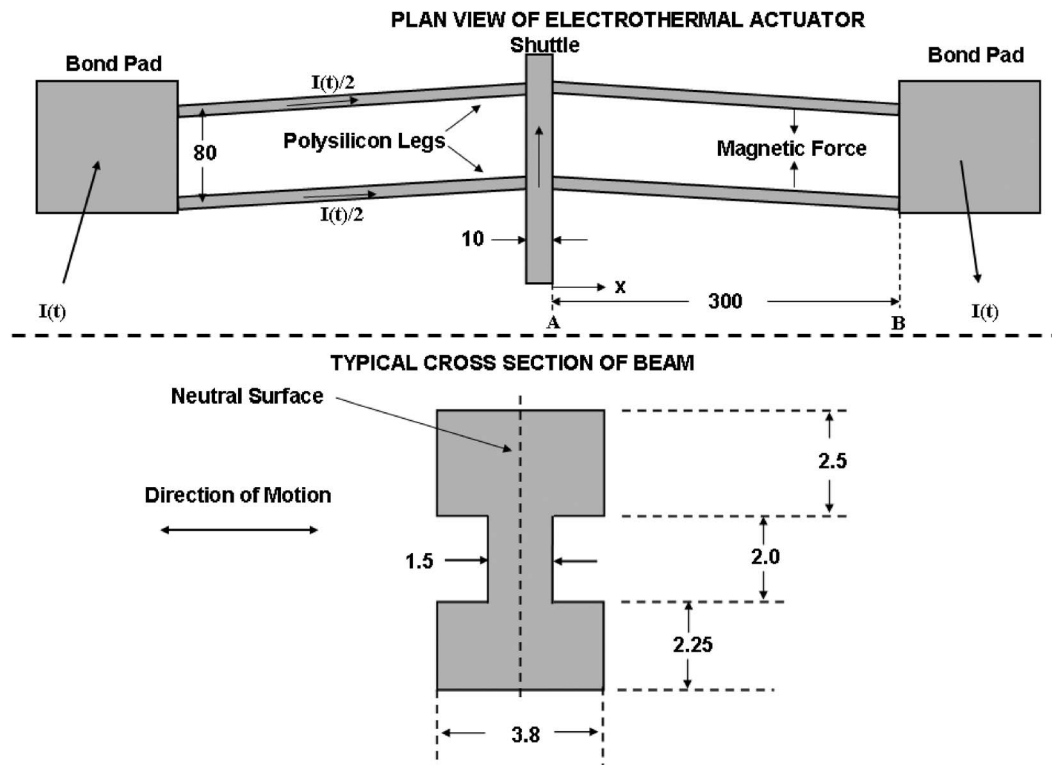


Fig. 2 Line drawing of Fig. 1 and a cross section of its polysilicon legs. The time-dependent current, $I(t)$, is assumed to be equally divided between the legs. All dimensions are in microns.

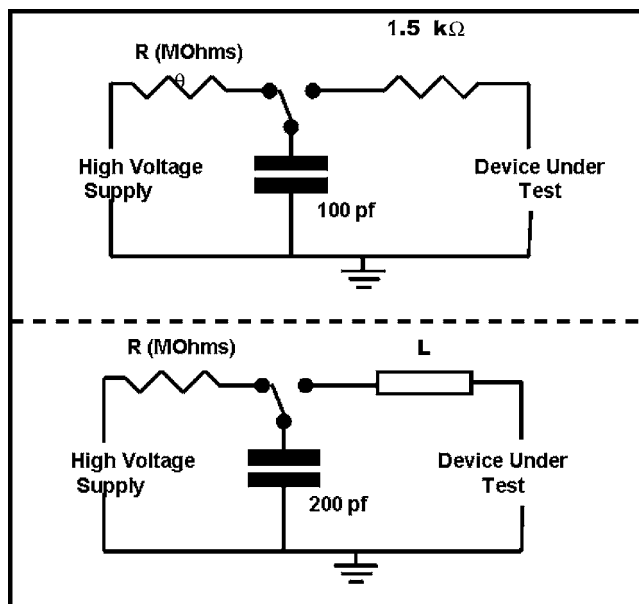


Fig. 3 Circuits used to implement the human body model (top) and the machine model (bottom) of ESD

voltage power supply and, after switching, it is rapidly discharged through $1.5 \text{ k}\Omega$ resistor in series with the device under test (here, the electrothermal actuator). These circuit parameters approximate the capacitance to ground and the resistance of a human. A small parasitic inductance that is part of this circuit is not shown in the figure. The ideal result of this discharge is one of the current wave forms shown in Fig. 4 [9], where the voltage referred to is that across the capacitor at the instant before discharge. The HBM wave form is that of an overdamped RLC (i.e., resistance, inductance, capacitance) circuit and can be fit to (where residual inductance is obviously implied)

$$I(t)_{\text{HBM}} = 0.000668V_0[\exp(-0.00651t) - \exp(-0.250t)] \quad (1)$$

where $I(t)_{\text{HBM}}$ is the time-dependent current in amps; V_0 is the initial voltage across the capacitor; and t is the time in nanoseconds.

Also discussed in Ref. [8] is ESD involving the discharge of a charged machine tool or fixture through the device. A simulation of this event is known as the machine model (MM) and it is

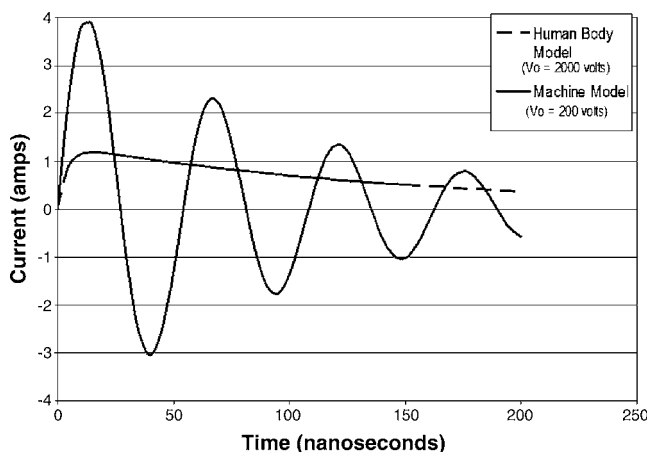


Fig. 4 Typical current wave forms for the human body model and the machine model (from Ref. [9])

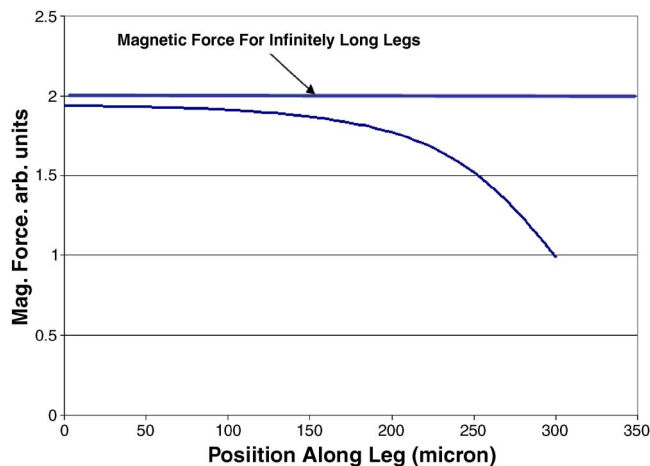


Fig. 5 Magnetic force per unit length per current squared versus position along a leg

implemented with the standard circuit shown in the bottom half of Fig. 3 The MM wave form shown in Fig. 4 [9] is that of an underdamped RLC circuit and can be fit to

$$I(t)_{\text{MM}} = 0.0228V_0[\exp(-0.01t)\sin(0.116t)] \quad (2)$$

Although the ESD testing described in Ref. [4] was conducted to assure optimum fidelity to the ideal MM and HBM wave forms, the actual wave forms were not recorded. Nevertheless, these ideal ESD wave forms will be used in the magnetic analysis that follows.

Response To Magnetic Forces

Formulation of the Calculation. This calculation is based on the configuration shown in Fig. 2, ignoring the slight inclination of the legs to the horizontal that is normally included to produce motion of the shuttle in the desired direction. The current, $I(t)$, entering the bond pad is presumed to divide equally between the two inline pairs of legs and passes through the shuttle unperturbed. Since the center-to-center separation between the pairs of legs ($80 \mu\text{m}$) is much larger than the cross-sectional dimension of the legs, the current is assumed to be filamentary within any leg and within the shuttle. Attention is focused on one leg, such as that between points A and B , which is assumed to be built into the stationary shuttle and the bond pad. Thus, the displacement of the leg and its derivative are zero at the ends. By a straightforward integration of the Biot-Savart law for the magnetic field produced by a current element, and the expression for the force on a current element in a magnetic field, one can show that the magnitude of the magnetic force per unit length, $F_M(x, t)$, at point x and at time t is

$$F_M(x, t) = (I(t))^2(1/40r)[(L-x)/\{r^2 + (L-x)^2\}^{1/2} + (L+t+x)/\{r^2 + (L+t+x)^2\}^{1/2}] \quad (3)$$

where the unit of this "linear force density" is N/m and its direction is as shown in Fig. 2. The parameters r , t , and L are the separation, shuttle width, and the length of the leg. These are $80 \mu\text{m}$, $10 \mu\text{m}$, and $300 \mu\text{m}$, respectively. Plotted in Fig. 5 is the spatial factor in brackets, demonstrating how F_M varies with position along the leg. It will be assumed in the following analysis that r remains essentially constant, despite the vibrations of the beam. If the amplitude of the vibrations were not negligible compared to the separation r , Eq. (3) would not be valid, since it is

only true for parallel current elements.

Based on standard vibration theory, one can show that the response of the leg to a distributed force, such as that in Eq. (3), is as follows [10]

$$Y(x,t) = (1/\rho A) \sum_m (\Phi_m(x)/\omega_m) \int_0^L G(x') \Phi_m(x') dx' \times \int_0^t (I(t'))^2 \sin(\omega_m(t-t')) dt' \quad (4)$$

In the absence of damping, this is an expression for the perpendicular displacement of the leg at point and time t in terms of its normal modes m . $Y(x,t)$ is the displacement; ρ is the mass density of the polysilicon (2330 kg/m^3); A is the cross-sectional area of the leg ($21 \times 10^{-12} \text{ m}^2$); ω_m is the angular frequency of the m th normal mode; $\Phi_m(x)$ is the m th characteristic spatial function (or eigenfunction) of the leg, normalized to 1; and $G(x)$ is everything in Eq. (3) not involving the current. The form of the functions $\Phi_m(x)$ is

$$\Phi_m(x) = N_m [\cos(k_m x) - \cosh(k_m x) + Q_m (\sin(k_m x) - \sinh(k_m x))] \quad (5)$$

where

$$Q_m = (\cos(c_m) - \cosh(c_m)) / (\sinh(c_m) - \sin(c_m)) \quad (6)$$

$$k_m = c_m / L \quad (7)$$

and c_m is determined from the eigenvalue condition that

$$\cos(c_m) \cosh(c_m) = 1 \quad (8)$$

N_m is determined from the normalization condition that

$$\int_0^L (\Phi_m(x))^2 dx = 1 \quad (9)$$

It turns out that all N_m are very close to $1/\sqrt{L}$. In this calculation, only the first five modes have been included in the sum. The values of c_m for $m=1-5$ are: 4.730, 7.853, 10.996, 14.137, and 17.279. A value of zero is also possible, but it results in zero for its associated eigenfunction. It is clear from Eq. (8) that c_m rapidly approaches $(m+1/2)\pi$ with increasing m . In fact, it is within a fraction of a percent of that value even for $m=1$. The various angular frequencies, ω_m , associated with the modes are given by

$$\omega_m = b k_m^2 \quad (10)$$

where

$$b = (EI/\rho A)^{1/2} \quad (11)$$

I is the moment of inertia associated with direction of motion shown in Fig. 2. For the structure under consideration, $I = 22.3 \times 10^{-24} \text{ m}^4$ and $\rho A = 48.9 \times 10^{-9} \text{ kg/m}$. In addition, the Young's modulus, $E = 164 \text{ GPa}$ for polysilicon in MEMS devices [11]. These numbers yield $8.64 \times 10^{-3} \text{ m}^2/\text{s}$ for b . In MHz, the angular frequencies for the first five modes are: 2.15, 5.92, 11.6, 19.2, and 28.7. This means that the fundamental vibrational frequency, $\omega_1/2\pi$, of a leg is about 342 kHz.

In order to determine the tensile stress at any point within the leg, $Y''(x,t)$ must be obtained from Eq. (4) and multiplied by the product of E and a distance, s , from the neutral surface. This surface, which bisects the cross section, is shown in the lower half of Fig. 2. The stress, $\sigma(x,s,t)$, is then

$$\sigma(x,s,t) = EY''(x,t)s \quad (12)$$

The points of maximum stress occur at the anchor points ($x=0, L$) and at the top and bottom surfaces, which are farthest from the neutral surface. Denoting the position of these points relative

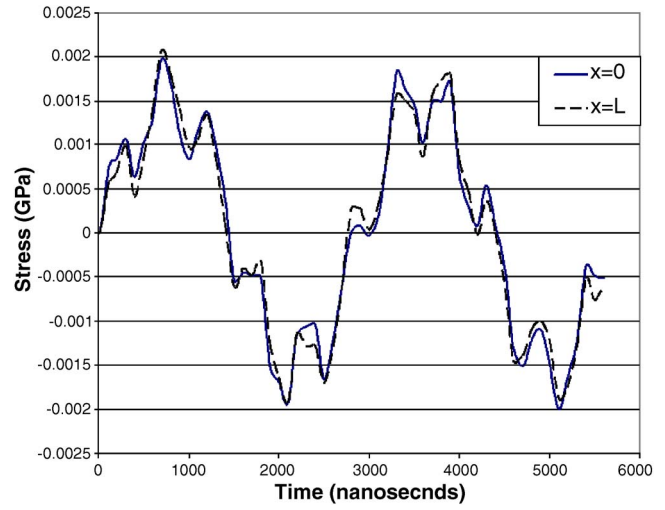


Fig. 6 Time-dependent stress at the two ends of a leg for an initial capacitor voltage of 6500 V (human body model). This curve scales with the square of the initial capacitor voltage.

to the neutral surface as $(w/2)$, the half-width of the leg, the maximum stress is

$$\sigma(0 \text{ or } L, w/2, t) = EY''(0 \text{ or } L, t)(w/2) \text{ (SCF)} \quad (13)$$

where a stress concentration factor (SCF) has been included at the anchor points. In our case, the maximum value of $(w/2)$ is $1.9 \mu\text{m}$ and the chosen SCF is 1.6. The calculated stress will have to be compared at any time to the tensile strength of polysilicon.

Differentiating Eq. (5) twice with respect to x and evaluating it at 0 or L gives us the spatial contribution to $Y''(0 \text{ or } L, w/2, t)$. Due to the symmetry or antisymmetry of the eigenfunctions with respect to the center of the leg, the second derivatives of a given mode are either equal to or are the negative of each other at the two anchor points. This useful characteristic of the eigenfunctions allows us to bypass concerns over numerical accuracy at $x=L$ discussed in Refs. [12,13]. In addition, since the second derivative is proportional to k_m^2 , Eq. (10) tells us that $(\Phi_m''(0)/\omega_m)$ is independent of the index m . Using this fact and Eq. (13), one obtains expressions for the stress at the two anchor points

$$\sigma(0, w/2, t) = 1.36 \times 10^5 \sum_m S_m T_m(t) \text{ GPa} \quad (14)$$

and

$$\sigma(L, w/2, t) = 1.36 \times 10^5 \sum_m (-1)^m S_m T_m(t) \text{ GPa} \quad (15)$$

where S_m and $T_m(t)$ are the modal space and time integrals, respectively, in Eq. (4). The factor $(-1)^m$ appears in Eq. (15) because the sign of the second derivative at $x=L$ alternates with the mode.

Results of the Calculation. Figure 6 shows the result of the stress calculation for the HBM, using Eqs. (1), (14), and (15). As stated earlier, the first five modes were included. In Fig. 6, $V_0 = 6500 \text{ V}$, which is the highest voltage used in the ESD testing, but the stress is still far below the nominal tensile strength of polysilicon, or about 2 GPa. This result leads to the conclusion that magnetic forces alone cannot explain the failure resulting from HBM testing. On the other hand, the similar looking, but much larger, results for the MM shown in Fig. 7 could explain failure, particularly if the maximum tensile strength is somewhat less than half of 2 GPa, due to the factors mentioned earlier. In addition, normal dimensional variations in the legs could be in the right direction to enhance the stress at the points of attachment.

It is of interest to observe the relative contributions from the first five modes and determine whether or not the inclusion of additional modes would have changed the conclusion. Figure 8(a)

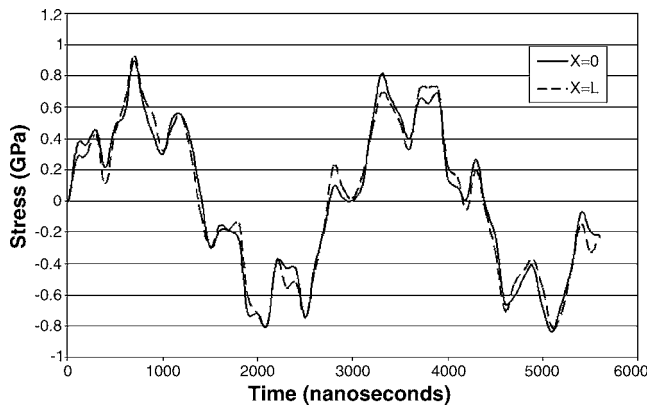


Fig. 7 Time-dependent stress at the two ends of a leg for an initial capacitor voltage of 6500 V (machine model). This curve scales with the square of the initial capacitor voltage.

is a scaled plot of the maximum stress at $x=0$ versus the number of modes, from one to five, for both HBM and MM. That is, aside from a multiplication factor, it is the maximum value achieved in time by the sum in Eq. (14). The character of both sets of points is the same: at least 2/3 of the stress is accounted for by the lowest-order mode alone. Additional modes add to the stress both very gradually and in larger steps, and both series are flattening out, as they must for convergence. Other information is contained in Fig. 8(b), in which the individual (scaled) values, $T_m S_m$, that contribute to the sum for the HBM in Fig. 8(a) are plotted. It is noted how they decay away with mode number in an oscillatory fashion, at least up to $m=5$, thereby explaining the character of Fig. 8(a). The space integrals, S_m , are also plotted and exhibit a similar behavior, except that they lie somewhat above $T_m S_m$, except for $m=1$ because all the points were scaled to be equal for that value. This difference suggests that T_m is decaying with mode number. Figure 8(c) illustrates this decay explicitly for both HBM and MM. With the following, one can see that these trends will continue, even at an accelerated pace. Under the excellent assumptions that $C_m = (m+1/2)\pi$ and all $Q_m \approx -1$, and the somewhat cruder assumption that magnetic force is constant with x , one can prove that S_m varies as $1/(m+1/2)$ for odd m and is zero for even m . This simplified function, appropriately scaled and labeled $\text{Simp.}(S_m)$, is also plotted in Fig. 8(b). It reproduces the essence of S_m quite well. The time integral is expected to decay with modal angular frequency, ω_m , because of the cancellation introduced by the factor, $\sin(\omega_m(t-t'))$, in Eq. (4). It is also expected to do so in an oscillatory manner because of the varying phase of the trigonometric functions appearing in the analytic expression for T_m (not shown). In addition, successive integration by parts of T_m results in an asymptotic series for it whose lead term is $(1/\omega_m)$. Thus, aside from any oscillations, by Eq. (10), T_m should decay with mode number as $1/(m+1/2)^2$ for large m . This trend is plotted along with T_m in Fig. 8(c) for the first five modes. It is clear, despite the decay in T_m , that “asymptopia” has not yet been reached up to $m=5$. Nevertheless, for large m , $T_m S_m$ is expected to vary as $1/(m+1/2)^3$. It is clear from the foregoing discussion that adding more modes to the stress calculation will not change the basic conclusion stated in the first paragraph of this section.

Thermal Expansion and Stress Amplification

We have seen that the magnetic attraction between the two pairs of legs provides a possible explanation for the failure of the electrothermal actuators under MM ESD testing and that it cannot under HBM testing. However, this explanation does not include the thermal expansion of the legs that takes place during this process. Alone or in conjunction with the magnetic attraction, an

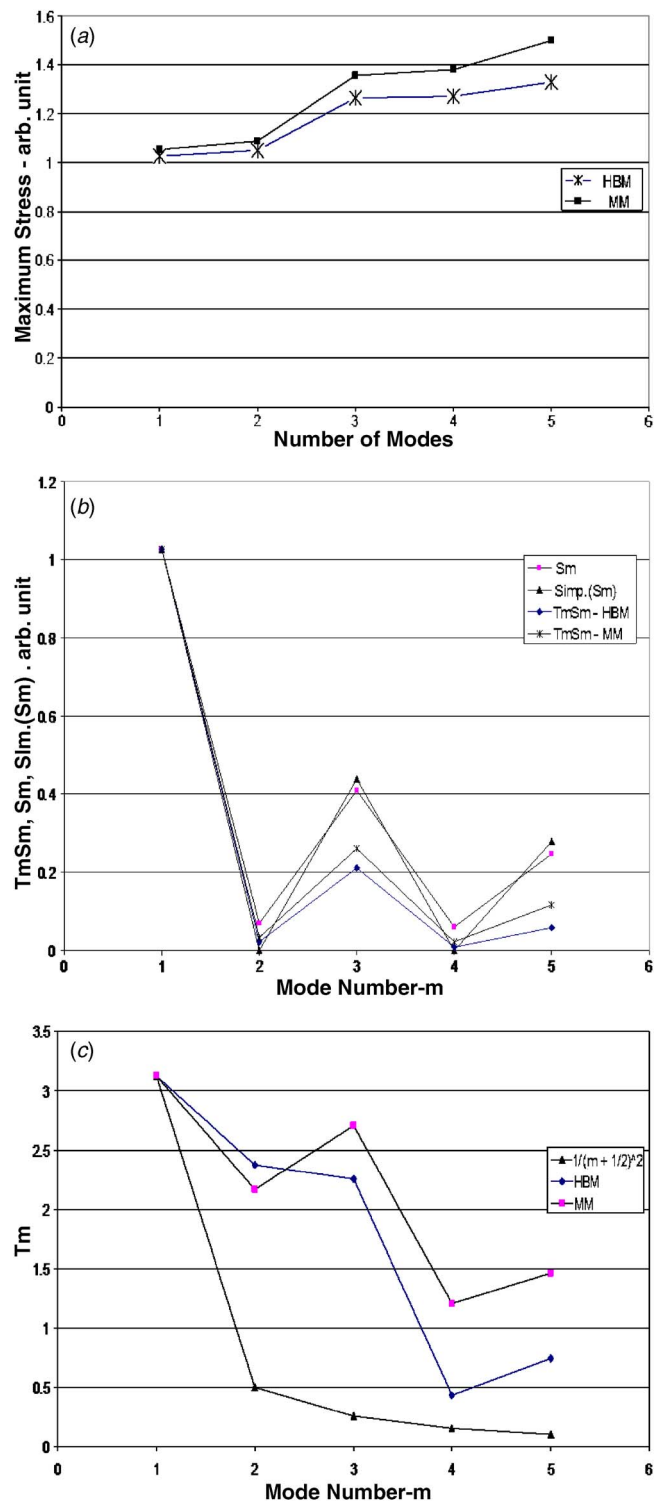


Fig. 8 (a) Maximum stress achieved at $x=0$ versus the number of modes included in the calculation, for both HBM and MM; (b) the product of integrals, $T_m S_m$, used to calculate the stress in Fig. 8(a) for the HBM versus mode number. Similarly for S_m alone and a simplified calculation of S_m ($\text{Simp.}(S_m)$) illustrating the basic dependence of S_m on mode number; and (c) time integral, T_m , versus mode number for HBM and MM. Also, the asymptotic behavior of T_m versus mode number ($1/(m+1/2)^2$).

attempt will be made to demonstrate that it does provide a feasible explanation for failure, under the assumption to be justified later, that negligible motion of the shuttle occurs during this rapid excursion in temperature.

As a result of resistive heating by the electric current, any leg tends to expand. If it is constrained longitudinally by the shuttle and bond pad, a compressive load develops within the leg. If this load reaches a critical value, the leg will buckle even in the absence of any lateral force, such as the magnetic attraction discussed above. If the buckling condition is exceeded in the absence of a lateral force, the initially straight leg can assume a sinusoidal shape of finite amplitude, which can be determined from the thermally induced change in length. Both the spatial frequency of the sinusoid and the critical value depend on the boundary conditions. Since the leg is assumed to be built in, the boundary conditions are that its displacement and the derivative of its displacement are zero at either end. Since the conditions are quite different, mechanically, the discussion will be divided into pre- and post-buckling, including the effects of a lateral load. However, the critical compressive load and associated thermal issues will be considered first.

The Critical Compressive Load and Associated Thermal Issues. As stated in Ref. [14], the critical compressive load, P_{crit} , for a beam built in at both ends is given by the well-known Euler formula

$$P_{crit} = 4\pi^2 EI/L^2 \quad (16)$$

For a leg of the thermal actuator, this value is 1.6×10^{-3} N. The compressive load, P_0 , generated by constrained thermal expansion is, for small length changes, given by

$$P_0 = EA\epsilon_t = EA\Delta L/L = EA \int_{300}^T \alpha(T')dT' \quad (17)$$

where $\alpha(T)$ is the temperature-dependent linear thermal expansion coefficient. Substituting Eq. (16) into Eq. (17), one deduces that criticality is reached when

$$\int_{300}^T \alpha(T')dT' = 4\pi^2 I/AL^2 = 0.000466 \quad (18)$$

In other words, if the beam expands in length by about 42/3 parts per 10,000, criticality is reached. It is important to note that this condition is for the nominal values of the various dimensions in Fig. 2, several of which are raised to a power greater than one in Eq. (18). Thus, the inevitable manufacturing variations in those dimensions could have a significant impact on failure, if criticality is closely approached.

In order to determine whether criticality is reached or is closely approached, assuming the nominal dimensions, the integral will be performed for various elevated temperatures. A fit to the data on single crystal silicon yields the following for $\alpha(T)$ [15]

$$\alpha(T) = 10^{-6}[3.725 - 7.723 \exp(-0.00588T) + 0.0005548T]K^{-1} \quad (19)$$

When the integral is performed to the melting temperature of 1688 K, the results displayed in Fig. 9 are obtained. The calculation indicates that criticality is reached at about 452 K (179°C). But, how hot is the leg likely to get during ESD testing?

The maximum achievable temperature can be obtained by assuming that all of the energy initially stored in the capacitor, $CV_0^2/2$, goes into resistive heating. In the case of the HBM, about 17% of it is available for heating the legs of the thermal actuator, since, for this circuit, the capacitor discharges through a 1500 Ω resistor in series with the electrothermal actuator [4]. Its total resistance is about 300 Ω . Only 25% of the remaining energy goes into heating any one of the four legs. In the case of the MM, there is essentially no external series resistor; thus, 25% of the total is available to heat any of the four legs. As for radiation, a simple estimate based on the surface area of the legs and an emissivity of 1 indicates that a negligible fraction of the total available electrical energy is converted into this form during the microsecond, or

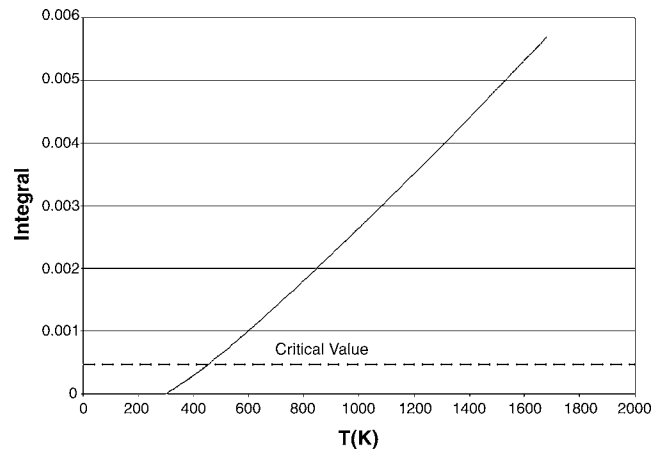


Fig. 9 Integral of the linear thermal expansion coefficient from 300 K to some elevated temperature T

so, that is required for discharge. Thermal conduction during this brief period is also assumed to be negligible because it is a relatively slow process.

Thus, an upper limit to the temperature achievable during ESD testing is obtained from

$$\Gamma CV_0^2/2 = m \int_{300}^T C_p(T')dT' \quad (20)$$

where $\Gamma=0.0425$ in the case of HBM and $\Gamma=0.25$ in the case of MM. The symbols m and $C_p(T)$ are the mass of a leg and the heat capacity per unit mass of polysilicon. Between room temperature and its melting point, the heat capacity of silicon, in J/(kg K), is the following function of temperature [16]

$$C_p(T) = 843.6 + 0.1176T - 151.4 \times 10^{-5}T^{-2} \quad (21)$$

The result is Fig. 10, which is a plot of the calculated temperature rise versus the initial capacitor voltage for the two values of capacitance considered here. Considering the voltages used in the ESD testing [4], it appears inevitable that criticality was reached for some of them.

One concern regarding the calculation of the maximum temperature achievable is that at the higher voltages, particularly for the MM, it predicts that the melt can be exceeded, even though no evidence of melting was associated with ESD-induced failures.

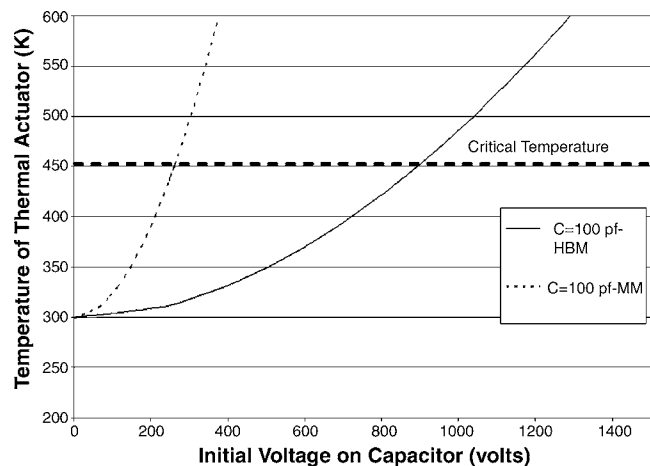


Fig. 10 Maximum achievable temperature of the thermal actuator during ESD testing versus initial voltage on the capacitor

This problem can be ameliorated by noting that the latent heat of fusion for silicon is high enough that ~ 1.4 times the energy required to elevate the silicon to the melting temperature is required to melt it. The latent heat of fusion is ~ 1.81 MJ/kg [17], while the energy required for elevation to the melt, obtained by integrating Eq. (21), ~ 1.29 MJ/kg. Thus, achieving its melting temperature and actually melting the silicon are quite different. It may be asked if a dissipative mechanism such as thermal conduction may have played a larger role in reducing the maximum achievable temperature than originally thought. This is unlikely, in view of the analysis of the transient temperature response of these actuators by Lott et al. [18]. They have demonstrated that the relevant thermal time constant varies from a significant fraction of a millisecond for devices in air to about 30 ms for devices in vacuum. On the other hand, the possible existence of resistance in the circuits not shown in Fig. 3 would have reduced the fraction of the original energy available to heat the silicon. Furthermore, if melting did occur, the leg may have been in the liquid state for a short enough time that shape changes were negligible. Resolving this problem is impossible here, particularly because the current wave forms were not recorded, but suffice it to say that when post-buckling behavior is examined, temperatures up to the melt should be considered.

Prebuckling. As the critical value of compression is approached in the presence of even a weak lateral force, small-deflection theory predicts that the amplitude of the deflected beam increases without limit [12]. Although small-deflection theory cannot be considered accurate at large amplitudes, this prediction might still lead one to expect that all of the legs will break if the ESD heating brings them even close to the critical temperature of 452 K. However, many did not. The reason for this discrepancy is that the first statement of this paragraph presupposes that the compression is constant, regardless of the amplitude. This condition is unattainable here; since the separation between the end points is fixed, the compression is reduced when the beam is deflected because it elongates in the process. Thus, as is now demonstrated, even at the critical temperature, a self-consistent solution can be found for which the deflection is finite.

In this discussion, small-deflection theory will be assumed and, for simplicity, the lateral force per unit length, q , is spatially constant. In view of Fig. 5, a uniform force is a reasonable assumption. Regarding the accuracy of small-deflection theory, a simple calculation assuming a uniform lateral force and zero compressive force indicates that the nominal tensile stress of 2 GPa is reached at the anchor points while small-deflection theory is still quite accurate (i.e., all $y'^2(x) \ll 1$, where $y(x)$ is the deflection curve. In fact, its maximum value in this example ≈ 0.013). In the presence of both lateral and compressive forces, the differential equation for the deflection of a long, slender beam, $y(x)$, is [12]

$$EI d^4y/dx^4 + P_c d^2y/dx^2 = q \quad (22)$$

For the moment, the compressive force, P_c , is considered constant and equal to P_0 , the compression at zero deflection. Placing the origin of the coordinate system at its center, in anticipation of symmetry, the beam extends over the range: $-L/2 \leq x \leq L/2$. For any of the legs, $L=300 \mu\text{m}$. Since the beam is fixed at either end, $y(x)=y'(x)=0$ at $x=\pm L/2$. Then the solution to Eq. (22) is

$$y(x) = (qL^2/4P_0U_0 \sin U_0)(\cos 2U_0x/L - \cos U_0) + (q/2P_0)(x^2 - (L/2)^2) \quad (23)$$

where $U_0=(L/2)\sqrt{(P_0/EI)}$. Also defined is $U_c=(L/2)\sqrt{(P_c/EI)}$, for use later. Of particular interest is the second derivative at the end points, $y''(\pm L/2)$, because at these points the tensile stress induced by flexure is proportional to it. The total tensile stress is obtained by subtracting P_0/A , the overall compressive stress, from this quantity. The maximum value of P_0/A occurs at the critical compression, where it is 0.076 GPa, or 3.8% of the nomi-

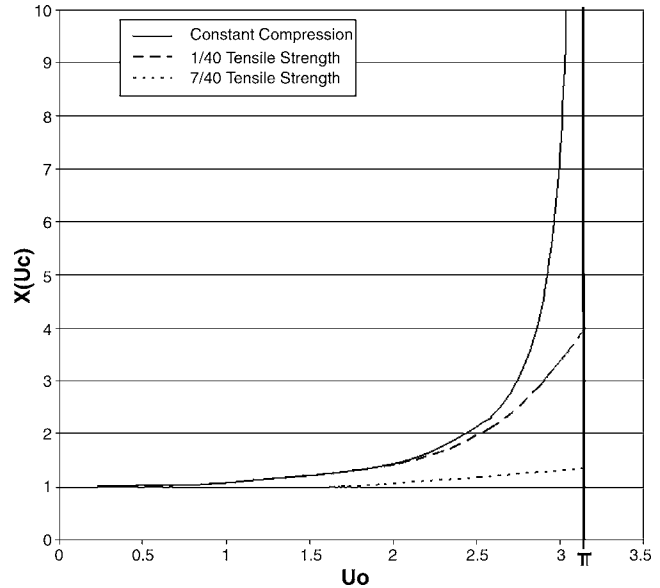


Fig. 11 Amplification factor, $X(U_c)$ versus U_0 for constant compression, where $P_c=P_0$, and for values of q that would produce flexural stress at the end points of 7/40 and 1/40 the nominal stress of silicon, according to simple beam theory and $P_c=0$. For all $q>0$, $P_c<P_0$. $U_{0,c}=(L/2)\sqrt{(P_{0,c}/EI)}$.

nal tensile strength. From Eq. (23), the second derivative at the end points is

$$y''(\pm L/2) = (qL^2/12EI)X(U_0) \quad (24)$$

where

$$X(U_0) = (3/U_0^2)(1 - U_0/\tan U_0) \quad (25)$$

The expression multiplying $X(U_0)$ in Eq. (24) is the second derivative at zero compression. Thus, the second derivative at finite compression is that at zero compression amplified by $X(U_0)$. This function is infinite at criticality, where $U_0=\pi$, and it serves as the basis for the statement made at the beginning of this section. It is plotted in Fig. 11 as the curve labeled "constant compression." However, as stated above, the compression is not constant with deflection. The actual compression, P_c , at a finite deflection is given by

$$P_c = P_0 - EA\Delta/L \quad (26)$$

where Δ is the change in length caused by the deflection. It is given by

$$\Delta = \int_{-L/2}^{L/2} \sqrt{1 + y'^2(x)} dx - L \approx (1/2) \int_{-L/2}^{L/2} y'^2(x) dx \quad (27)$$

with the approximation having been made in the spirit of small-deflection theory. Using Eqs. (23) and (27), the following is obtained for P_c

$$P_c = (2/L)^2 EI U_c^2 = P_0 - (EAL^6/256)(q/EI)^2 [(1/U_c^4)(2/3 + \csc^2 U_c + (3 \cot U_c)/U_c - (2/U_c^2))] \quad (28)$$

The solution of this equation yields a self-consistent value of the compression in the deflected beam as a function of P_0 and q , along with the other parameters that characterize the geometry and rigidity of the beam. The function in brackets is relatively flat around $U_c=0$ and it $\rightarrow\infty$, as $U_c\rightarrow\pi$. It is similar in appearance to $X(U_0)$. The result of graphically solving for P_c as a function of P_0 is shown in Fig. 11 for two values of q , conveniently chosen to be illustrative of the point being made. Including the stress concentration factor, they produce a flexural stress at the attachments

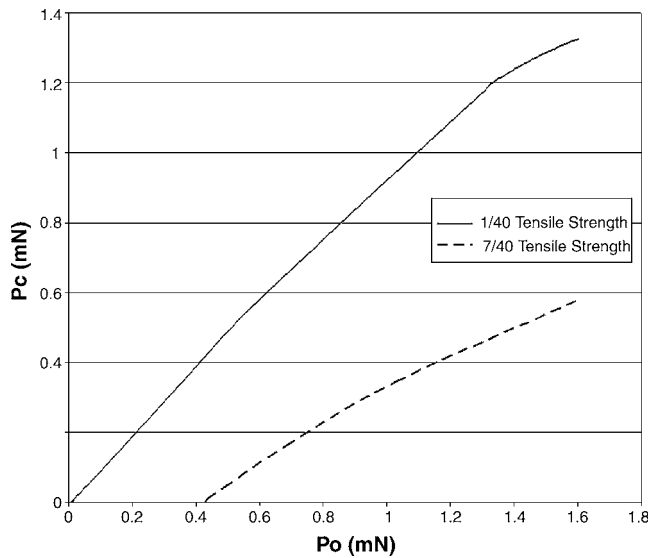


Fig. 12 P_c versus P_0 for the two values of q in Fig. 11

points of $\sim 1/40$ and $\sim 7/40$ of the nominal tensile strength of the silicon, assuming that $P_c = 0$ in Eq. (22). (The dynamic equivalent was used to calculate the magnetic response earlier.) The curves demonstrate that the actual compression at a given deflection is, as expected, less than P_0 . One can also see from the dashed curve that the lateral force can drive an originally compressed beam into tension, where $P_c < 0$, although that region is not shown. The curves in Fig. 12 for these two cases illustrate the dramatic reduction in the amplification factor near buckling. This dramatic reduction suggests that thermal expansion below the critical temperature will not greatly enhance the magnetic interaction as a probable cause of fracture.

Postbuckling. The two curves in Fig. 12 could be smoothly continued beyond the critical temperature, since the actual compression of the beam, as determined by Eq. (28), never reaches the critical compression. However, this treatment predicts a deflection of the beam that is zero when the lateral force is removed. This is true below the critical temperature, but beyond it the beam will begin to buckle and acquire a deflection even in the absence of a lateral force. (Zero deflection is mathematical allowable, but is physically unstable.) Thus, a discontinuity of sorts does occur at the critical temperature. It is presumed, therefore, that the effects of the lateral force can be superimposed on those caused by buckling. Within small-deflection theory [12], a buckled beam, built in at both ends, has the following deflection for $0 \leq x \leq L$

$$y_b(x) = A(1 - \cos 2\pi x/L) \quad (29)$$

where the amplitude A is indeterminate for strictly mechanically induced buckling. In addition, the compressive stress within the buckled beam is pinned at the critical stress. In the case of thermally induced buckling, the amplitude can be determined from the length change beyond the critical temperature, T_c . In other words

$$\Delta = L \int_{T_c}^T \alpha(T') dT' \approx (1/2) \int_0^L y_b'^2(x) dx = (1/L)(\pi A)^2 \quad (30)$$

Thus,

$$A \approx (L/\pi) \sqrt{\int_{T_c}^T \alpha(T') dT'} \quad (31)$$

The maximum value of A occurs at 1688 K, the melting temperature. At that temperature, $A \approx 6.9 \mu\text{m}$ for $L = 300 \mu\text{m}$, which results in a maximum value of $y_b'^2(x)$ of 0.021. Its small size com-

pared to unity justifies the use of small-deflection theory, post-buckling. Including the stress concentration factor of 1.6 and the overall compression of 0.076 GPa, the corresponding stress at the anchor points is ≈ 1.43 GPa, which is well within the range of the estimated tensile strength of the silicon. Dimensional variations among the actuators, some by accident and others by design, will affect any numerical calculation of the critical temperature and the stress. The additional magnetic contribution, particularly in the case of the MM, could add nontrivially to any such result.

It was noted in Ref. [4] that devices subject to MM testing tended to fail at lower voltages than those subjected to HBM testing. The explanation that the higher current produced by MM testing resulted in larger buckling amplitudes and larger magnetic interactions can now be reasonably offered.

The discussion pertaining to buckling is valid for a static leg, although the condition of ESD testing is dynamic. For example, the fundamental, or dominant, frequency of vibration of the structure was stated earlier to be 342 KHz. Thus, its period is about $3 \mu\text{s}$. However, this seemingly small time should be viewed relative to the transit time of a crack through the width of the leg shown in the bottom half of Fig. 2. As discussed in Ref. [19], the terminal speed of crack propagation $\sim 0.38 \sqrt{E/\rho}$. For the values of silicon stated earlier, this speed $\sim 3.2 \times 10^5$ cm/s. Given that the largest transverse dimension of the leg is $3.8 \mu\text{m}$, the transit time straight through ~ 1.2 ns. Since this number is 2500 times smaller than the fundamental period of vibration, the leg can be considered essentially fixed during this time, and the presence of a few higher-order modes of lower amplitude has no material effect on this conclusion. Thus, the static treatment should yield insight into failure.

Residual Concerns

Eddy Currents. When the current pulse is applied to the legs, eddy currents are induced in the polysilicon substrate a few microns below the bottom of the legs, but they have been ignored in the analysis up to this point. Since they are of obvious concern, certain general comments concerning these currents will be presented here, although their detailed study is beyond the scope of this paper.

An expression, in the form of a rather complicated integral, for the current density induced within a large conducting plate of finite thickness below a long filamentary current has been derived elsewhere [20]. The filament is presumed to be parallel to the plate, as in Fig. 13(a), and the current oscillates at a single frequency. For a current pulse, such as that in Eqs. (1) and (2), the expression in Ref. [20] can be generalized by integrating it over the frequency content of the pulse. For a continuous current distribution within an actual wire, the same expression can be generalized by integrating it over the cross section of the wire, such as the leg in the bottom of Fig. 2. However, this additional analysis is beyond the scope of this paper. In what follows, therefore, the results for a filamentary current oscillating at a single frequency will be considered sufficient.

The induced current density within the plate is parallel to the filament and is obviously symmetric about a plane perpendicular to the plate and passing through the filament. Thus, the force that these eddy currents exert on the filament is perpendicular to the plane of the plate and, therefore, perpendicular to the in-plane motion that has been assumed throughout this paper. Thus, for deflections which are not large, the magnetic force exerted on a given leg caused by a neighboring leg can be treated independently of the force exerted on that same leg by its own eddy currents. However, a little thought will show that the tensile stress near an anchor point caused by the in-plane motion can be enhanced or diminished by this perpendicular motion.

It is now argued that the effect of the eddy currents is small compared to that of the other leg, although it is done only qualitatively. First, consider the distribution of the eddy current density

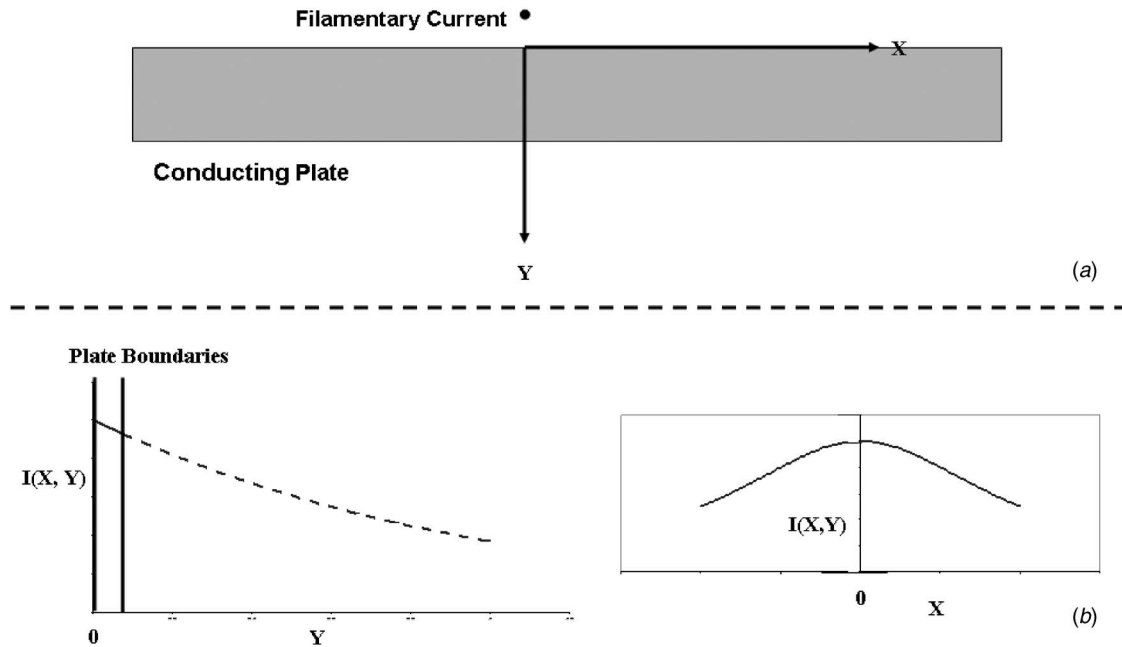


Fig. 13 (a) Geometry of eddy current discussion; and (b) Eddy current versus vertical and horizontal distances from the filament

within the plate. Its variation throughout the thickness of the plate is largely determined by the skin depth, which is the $(1/e)$ point of an exponentially decaying electromagnetic field of a given frequency within a conductor. Now, the thickness of the polysilicon substrate is about 5/8 mm. Given the dominant frequencies in the ESD pulses in Eqs. (1) and (2) and the resistivity of the phosphorus-doped, polysilicon substrate ($\sim 8 \Omega \text{ cm}$), one can show that the skin depth is much greater than this thickness. Thus, the induced current is essentially constant throughout the thickness of the substrate, which is more than six times the separation between the legs. This variation is shown qualitatively in Fig. 13(b). In addition, the lateral (i.e., parallel to the plate surface) distribution of current density is very wide, as suggested by certain examples in Ref. [20]. As shown in Fig. 13(b), it has a peak below the filament, decays slowly to zero with distance on either side, and likely extends beyond the second leg. The point is that the wide distribution of eddy currents implies that much of the contribution to the force on the leg is from current elements much farther away than the second leg. Thus, those contributions are expected to be relatively weak. In addition, it is only their vertical component that is of importance, as implied earlier. Furthermore, the phase of the eddy currents is spatially varying; therefore, some cancellation will occur from element to element. Finally, the beams were designed so as to resist out-of-plane motion: the vertical cross-sectional moment of inertia of the beam in Fig. 2 is 4.3 times its horizontal moment of inertia.

Motion of the Shuttle During the ESD Pulse. The analysis has presupposed essentially no upward motion of the shuttle during the ESD pulse, as it would in normal operation. This is important because substantial motion in that direction would relieve the tendency to buckle. In order to justify this assumption, the resonant frequencies of the system in Fig. 2 consisting of $600 \mu\text{m}$ legs and a point mass M in the center equal to that of the shuttle have been calculated. The legs are assumed to be built in at either end. Only the spatially symmetric modes of the “shuttleless” system, such as the fundamental, are affected by the shuttle; their frequency is lowered, but they remain symmetric. The antisymmetric modes have a node at the center point, where the additional inertia has no influence on their behavior. Utilizing symmetry, the boundary conditions on the displacement, $Y(x, t)$, are

$$Y(x, t) = Y'(x, t) = 0;$$

$$x = 0: 2EIY'''(x, t) = M d^2 Y(x, t)/dt^2$$

$$= -M\omega_m^2 Y(x, t); \quad x = L/2 \quad (32)$$

In these equations, the origin is now at the left bond pad in Fig. 2 and $x=L=600 \mu\text{m}$ is at the right bond pad. The first expresses the fixity condition at the left bond pad; the second is Newton's second law as applied to the point mass M under the combined action of the shear forces on either side of it. The mass and beam are assumed to be oscillating with modal angular frequency, ω_m . The resulting eigenvalue equation leading to ω_m is now

$$[\cos(c_m)\sinh(c_m) + \cosh(c_m)\sin(c_m/2)]/[c_m(1 - \cos(c_m)\cosh(c_m))] = R \quad (33)$$

where, in contrast to Eq. (7), $c_m = k_m L/2$, but Eq. (10) is unchanged. The quantity R is the ratio of the point mass to that of the entire beam. When $R=0$, one can readily show, as one must, that Eq. (33) can be manipulated into the form of Eq. (8). Using the typical value of R , 0.23, it is found that the fundamental vibrational frequency, $\omega_1/2\pi$, of the “shuttled” system is 67.6 kHz, and the first higher-order frequency of the symmetric modes is 404 kHz. All but the fundamental mode have vibrational nodes. Considering what the system would look like with a displaced shuttle, it is reasonable to assume that the fundamental mode would dominate, as it did in the magnetic analysis. Then the motion of the shuttle can be considered extremely sluggish on the time scale of the ESD pulses ($\sim 200 \text{ ns}$).

Summary

Two mechanisms have been postulated to explain the failure of electrothermal actuators during ESD testing: magnetic attraction between pairs of legs and resistive heating of the legs. An analysis of the magnetic attraction reveals that it can cause failure only near the highest currents generated by the tests. A modest amplification of the magnetic effects is achieved by thermally induced compression in the legs prior to buckling. Failure can readily be explained by resistive heating of the legs beyond the buckling temperature, the effects of which are added to the magnetic inter-

action. The analysis presented here presupposes essentially no bulk motion of the shuttle during the brief period of the ESD current and the negligible effect of eddy currents induced in the substrate during the same period. Arguments have been presented justifying these assumptions.

Acknowledgment

The author would like to thank Jeremy Walraven of Sandia National Laboratories for initially involving him in the failure analysis of these electrothermal actuators and for reviewing the manuscript. He also thanks Richard Plass and Michael Baker of Sandia National Laboratories for reviewing the manuscript. The comments of all three have, in the author's view, resulted in a significantly improved manuscript. He is also grateful to Michael Baker and Michael Shaw of Sandia National Laboratories for providing him with certain materials properties of polysilicon and a representative design of the electrothermal actuators. Sandia is multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under the National Nuclear Security Administration, Contract No. DE-AC04-94AL85000.

References

- [1] Que, L., Park, J. S., and Gianchandani, Y. B., 2001, "Bent-Beam Electrothermal Actuators—Part I: Single Beam and Cascaded Devices," *J. Microelectromech. Syst.*, **10**(2), pp. 247–254.
- [2] Park, J. S., Chu, L. L., Oliver, A. D., and Gianchandani, Y. B., 2001, "Bent-beam Electrothermal Actuators—Part II: Linear and Rotary Microengines," *J. Microelectromech. Syst.*, **10**(2), pp. 255–262.
- [3] Walraven, J. A., Soden, J. M., Tanner, D. M., Tangyung, P., Cole, E. I., Jr., Anderson, R. E., and Irwin, L. W., 2000, "Electrostatic Discharge/Electrical Overstress Susceptibility in MEMS: A New Failure Mode," *Proc. SPIE*, **4180**, pp. 30–39.
- [4] Walraven, J. A., Plass, R., Baker, M. S., and Shaw, M. J., 2004, "Failure Analysis of Electrothermal Actuators Subjected to Electrical Overstress (EOS) and Electrostatic Discharge," *Proceedings of the 30th International Symposium for Testing and Failure Analysis (ISTFA)*, Worcester, MA, Nov. 14–18.
- [5] Baker, M. S., Plass, R. A., Headley, T. J., and Walraven, J. A., 2004, "Final Report: Compliant Thermo-Mechanical MEMS, Actuators LDRD #52553," Sandia National Laboratories, Sandia, NM, Report No. SAND2004-6635.
- [6] Pilkey, W. D., 1997, *Peterson's Stress Concentration Factors*, Wiley, New York, Chap. 3.
- [7] Dugger, M. T., Boyce, B. L., Burchheit, T. E., and Prasad, S. V., 2004, "Mechanics and Tribology of MEMS Materials," Sandia National Laboratories, Sandia, NM, Report No. SAND2004-1319.
- [8] ESD Association, 2001, <http://www.esda.org/basics/part5.cfm>
- [9] National Semiconductor EOS/ESD Information Website, 2007, <http://www.national.com/appinfo/eosed>
- [10] Timoshenko, S. P., Young, D. H., and Weaver, W., 1974, *Vibration Problems In Engineering*, 4th ed., Wiley, New York, Chap. 5.
- [11] Jensen, B. D., et al., 2001, "Interferometry of Actuated Microcantilevers to Determine Material Properties and Test Structure Nonidealities in MEMS," *J. Microelectromech. Syst.*, **10**(3), pp. 336–346.
- [12] Chang, T., and Craig, R. R., 1969, "Normal Modes of Uniform Beams," *J. Eng. Mech.*, **95**(EM4), pp. 1027–1031.
- [13] Tang, Y., 2003, "Numerical Evaluation of Uniform Beam Modes," *J. Eng. Mech.*, Vol. **129**(12), pp. 1475–1477.
- [14] Timoshenko, S. P., and Gere, J. M., 1961, *Theory of Elastic Stability*, 2nd ed., McGraw-Hill, New York, Chaps. 1 and 2.
- [15] Okada, Y., and Tokumaru, Y., 1984, "Precise Determination of Lattice Parameter and Thermal Expansion Coefficient of Silicon Between 300 and 1500 K," *J. Appl. Phys.*, **56**(2), pp. 314–320.
- [16] Gurvich, L. V., Vetyts, I. V., and Alcock, C. B., eds., 1990, *Thermodynamic Properties of Inorganic Substances*, Hemisphere, New York, Vol. 2.
- [17] Brandes, E. A., Brook, G. B., and Smithells, C. J., 1998, *Smithells Metals Reference Book*, 7th ed., Butterworth-Heinemann, Boston, MA.
- [18] Lott, C. D., McLain, T. W., Harb, J. N., and Howell, L. L., 2002, "Modeling the Thermal Behavior of a Surface-Micromachine Linear-Displacement Thermomechanical Microactuator," *Sens. Actuators, A*, **101**, pp. 239–250.
- [19] Anderson, T. L., 1995, *Fracture Mechanics: Fundamentals and Applications*, 2nd ed., CRC, Boca Raton, FL, Chap. 4.
- [20] Tegopoulos, J. A., and Kiezis, E. E., 1985, *Eddy Currents in Linear Conducting Media*, Elsevier, Amsterdam, The Netherlands.

Shi Chen

Xi'an Jiaotong University,
Xi'an 710049, China;
School of Physics and Information Technology,
Shaanxi Normal University,
Xi'an 710062, China
e-mail: chenshi@snnu.edu.cn

Tiantong Tang

Zhaohong Wang

Xi'an Jiaotong University,
Xi'an 710049, China

A One-Dimensional Theory for the Analysis of Strip Surface Acoustic Waveguides

In this paper, an effective and new one-dimensional theory for the analysis of strip surface acoustic waveguides is presented. Compared with the known methods, a prominent advantage of this method proposed in this paper is that all the field variables can be calculated. [DOI: 10.1115/1.2723815]

Keywords: strip surface acoustic waveguides, eigenmodes, Mindlin plate theory, dispersion curves

1 Introduction

Because of their excellent properties to confine acoustic waves, strip surface acoustic waveguides are important components in some acoustic devices, such as acousto-optic (AO) polarization converters [1] and AO tunable filters [2] in wavelength division multiplexing optical communication systems.

It is very important for the design of the acoustic devices to accurately model these waveguides. Some approximate methods have been established in previously published papers. For example, a scalar potential method (SPM) [3] was proposed. Drawbacks of this method are twofold: one is that only scalar potential is considered in the calculation of the eigenmodes, and the vector potential is ignored, thus it is impossible to obtain the field distribution; the other is that it is only valid for isotropic materials. Lately an effective index method (EIM) [4] can be regarded as a generalization of the SPM for anisotropic materials, and in this method only a single component of the elastic displacements is considered in the calculation. Although this method is simple, it is subject to some drawbacks, such as its low precision and the impossibility to calculate the field distribution. There is also another theory [5], which is completely different from the SPM and the EIM, and is based on the microwave equivalent network formulation (MENF) [5]. A drawback of the MENF is that the field distribution cannot be obtained by this method.

In this paper, a new and effective one-dimensional theory for the analysis of strip surface acoustic waveguides is presented. It is based on a two-dimensional theory [6] for the surface acoustic wave (SAW) propagation in a semi-infinite solid, where from the general Mindlin plate theory [7] Wang and his co-workers derived the two-dimensional theory for SAWs near the surface of a semi-infinite solid without the discontinuities of mechanical boundary conditions. Compared with the work of Wang et al., our contributions are as follows: (1) The real number variation formulation in the paper [6] is replaced by the complex variation formulation in our paper. The advantage of the complex formulation is obvious, which avoids the complex expression of the eigenmodes; (2) The discontinuities of the mechanical boundary conditions on the surface are considered fully, which enables the two-dimensional SAW theory to be used for the design of practical devices; and (3) The mass loaded effect is also investigated generally. Compared with the SPM, EIM, and MENF, a prominent advantage of the method is: the field distribution of the eigenmodes can be calculated by our proposed method, and this is very important in the design of some acoustic devices.

2 Theory Formulation

A section of a strip surface acoustic waveguide is shown in Fig. 1, where the anisotropic solid occupies the semi-infinite space $x_1 \leq 0$, the thickness, and the width of the film are h and $2w$, respectively; and the other region is free space.

2.1 A One-Dimensional Theory for SAWs on a Semi-Infinite Anisotropic Solid With a Uniform Surface. The three-dimensional field equations in anisotropic solid materials are

$$T_{ij,j} - \rho \ddot{u}_i = 0 \quad (1)$$

The stress tensor T_{ij} can be expressed as

$$T_{ij} = c_{ijkl} u_{k,l} \quad (2)$$

Here u_j and c_{ijkl} denote the elastic displacement vector and the elastic tensor, respectively; and ρ is the mass density.

Provided that the anisotropic solid is infinite in the x_1 and x_2 directions, finite in the x_3 direction, and assuming that the SAW propagates in the x_2 direction, all the field variables have a common factor $\exp(Kx_2)$, which will be omitted hereafter. The elastic displacements can be expressed as

$$u_i(x_1, x_3) = \sum_{m=1}^3 u_i^{(m)}(x_3) \exp(\beta_m x_1) \quad (3)$$

where K is a pure imaginary wave vector in the x_2 direction; and β_m ($m=1, 2, 3$) are complex numbers with positive real parts, which are from the eigensolutions for a infinite anisotropic solid and express approximately the attenuation behavior of the SAWs in the depth direction in the finite structure. In other words, $(i\beta_m, iK, 0)$ is the wave vector for the infinite case, and the waves are inhomogeneous in the x_1 direction. Because only monochromatic waves are studied in the paper, all field variables have a common factor $\exp(i\omega t)$, which will be omitted hereafter. Here ω denotes the angular frequency of the waves.

According to the complex variation principle [8], the following variation equation is listed

$$\iiint_{\Omega} [(T_{ij,j} + \rho \omega^2 u_i) \delta u_i^*] dx_1 dx_2 dx_3 = 0 \quad (4)$$

where Ω denotes the space occupied by the anisotropic solid, which is finite in the x_3 direction.

The integral regarding x_2 is a constant, and it can be dropped from Eq. (4)

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received August 24, 2006; final manuscript received February 1, 2007. Review conducted by Sridhar Krishnaswamy.

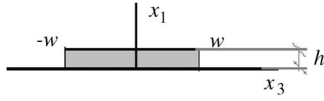


Fig. 1 A section of a strip surface acoustic waveguide

$$\int dx_3 \int_{-\infty}^0 dx_1 (T_{ij',j'} + \rho \omega^2 u_i + T_{i1,1}) \delta u_i^* = 0 \quad (5)$$

here $j' = 2, 3$ and $\partial_2 \equiv K$. The following equation is obvious

$$\begin{aligned} \int_{-\infty}^0 dx_1 T_{i1,1} \delta u_i^* &= T_{i1} \delta u_i^*|_{-\infty}^0 - \int_{-\infty}^0 dx_1 T_{i1} (\partial_1 \delta u_i^*) \\ &= T_{i1}(0) \delta u_i^*(0) - \int_{-\infty}^0 dx_1 T_{i1} (\partial_1 \delta u_i^*) \end{aligned} \quad (6)$$

where boundary conditions $u_i(x_1 = -\infty) = 0$ are used.

Submitting expression (6) into Eq. (5) yields

$$\begin{aligned} \int dx_3 \int_{-\infty}^0 dx_1 [(T_{ij',j'} + \rho \omega^2 u_i) \delta u_i^* - T_{i1} \partial_1 u_i^*] + \int dx_3 T_{i1}(0) \delta u_i^*(0) \\ = 0 \end{aligned} \quad (7)$$

Submitting expression (3) into Eq. (7) and completing the integral calculation regarding x_1 yields the following one-dimensional equations

$$\sum_{n=1}^3 \int dx_3 \sum_{m=1}^3 [(\tilde{T}_{ij',j'}^{(m)} - \beta_n^* \tilde{T}_{i1}^{(m)} + \rho \omega^2 u_i^{(m)}) A_{mn} + \tilde{T}_{i1}^{(m)}] \delta u_i^{(n)*} = 0 \quad (8)$$

Here

$$A_{mn} = \int_{-\infty}^0 dx_1 \exp[(\beta_m + \beta_n^*) x_1] = 1/(\beta_n^* + \beta_m)$$

and

$$\tilde{T}_{ij}^{(m)} = c_{ijkl} u_{k,l}^{(m)}(x_3) \quad (9)$$

In Eqs. (8) and (9), the algorithm of the derivatives are: $u_{j,1}^{(m)} \equiv \beta_m u_j^{(m)}(x_3)$, $u_{j,2}^{(m)} \equiv K u_j^{(m)}(x_3)$, and $u_{j,3}^{(m)} \equiv \partial_j^{(m)}(x_3) / \partial x_3$. The algorithm is also adopted for $\tilde{T}_{ij',j'}^{(m)}$.

Because in Eq. (8) $\delta u_i^{(n)}$ are independent, the following equations can be derived from Eq. (8)

$$\begin{aligned} \sum_{m=1}^3 [(\tilde{T}_{ij',j'}^{(m)} - \beta_n^* \tilde{T}_{i1}^{(m)} + \rho \omega^2 u_i^{(m)}) A_{mn} + \tilde{T}_{i1}^{(m)}] \\ = 0; \quad (i, n = 1, 2, 3, \quad j' = 2, 3) \end{aligned} \quad (10)$$

Here $\sum_{m=1}^3 \tilde{T}_{i1}^{(m)}$ are the normal stresses on the surface $x_1 = 0$.

Now Eq. (10) is a one-dimensional field equation without considering the boundary conditions at the surface $x_1 = 0$.

2.2 Boundary Conditions for the One-Dimensional Theory for SAWs

2.2.1 The Boundary Conditions for a Surface Without Mass Loaded Effect. For a surface without mass loaded effect, the traction free boundary conditions on the surface are

$$T_{1j} = \sum_{m=1}^3 \tilde{T}_{1j}^{(m)} = 0, \quad (j = 1, 2, 3) \quad (11)$$

Submitting Eq. (11) into Eq. (10) yields the following one-dimensional field equations

$$\sum_{m=1}^3 [(\tilde{T}_{ij',j'}^{(m)} - \beta_n^* \tilde{T}_{i1}^{(m)} + \rho \omega^2 u_i^{(m)}) A_{mn}] = 0, \quad (i, n = 1, 2, 3; \quad j' = 2, 3) \quad (12)$$

It is worth indicating that Eq. (12) includes the boundary conditions on the surface.

2.2.2 The Boundary Conditions for a Surface With Mass Loaded Effect. Provided that the surface is covered by an anisotropic solid film with thickness h , some difficulties will arise. It is fortunate that by turning to the approximate theory [9,10] for the wave propagation in an infinite plate covered by a film the difficulties can be overcome. In this paper the approximate theory will be generalized for studying a finite structure and the validity of the approximate theory will be investigated as well.

The field equations in the film are

$$\tilde{T}_{ij,j} + \tilde{\rho} \omega^2 \tilde{u}_i = 0 \quad (13)$$

The components of the stress tensor $\tilde{T}$ in the film are

$$\tilde{T}_{ij} = \tilde{c}_{ijkl} \tilde{u}_{k,l} \quad (14)$$

where a variable with a symbol $\sim$ on it implies that the variable is defined in the film.

Due to the traction free boundary conditions $\tilde{T}_{1j}(h) = 0$ on the top surface of the film ($x_1 = h$), the traction continuous boundary conditions at the interface $x_1 = 0$ can be expressed as

$$\begin{aligned} T_{1j}(0) = \tilde{T}_{1j}(0) - \tilde{T}_{1j}(h) = \tilde{T}_{1j}(0) - [\tilde{T}_{1j}(0) + \tilde{T}_{1j,1}(0)h \\ + (1/2)\tilde{T}_{1j,11}(0)h^2 + \delta(h^3)] \end{aligned} \quad (15)$$

because the thickness of the film is very small, so higher-order terms in Eq. (15) can be dropped. Thus Eq. (15) can be expressed as

$$\tilde{T}_{1j}(0) = -\tilde{T}_{1j,1}(0)h \quad (16)$$

From Eq. (13), the following equation is derived

$$\tilde{T}_{j1,1} + K\tilde{T}_{j2} + \tilde{T}_{j3,3} + \tilde{\rho} \omega^2 \tilde{u}_j = 0 \quad (17)$$

Because practical waveguides are generally single mode waveguides, the module of the component of the wave vector in the x_3 direction is small and can be ignored compared with the module of K , which will be called single mode waveguide approximation hereafter. So the $\tilde{T}_{j3,3}$ term in Eq. (17) can be ignored. Thus Eq. (17) can be rewritten as

$$\tilde{T}_{1j,1} + K\tilde{T}_{2j} = -\tilde{\rho} \omega^2 \tilde{u}_j \quad (18)$$

After carrying out similar approximate procedures from Eq. (14), the following equations can be derived

$$\tilde{T}_{1j} = \tilde{c}_{1jk1} \tilde{u}_{k,1} + K\tilde{c}_{1jk2} \tilde{u}_{k,1}, \quad (j = 1, 2, 3) \quad (19)$$

$$\tilde{T}_{1j,1} = \tilde{c}_{1jk1} \tilde{u}_{k,11} + K\tilde{c}_{1jk2} \tilde{u}_{k,1}, \quad (j = 1, 2, 3) \quad (20)$$

$$\tilde{T}_{2j} = \tilde{c}_{2jk1} \tilde{u}_{k,1} + K\tilde{c}_{2jk2} \tilde{u}_{k,1}, \quad (j = 1, 2, 3) \quad (21)$$

Submitting Eqs. (20) and (21) into Eq. (18) yields

$$\begin{aligned} \tilde{c}_{1jk1} \tilde{u}_{k,11} + K\tilde{c}_{1jk2} \tilde{u}_{k,1} + K\tilde{c}_{2jk1} \tilde{u}_{k,1} \\ = -\tilde{\rho} \omega^2 \tilde{u}_j - K^2 \tilde{c}_{2jk2} \tilde{u}_{k,1}, \quad (j = 1, 2, 3) \end{aligned} \quad (22)$$

Equations (22) and (19) can be rewritten as a matrix equation as

$$A(K)X = B(\omega, K)Y \quad (23)$$

where $A(K)$ and $B(\omega, K)$ are 6×6 matrices, $X = [\tilde{u}_{1,11}, \tilde{u}_{2,11}, \tilde{u}_{3,11}, \tilde{u}_{1,1}, \tilde{u}_{2,1}, \tilde{u}_{3,1}]^T$, and $Y = [\tilde{T}_{11}, \tilde{T}_{12}, \tilde{T}_{13}, \tilde{u}_1, \tilde{u}_2, \tilde{u}_3]^T$. The expressions of the matrices $A(K)$ and $B(\omega, K)$ can be derived easily and directly, and they are not listed here.

From Eq. (23), the following expression is derived

$$X = A^{-1}(K)B(\omega, K)Y \quad (24)$$

From Eqs. (20) and (24), it is detected that the $\tilde{T}_{1j,1}$ have the following forms

$$\tilde{T}_{1j,1} = f_j(\tilde{T}_{11}, \tilde{T}_{12}, \tilde{T}_{13}, \tilde{u}_1, \tilde{u}_2, \tilde{u}_3) \quad (25)$$

Now submitting Eq. (25) to Eq. (16) and considering the continuous boundary conditions of the elastic displacement and the normal stresses yields

$$T_{1j}(0) = -hf_j(T_{11}, T_{12}, T_{13}, u_1, u_2, u_3), \quad (x_1 = 0) \quad (26)$$

Equation (26) can be rewritten as follows

$$T_{1j}(0) = g_j(u_1(0), u_2(0), u_3(0)) \quad (27)$$

Equation (27) denotes the mechanical boundary conditions at the interface $x_1 = 0$.

Submitting Eq. (27) to Eq. (10) yields

$$\sum_{m=1}^3 [(\tilde{T}_{ij'}^{(m)} - \beta_n^* \tilde{T}_{il}^{(m)} + \rho \omega^2 u_i^{(m)}) A_{mn} + g_i^{(m)}] = 0, \quad (i, n = 1, 2, 3; \quad j' = 2, 3) \quad (28)$$

It is also worth indicating that Eq. (28) includes all the boundary conditions on the interface.

2.2.3 The Solution of the Eigenmode Problem of a Strip Surface Acoustic Waveguide. A section of a strip surface acoustic waveguide is shown in Fig. 1.

The field variables in the nonmass-loaded regions can be derived from Eq. (12). Assuming all the field variables have the form: $u_\alpha^{(m)}(x_3) = u_\alpha^{(m)} \exp(\gamma x_3)$, $\alpha = 1, 2, 3$, a matrix eigenequation can be derived from Eq. (12), and it is expressed as

$$\gamma^2 CY + \gamma DY + EY = 0 \quad (29)$$

Here C , D , and E are 9×9 matrices, and their components can be expressed as

$$C_{3(n-1)+i, 3(m-1)+j} = A_{mn} c_{i3j3}$$

$$D_{3(n-1)+i, 3(m-1)+j} = A_{mn} (c_{i2j3} K + c_{i3j1} \beta_m + c_{i3j2} K - \beta_n^* c_{i1j3}) + c_{i1j3}$$

$$E_{3(n-1)+i, 3(m-1)+j} = A_{mn} (c_{i2j1} K \beta_m + c_{i2j2} K^2 - c_{i1j1} \beta_n^* \beta_m - c_{i1j2} \beta_n^* K + \rho \omega^2 \delta_{ij}) + c_{i1j1} \beta_m + c_{i1j2} K$$

Here $j = 1, 2, 3$. Y is a 9×1 vector, and $Y = [u_1^{(1)}, u_2^{(1)}, u_3^{(1)}, \dots, u_1^{(3)}, u_2^{(3)}, u_3^{(3)}]^T$, where $u_\alpha^{(m)}$ are unknown constants to be determined.

The solution of Eq. (29) is well known, and it is not described here. From Eq. (29), the eigenvalues and eigenvectors can be obtained: The eigenvalues with positive real parts are: $\gamma_1, \gamma_2, \dots, \gamma_9$, and the corresponding eigenvectors are: $Y^{(1)}, Y^{(2)}, \dots, Y^{(9)}$. The eigenvalues with negative real parts are: $\gamma_{10}, \gamma_{11}, \dots, \gamma_{18}$, and the corresponding eigenvectors are: $Y^{(10)}, Y^{(11)}, \dots, Y^{(18)}$.

For the $x_3 > w$ region, the field variables can be expressed as: $u_\alpha^{(m)}(x_3) = \sum_{L=1}^9 H_{L+9} \exp(\lambda_{L+9} x_3) Y_{\alpha+3m-3}^{(L+9)}$. For the $x_3 < -w$ region, the field variables can be expressed as: $u_\alpha^{(m)}(x_3) = \sum_{L=1}^9 H_L \exp(\lambda_L x_3) Y_{\alpha+3m-3}^{(L)}$. Here H_L ($L = 1, 2, \dots, 18$) are unknown constants to be determined, and $Y_j^{(L)}$ denotes the j th com-

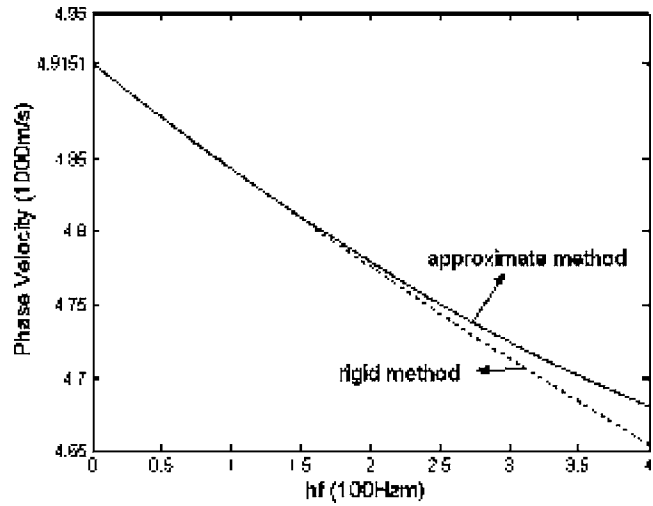


Fig. 2 The relationship between the phase velocity of SAWs and the thickness–frequency product in a SiO_2/Si structure, where the film SiO_2 is infinite in the x_2 and x_3 directions

ponent of the L th eigenvector.

According to a similar method, the field variables in the mass-loaded region can be derived from Eq. (28). Assuming all the field variables have the form: $u_\alpha^{(m)}(x_3) = u_\alpha^{(m)} \exp(\gamma x_3)$, a matrix eigenequation can also be derived from Eq. (28), and it is expressed as

$$\bar{\gamma}^2 \bar{C} \bar{Y} + \bar{\gamma} \bar{D} \bar{Y} + \bar{E} \bar{Y} = 0 \quad (30)$$

where $\bar{C}$, $\bar{D}$, and $\bar{E}$ are 9×9 matrices; $\bar{Y}$ is a 9×1 vector; and $\bar{Y} = [u_1^{(1)}, u_2^{(1)}, u_3^{(1)}, \dots, u_1^{(3)}, u_2^{(3)}, u_3^{(3)}]^T$.

The eigenvalues of Eq. (30) are: $\bar{\gamma}_1, \bar{\gamma}_2, \dots, \bar{\gamma}_{18}$, and the corresponding eigenvectors are: $\bar{Y}^{(1)}, \bar{Y}^{(2)}, \dots, \bar{Y}^{(18)}$. The field variables in the $-w < x_3 < w$ region can be expressed as: $u_\alpha^{(m)}(x_3) = \sum_{L=1}^{18} \bar{H}_L \exp(\bar{\lambda}_L x_3) \bar{Y}_{\alpha+3m-3}^{(L)}$. Here $\bar{H}_L$ ($L = 1, 2, \dots, 18$) are unknown constants to be determined.

In order to obtain the characteristics of the eigenmodes, some additional boundary conditions must be introduced. They are: the $u_\alpha^{(m)}(x_3)$ and $u_{\alpha,3}^{(m)}(x_3)$ must be continuous at $x_3 = \pm w$. The number of the continuous conditions is 36. They result in a matrix equation

$$P(\omega, K)Q = 0 \quad (31)$$

where $P(\omega, K)$ is a 36×36 matrix; Q is a 36×1 vector; and $Q = [H_1, \dots, H_{18}, \bar{H}_1, \dots, \bar{H}_{18}]^T$. In order to get a nontrivial solution, the determinant of the matrix P must be zero, i.e.

$$\det(P(\omega, K)) = 0 \quad (32)$$

Equation (32) is the dispersion equation of the strip surface waveguide. The phase velocity and the field distribution of the eigenmodes can be obtained from Eq. (32).

3 Some Numerical Examples and the Comparison With Other Methods

In order to check the precision of the approximate boundary conditions (Eq. (27)) and the validity of Eq. (28), the relationship between the phase velocity of SAWs and the thickness–frequency product in a SiO_2/Si structure, where the film SiO_2 is infinite in the x_2 and x_3 directions, is investigated by the rigid method [11] and the approximate method (i.e., Eq. (28)). The computation results are shown in Fig. 2. In the figure the dot and solid lines denote the curves calculated by the rigid and approximate methods, respectively. It is detected that the precision of the approxi-

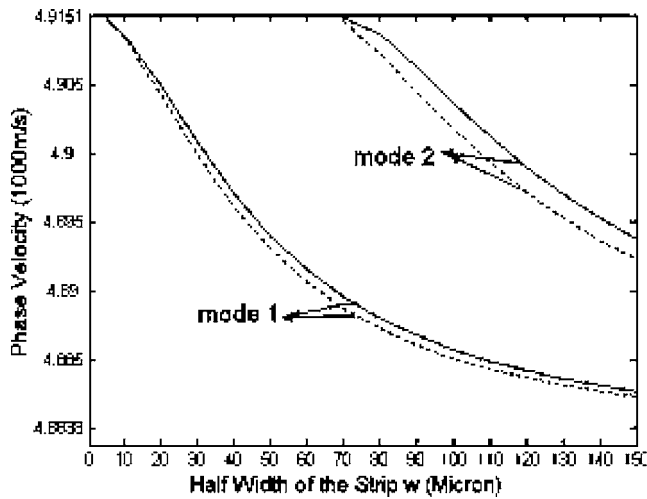


Fig. 3 The eigenmodes in the strip waveguide (SiO_2/Si) structure, where the thickness of the film SiO_2 is $h=0.25 \mu\text{m}$, and the frequency of waves is $f=\omega/2\pi=0.17 \text{ GHz}$

mate boundary conditions (Eq. (27)) is enough for the practical range of hf , such as $hf < 200 \text{ Hz m}$, where f denotes the frequency of the waves. For obtaining higher precision, high-order terms in Eq. (15) must also be considered and it is obvious that Eq. (28) is valid for uniform multilayered structures.

The relationship between the phase velocity of the eigenmodes in the strip waveguide (SiO_2/Si) structure and the half-width of the strip w is shown in Fig. 3. In Fig. 3 the dot and solid lines denote the curves calculated by the EIM and the method proposed in this paper, respectively. It is detected that the results obtained from the two methods are consistent. A slight larger difference between the two methods is observed for the second mode, which may be resulted from the single mode waveguide approximation or the error of the EIM. It may be worth studying in more detail in the future.

The field distribution of the first and second eigenmodes is shown in Figs. 4–7. Because u_3 is almost zero anywhere for the first and second eigenmodes, it is not shown in the figures. From these figures, we can observe the following results: (1) the first mode is symmetric, and the second mode is antisymmetric; and (2) although there is the uncontinuity of mechanical boundary

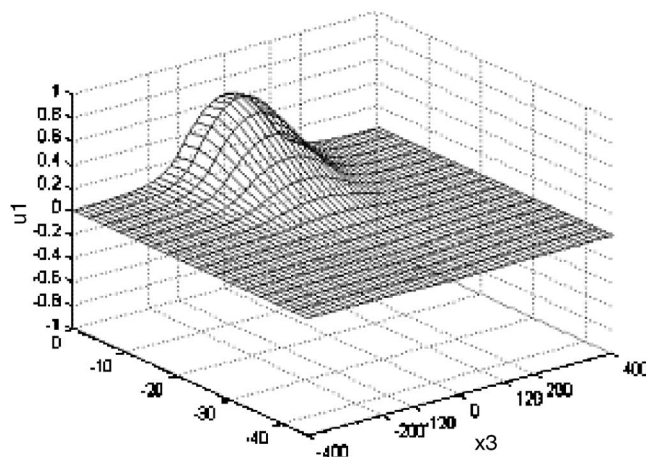


Fig. 4 The relative amplitude of u_1 of the first mode, where $h=0.25 \mu\text{m}$, $w=120 \mu\text{m}$, and $f=0.17 \text{ GHz}$

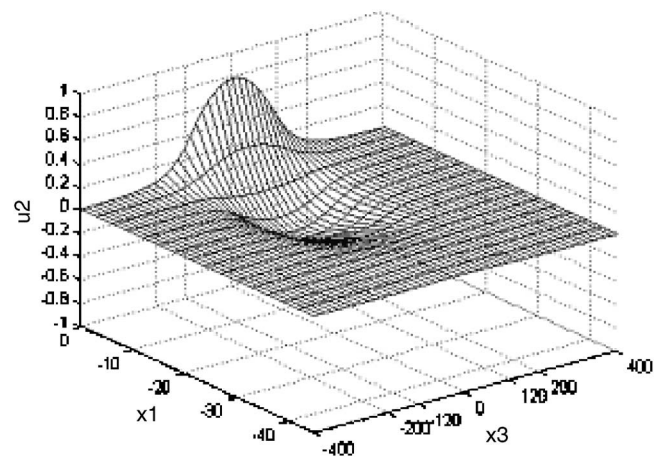


Fig. 5 The relative amplitude of u_2 of the first mode, where $h=0.25 \mu\text{m}$, $w=120 \mu\text{m}$, and $f=0.17 \text{ GHz}$

conditions, the elastic displacement component u_3 is small and can be ignored. In other words, the waves in the waveguides are Rayleigh waves.

In the numerical calculation, all the material parameters are taken from the literature [12].

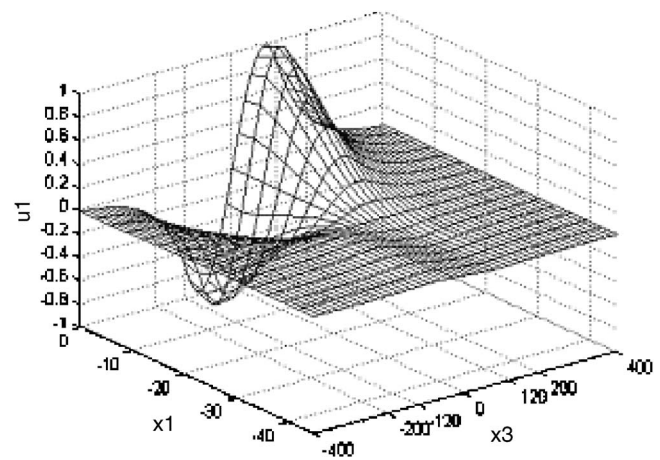


Fig. 6 The relative amplitude of u_1 of the second mode, where $h=0.25 \mu\text{m}$, $w=120 \mu\text{m}$, and $f=0.17 \text{ GHz}$

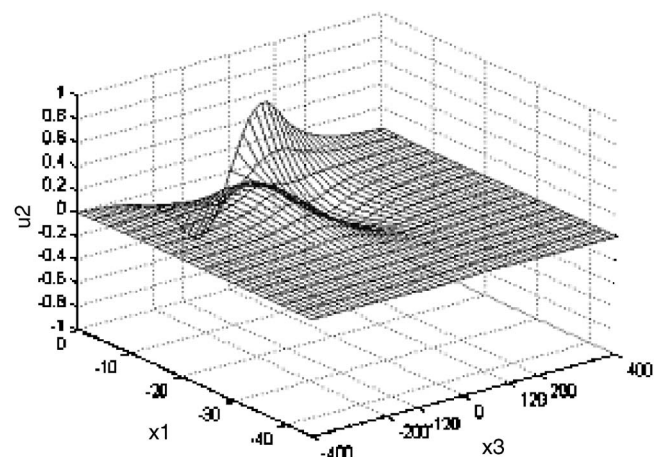


Fig. 7 The relative amplitude of u_2 of the second mode, where $h=0.25 \mu\text{m}$, $w=120 \mu\text{m}$, and $f=0.17 \text{ GHz}$

4 Conclusions

In this paper, a new and effective one-dimensional theory for studying characteristics of SAWs in strip waveguides is presented. The prominent advantage is that the field distribution can be calculated, and this is very important for the design of some acoustic devices.

References

- [1] Peverini, O. A., and Orta, R., 2004, "Film-Loaded SAW Waveguides for Integrated Acousto-Optical Polarization Converters," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, **51**(10), pp. 1298–1307.
- [2] Saitoh, K., Koshihara, M., and Tsuji, Y., 1998, "Numerical Analysis of Integrated Acoustooptic Tunable Filters With Weighted Coupling," *IEEE Ultrasonic Symposium*, May 10, Sendai, Japan, pp. 1293–1296.
- [3] Tiersten, H. F., 1969, "Elastic Surface Waves Guided by Thin Films," *J. Appl. Phys.*, **40**(2), pp. 770–789.
- [4] Peverini, O. A., Orta, R., and Tascone, R., 1993, "Analysis of Piezoelectric Strip Waveguides Based on the Effective Index and Pseudospectral Element Methods," *J. Acoust. Soc. Am.*, **112**, pp. 2623–2633.
- [5] Yen, K. H., and Oliner, A. A., 1976, "The Strip Acoustic Surface Waveguide: Comparison Between Measurement and Theory," *Appl. Phys. Lett.*, **28**(7), pp. 368–370.
- [6] Wang, J., and Lin, J., 2005, "A Two-Dimensional Theory for Surface Acoustic Wave Propagation in Finite Piezoelectric Solids," *J. Intell. Mater. Syst. Struct.*, **16**, pp. 623–629.
- [7] Mindlin, R. D., 1972, "High-Frequency Vibrations of Piezoelectric Crystal Plates," *Int. J. Solids Struct.*, **8**, pp. 891–906.
- [8] Abe, H., and Sato, T., 2000, "Boundary Integral Equations From Hamilton Principle for Surface Acoustic Waves Under Periodic Metal Gratings," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, **47**(6), pp. 1601–1603.
- [9] Bovik, P., 1994, "On the Modelling of Thin Interface Layers in Elastic and Acoustic Scattering Problems," *Q. J. Mech. Appl. Math.*, **47**, pp. 17–42.
- [10] Bovik, P., 1996, "A Comparison Between the Tiersten Model and $O(h)$ Boundary Conditions for Elastic Surface Waves Guided by Thin Layers," *ASME Trans. J. Appl. Mech.*, **63**, pp. 162–167.
- [11] Adler, E. L., 1990, "Matrix Method Applied to Acoustic Waves in Multilayers," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, **37**(6), pp. 485–490.
- [12] Auld, B. A., 1990, *Acoustic Fields and Waves in Solids*, Krieger, New York, Vol. I, Chap. 8.

Series Solution of Three-Dimensional Unsteady Laminar Viscous Flow Due to a Stretching Surface in a Rotating Fluid

Yue Tan

e-mail: tanyue@sjtu.edu.cn

Shi-Jun Liao¹

e-mail: sjliao@sjtu.edu.cn

State Key Laboratory of Ocean Engineering,
School of Naval Architecture, Ocean, and Civil
Engineering,
Shanghai Jiao Tong University,
Shanghai 200030, China

An analytic technique, namely the homotopy analysis method, is applied to solve the Navier–Stokes equations governing unsteady viscous flows due to a suddenly stretching surface in a rotating fluid. Unlike perturbation methods, the current approach does not depend upon any small parameters at all. Besides contrary to all other analytic techniques, it provides us with a simple way to ensure the convergence of solution series. In contrast to perturbation approximations which have about 40% average errors for the considered problem, our series solutions agree well with numerical results in the whole time region $0 \leq t < +\infty$. Explicit analytic expressions of the skin friction coefficients are given, which agree well with numerical results in the whole time region $0 \leq t < +\infty$. This analytic approach can be applied to solve some complicated three-dimensional unsteady viscous flows governed by the Navier–Stokes equations. [DOI: 10.1115/1.2723816]

Keywords: unsteady, Navier–Stokes equation, series solution, rotating fluid, homotopy analysis method

1 Introduction

The flow of a rotating fluid past a stretching surface is encountered in many technical and industrial applications which include the cooling of an infinite metallic plate in a cooling bath, the boundary layer along material handling conveyers, the aerodynamic extrusion of plastic sheets, the boundary layer along a liquid film and condensation processes, the cooling and drying of paper and textiles, the glass fiber production, etc. In particular, in extrusion of a polymer in a melt-spinning process, the extrusion from the die is generally drawn and simultaneously stretched into a sheet, which is then solidified throughout quenching or gradual cooling by direct contact with water.

Wang [1] first considered the two-dimensional stretching of a surface in a rotating fluid. However, relatively little work has been done on unsteady boundary layer flow due to impulsive starting from rest of a stretching sheet in a viscous fluid [2,3]. Nazar et al. [4] solved the unsteady boundary layer flows due to a stretching surface in a rotating fluid by means of both the Keller-box numerical method [5] and the perturbation technique. However, the perturbation approximations given by Nazar et al. [4] have about 40% average error, and are *not* accurate enough in the *whole* time region, as shown in Figs. 4 and 5. Besides, it becomes more and more difficult to get higher-order perturbation approximations. This is mainly because perturbation techniques depend too strongly upon small parameters.

An analytical method for strongly nonlinear problems, namely the homotopy analysis method (HAM) [6–11], has been developed since 1992. In contrast to perturbation techniques, the homotopy analysis method is *independent* of any small parameters at all. Besides, it provides us with a simple way to ensure the convergence of the solution series, so that we can always get accurate enough approximations. Furthermore, it provides us with freedom to choose better basis functions to approximate nonlinear problems. Finally, as proved by Liao [7,9], the homotopy analysis

method logically contains the so-called nonperturbation methods such as Lyapunov's artificial small parameter method, the δ -expansion method, and Adomian's decomposition method. Using the relationship between the homotopy analysis method and Adomian's decomposition method, Allan [12] investigated the accuracy of approximations given by the Adomian's decomposition method. Currently, Hayat et al. [13], Sajid et al. [14], and Abbasbandy [15,16] pointed out that the so-called "homotopy perturbation method" [17] proposed in 1999 is also a special case of the homotopy analysis method [6,7] propounded in 1992. Thus, the homotopy analysis method is rather general. The homotopy analysis method has been successfully applied to many nonlinear problems in science and engineering, such as the similarity boundary-layer flows [9,18–21], nonlinear heat transfer [15], nonlinear evaluation equations [22], nonlinear waves [16], viscous flows of non-Newtonian fluid [13,14,23,24], Thomas–Fermi atom model [7], Volterra's population model [7], etc. It has been applied in many fields of research. For example, Zhu [25,26] applied the HAM to give, for the first time, an explicit series solution of the famous Black–Scholes type equation in finance for American put option, which is a system of nonlinear partial differential equations (PDEs) with an unknown moving boundary. Besides, the HAM has been successfully applied to solve some PDEs in fluid mechanics and heat transfer, such as the unsteady boundary-layer viscous flows [10], the unsteady nonlinear heat transfer problem [27], etc. In this paper, we further employ it to give much more *accurate* analytic approximations (with less than 0.5% error in the *whole* time region) of the unsteady nonlinear problem at hand.

2 Mathematical Description

Let (u, v, w) be the velocity components in the direction of Cartesian axes (x, y, z) , respectively, with the axes rotating at an angular velocity Ω in the z direction. Consider the unsteady boundary-layer flows caused by a stretching surface at $z=0$ in a rotating fluid. When $t < 0$, the surface rotates at an angular velocity Ω in the z direction so that the fluid is at rest relative to the surface. At time $t=0$, the surface at $z=0$ is impulsively stretched

¹Corresponding author.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received September 8, 2006; final manuscript received January 29, 2007. Review conducted by Nesreen Ghaddar.

in the x direction. Due to the Coriolis force, the fluid motion is three dimensional and is governed by the continuity equation and the unsteady Navier–Stokes equations [4]

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0 \quad (1a)$$

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} + w \frac{\partial u}{\partial z} - 2\Omega v = -\frac{1}{\rho} \frac{\partial p}{\partial x} + \nu \nabla^2 u \quad (1b)$$

$$\frac{\partial v}{\partial t} + u \frac{\partial v}{\partial x} + v \frac{\partial v}{\partial y} + w \frac{\partial v}{\partial z} + 2\Omega u = -\frac{1}{\rho} \frac{\partial p}{\partial y} + \nu \nabla^2 v \quad (1c)$$

$$\frac{\partial w}{\partial t} + u \frac{\partial w}{\partial x} + v \frac{\partial w}{\partial y} + w \frac{\partial w}{\partial z} = -\frac{1}{\rho} \frac{\partial p}{\partial z} + \nu \nabla^2 w \quad (1d)$$

where p denotes the pressure; ρ the density; ν the kinematic viscosity; and ∇^2 the three-dimensional Laplacian, respectively. The initial and boundary conditions are

$$t < 0: \quad u = v = w = 0 \quad \text{for any } x, y, z \quad (1e)$$

$$t \geq 0: \quad u = ax, \quad v = w = 0 \quad \text{at } z = 0 \quad (1f)$$

$$u \rightarrow 0, \quad v \rightarrow 0, \quad w \rightarrow 0 \quad \text{as } z \rightarrow \infty \quad (1g)$$

where the constant a ($a > 0$, with the dimension of t^{-1}) represents the stretching rate.

Using Williams and Rhyne's similarity transformation [28]

$$\xi = 1 - e^{-\tau}, \quad \tau = at \quad (2)$$

and introducing the following similarity variables

$$\eta = \sqrt{\frac{a}{\nu \xi}} z, \quad u = ax f'(\xi, \eta), \quad v = ax g(\xi, \eta), \quad w = -\sqrt{a \nu \xi} f(\xi, \eta) \quad (3)$$

Equations (1a)–(1d) become a system of two coupled nonlinear differential equations

$$f''' + \frac{1}{2}(1 - \xi) \eta f'' + \xi(f f'' - f'^2 + 2\lambda g) - \xi(1 - \xi) \frac{\partial f'}{\partial \xi} = 0 \quad (4a)$$

$$g'' + \frac{1}{2}(1 - \xi) \eta g' + \xi(f g' - f' g - 2\lambda f') - \xi(1 - \xi) \frac{\partial g}{\partial \xi} = 0 \quad (4b)$$

where the prime denotes the derivation with respect to the similarity variable η ; and $\lambda = \Omega/a$ is a dimensionless parameter. The corresponding boundary conditions read

$$f(\xi, 0) = 0, \quad f'(\xi, 0) = 1, \quad f'(\xi, +\infty) = 0, \quad g(\xi, 0) = 0, \quad g(\xi, +\infty) = 0 \quad (4c)$$

Note that, as $\tau \rightarrow 0$, we have $\xi \rightarrow \tau$. Thus, $\eta = \sqrt{a/\nu \xi} z$ is exactly the same as the traditional definition given in Ref. [4], which has meanings as $\tau \rightarrow 0$.

According to the definitions (2) and (3), we have the dimensionless skin friction coefficient in the x and y directions

$$C_{fx} = (\xi \text{Re}_x)^{-1/2} f''(\xi, 0), \quad C_{fy} = (\xi \text{Re}_x)^{-1/2} g'(\xi, 0) \quad (5)$$

where $\text{Re}_x = ax^2/\nu$ is the local Reynolds number.

3 Perturbation Approximations

3.1 Initial Solution as $\xi \rightarrow 0$. As $\xi \rightarrow 0$, corresponding to the initial flow, Eqs. (4a) and (4b) become

$$f''' + \frac{1}{2} \eta f'' = 0 \quad (6a)$$

$$g'' + \frac{1}{2} \eta g' = 0 \quad (6b)$$

subject to the boundary/initial conditions

$$f(0, 0) = 0, \quad f'(0, 0) = 1, \quad f'(0, \infty) = 0,$$

$$g(0, 0) = 0, \quad g(0, \infty) = 0 \quad (6c)$$

Its solution is given by

$$f(0, \eta) = \eta \text{erfc}(\eta/2) + \frac{2}{\sqrt{\pi}}(1 - e^{-\eta^2/4}) \quad (7)$$

$$g(0, \eta) = 0 \quad (8)$$

where $\text{erfc}(\eta)$ is the so-called complementary error function

$$\text{erfc}(\eta) = \frac{2}{\sqrt{\pi}} \int_{\eta}^{\infty} e^{-s^2} ds \quad (9)$$

Note that Eqs. (6a) and (6b) do not contain the parameter λ . Thus, the result as $\xi \rightarrow 0$ is independent of λ .

3.2 Steady Solution at $\xi = 1$. At $\xi = 1$, corresponding to the steady-state flow, Eqs. (4a) and (4b) read

$$f''' + f f'' - f'^2 + 2\lambda g = 0 \quad (10a)$$

$$g'' + f g' - f' g - 2\lambda f' = 0 \quad (10b)$$

subject to the same boundary and initial conditions as Eq. (6c). The steady-state boundary-layer flows were solved by Wang [1,2].

3.3 Perturbation Approximations for Small ξ . Regarding ξ as a small parameter, one has the perturbation expressions

$$f(\xi, \eta) = \hat{f}_0(\eta) + \hat{f}_1(\eta)\xi + \hat{f}_2(\eta)\xi^2 + \cdots \quad (11)$$

$$g(\xi, \eta) = \hat{g}_0(\eta) + \hat{g}_1(\eta)\xi + \hat{g}_2(\eta)\xi^2 + \cdots \quad (12)$$

Substituting them into Eqs. (4a)–(4c) and equating the coefficients of the like power of ξ , one has the zero-order perturbation equations

$$\hat{f}_0''' + \frac{1}{2} \eta \hat{f}_0'' = 0 \quad (13a)$$

$$\hat{g}_0'' + \frac{1}{2} \eta \hat{g}_0' = 0 \quad (13b)$$

$$\hat{f}_0(0) = 0, \quad \hat{f}_0'(0) = 1, \quad \hat{f}_0'(\infty) = 0, \quad \hat{g}_0(0) = 0, \quad \hat{g}_0(\infty) = 0 \quad (13c)$$

which are exactly the same as Eqs. (6a)–(6c), respectively. The first-order perturbation equations are given by

$$\hat{f}_1''' + \frac{1}{2} \eta \hat{f}_1'' - \hat{f}_1' = \frac{1}{2} \eta \hat{f}_0'' - \hat{f}_0 \hat{f}_0'' + \hat{f}_0'^2 \quad (14a)$$

$$\hat{g}_1'' + \frac{1}{2} \eta \hat{g}_1' - \hat{g}_1 = 2\lambda \hat{f}_0' \quad (14b)$$

$$\hat{f}_1(0) = 0, \quad \hat{f}_1'(0) = 0, \quad \hat{f}_1'(\infty) = 0, \quad \hat{g}_1(0) = 0, \quad \hat{g}_1(\infty) = 0 \quad (14c)$$

Its solution reads

$$\begin{aligned} \hat{f}_1(\eta) = & \left(\frac{1}{2} - \frac{2}{3\pi} \right) \left[\left(1 + \frac{\eta^2}{2} \right) \text{erfc}(\eta/2) - \frac{1}{\sqrt{\pi}} \eta e^{-\eta^2/4} \right] \\ & - \frac{1}{2} \left(1 - \frac{\eta^2}{2} \right) \text{erfc}^2(\eta/2) - \frac{3}{2\sqrt{\pi}} \eta e^{-\eta^2/4} \text{erfc}(\eta/2) + \frac{2}{\pi} e^{-\eta^2/2} \\ & - \frac{1}{\sqrt{\pi}} \left(\frac{1}{4} \eta + \frac{4}{3\sqrt{\pi}} \right) e^{-\eta^2/4} \end{aligned} \quad (15)$$

$$\hat{g}_1(\eta) = \lambda \eta^2 \operatorname{erfc}(\eta/2) - \frac{2}{\sqrt{\pi}} \lambda \eta e^{-\eta^2/4} \quad (16)$$

In general, one can transfer the original nonlinear initial/boundary-value problem into an infinite number of linear boundary-value problems

$$\mathcal{L}_f^p[f_m(\eta)] = R_m^p(\eta), \quad \mathcal{L}_g^p[g_m(\eta)] = G_m^p(\eta) \quad (17)$$

where the two linear operators $\mathcal{L}_f^p$ and $\mathcal{L}_g^p$ are defined by

$$\mathcal{L}_f^p \phi = \frac{\partial^3 \phi}{\partial \eta^3} + \frac{\eta}{2} \frac{\partial^2 \phi}{\partial \eta^2} - \frac{\partial \phi}{\partial \eta} \quad (18)$$

$$\mathcal{L}_g^p \psi = \frac{\partial^2 \psi}{\partial \eta^2} + \frac{\eta}{2} \frac{\partial \psi}{\partial \eta} - \psi \quad (19)$$

The common solutions of the linear equations

$$\mathcal{L}_f^p \phi = 0, \quad \mathcal{L}_g^p \psi = 0$$

are, respectively

$$\begin{aligned} \phi = & C_1 + C_2 \left(\eta + \frac{\eta^3}{6} \right) + C_3 \left[\frac{4}{3} \exp(-\eta^2/4) \left(1 + \frac{\eta^2}{4} \right) + \sqrt{\pi} \operatorname{erf}(\eta/2) \right. \\ & \left. \times \left(\eta - \frac{\eta^3}{6} \right) \right] \end{aligned} \quad (20)$$

and

$$\psi = C_1 \left(1 + \frac{\eta^2}{2} \right) + C_2 \left[\eta \exp(-\eta^2/4) + \sqrt{\pi} \operatorname{erf}(\eta/2) \left(1 + \frac{\eta^2}{2} \right) \right] \quad (21)$$

where C_1 , C_2 , and C_3 are integral constants. Note that the above common solutions contain the error function $\operatorname{erf}(\eta/2)$ and are rather complicated. Owing to this reason, it becomes more and more difficult to get higher-order perturbation approximations. It should be emphasized that these two linear operators $\mathcal{L}_f^p$ and $\mathcal{L}_g^p$ come directly from the original governing equations. Thus, the perturbation method does *not* provides us with any freedom to choose the linear operators of the linear subproblems.

The first-order perturbation approximation of the skin friction coefficients reads

$$C_f^x \operatorname{Re}_x^{1/2} \approx \xi^{-1/2} [\hat{f}_0''(0) + \xi \hat{f}_1''(0)] = \frac{1}{\sqrt{\pi}} \left[-\xi^{-1/2} + \left(-\frac{7}{4} + \frac{4}{3\pi} \right) \xi^{1/2} \right] \quad (22)$$

$$C_g^y \operatorname{Re}_x^{1/2} \approx \xi^{-1/2} [\hat{g}_0'(0) + \xi \hat{g}_1'(0)] = -\frac{2}{\sqrt{\pi}} \lambda \xi^{1/2} \quad (23)$$

For details, please refer to Nazar et al. [4]. Note that, due to the appearance of the error function $\operatorname{erfc}(\eta/2)$, it becomes more and more difficult to get higher-order perturbation approximations, as mentioned above. Note that the perturbation approximation of $C_f^x \operatorname{Re}_x^{1/2}$ is independent of λ . However, for large λ , these perturbation approximations are not accurate, and would have about 40% average error, as shown in Figs. 4 and 5.

4 Homotopy Analytic Solution

In this part, we solve Eqs. (4a)–(4c) by means of the homotopy analysis method [7,9]. First of all, it is well known that most of boundary-layer flows decay exponentially at infinity (i.e., $\eta \rightarrow +\infty$). Thus, according to the boundary conditions (4c), f' and g decay to zero at infinity *exponentially*. Besides, Eqs. (4a) and (4b) explicitly contain the terms ξ and η . Thus, $f(\xi, \eta)$ and $g(\xi, \eta)$ should be expressed by such a set of basis functions

$$\{\xi^k \eta^m \exp(-n\eta) | k \geq 0, m \geq 0, n \geq 0\} \quad (24)$$

in the following forms

$$f(\xi, \eta) = a_0^{0,0} + \sum_{k=0}^{+\infty} \sum_{m=0}^{+\infty} \sum_{n=1}^{+\infty} a_k^{m,n} \xi^k \eta^m \exp(-n\eta) \quad (25)$$

$$g(\xi, \eta) = \sum_{k=0}^{+\infty} \sum_{m=0}^{+\infty} \sum_{n=1}^{+\infty} b_k^{m,n} \xi^k \eta^m \exp(-n\eta) \quad (26)$$

where $a_k^{m,n}$ and $b_k^{m,n}$ are coefficients. Note that the solutions expressed in the above forms decay *exponentially* as $\eta \rightarrow +\infty$. As shown later, the above expressions are important in the frame of the homotopy analysis method for the choice of the initial guesses and the auxiliary linear operators. The initial guesses and auxiliary linear operators should be chosen in such a way that the approximations must be expressed by the above two expressions: this is so important that it is called *the rule of solution expressions* [7,9].

According to the solution expressions (25) and (26) and the boundary conditions (4c), it is straightforward to choose the initial guesses

$$f_0(\xi, \eta) = 1 - \exp(-\eta) \quad (27)$$

$$g_0(\xi, \eta) = 0 \quad (28)$$

Note that the above initial guesses satisfy the initial/boundary conditions (4c). Let $\mathcal{L}_f$ and $\mathcal{L}_g$ denote two auxiliary linear operators, which we will determine later. Here, we note that we have great freedom to choose $\mathcal{L}_f$ and $\mathcal{L}_g$. Based on Eqs. (4a) and (4b), we define the following two nonlinear operators

$$\begin{aligned} \mathcal{N}_f[F, G] = & \frac{\partial^3 F}{\partial \eta^3} + \frac{1}{2} (1 - \xi) \eta \frac{\partial^2 F}{\partial \eta^2} - \xi (1 - \xi) \frac{\partial^2 F}{\partial \xi \partial \eta} + \xi \left[F \frac{\partial^2 F}{\partial \eta^2} \right. \\ & \left. - \left(\frac{\partial F}{\partial \eta} \right)^2 + 2\lambda G \right] \end{aligned} \quad (29)$$

and

$$\begin{aligned} \mathcal{N}_g[F, G] = & \frac{\partial^2 G}{\partial \eta^2} + \frac{1}{2} (1 - \xi) \eta \frac{\partial G}{\partial \eta} - \xi (1 - \xi) \frac{\partial G}{\partial \xi} + \xi \left(F \frac{\partial G}{\partial \eta} - G \frac{\partial F}{\partial \eta} \right. \\ & \left. - 2\lambda \frac{\partial F}{\partial \eta} \right) \end{aligned} \quad (30)$$

Then, we construct the so-called zero-order deformation equations

$$(1 - q) \{ \mathcal{L}_f[F(\xi, \eta; q)] - \mathcal{L}_f[f_0(\xi, \eta)] \} = q \hbar \mathcal{N}_f[F(\xi, \eta; q), G(\xi, \eta; q)] \quad (31a)$$

$$(1 - q) \{ \mathcal{L}_g[G(\xi, \eta; q)] - \mathcal{L}_g[g_0(\xi, \eta)] \} = q \hbar \mathcal{N}_g[F(\xi, \eta; q), G(\xi, \eta; q)] \quad (31b)$$

subject to the corresponding boundary/initial conditions

$$F(\xi, 0; q) = 0, \quad \left. \frac{\partial F(\xi, \eta; q)}{\partial \eta} \right|_{\eta=0} = 1, \quad \left. \frac{\partial F(\xi, \eta; q)}{\partial \eta} \right|_{\eta=\infty} = 0 \quad (31c)$$

and

$$G(\xi, 0; q) = 0, \quad G(\xi, \infty; q) = 0 \quad (31d)$$

where $q \in [0, 1]$ is the embedding parameter; $\hbar$ denotes a nonzero auxiliary parameter; and $F(\xi, \eta; q)$ and $G(\xi, \eta; q)$ are unknown functions related to $f(\xi, \eta)$ and $g(\xi, \eta)$, respectively.

When $q=0$, since the initial guesses $f_0(\xi, \eta)$ and $g_0(\xi, \eta)$ satisfy the initial/boundary conditions (4c), the above zero-order deformation equations have the solution

$$F(\xi, \eta; 0) = f_0(\xi, \eta), \quad G(\xi, \eta; 0) = g_0(\xi, \eta) \quad (32)$$

When $q=1$, since $\hbar \neq 0$, the above zero-order deformation equations are equivalent to the original Eqs. (4a)–(4c), provided

$$F(\xi, \eta; 1) = f(\xi, \eta), \quad G(\xi, \eta; 1) = g(\xi, \eta) \quad (33)$$

respectively. Thus, as q increases from 0 to 1, $F(\xi, \eta; q)$ and $G(\xi, \eta; q)$ vary (or *deform*) from the known initial guesses $f_0(\xi, \eta)$ and $g_0(\xi, \eta)$ to the unknown solutions $f(\xi, \eta)$ and $g(\xi, \eta)$ of the original Eq. (4a)–(4c), respectively. This is the reason why Eqs. (31a)–(31d) are called zero-order deformation equations.

By Taylor's theorem and using Eq. (32), we have the power series

$$F(\xi, \eta; q) = f_0(\xi, \eta) + \sum_{n=1}^{+\infty} f_n(\xi, \eta) q^n \quad (34)$$

$$G(\xi, \eta; q) = g_0(\xi, \eta) + \sum_{n=1}^{+\infty} g_n(\xi, \eta) q^n \quad (35)$$

where

$$f_n(\xi, \eta) = \frac{1}{n!} \left. \frac{\partial^n F(\xi, \eta; q)}{\partial q^n} \right|_{q=0}, \quad g_n(\xi, \eta) = \frac{1}{n!} \left. \frac{\partial^n G(\xi, \eta; q)}{\partial q^n} \right|_{q=0} \quad (36)$$

Note that Eqs. (31a) and (31b) contain the auxiliary parameter $\hbar$. Obviously, the convergence of the series (34) and (35) is dependent not only upon the auxiliary linear operators $\mathcal{L}_f$ and $\mathcal{L}_g$ but also the auxiliary parameter $\hbar$, and more importantly, we have great freedom to choose all of them. Assuming that the auxiliary parameter $\hbar$ and the auxiliary linear operators $\mathcal{L}_f$ and $\mathcal{L}_g$ are so properly chosen that the series (34) and (35) converge at $q=1$, we have, using Eq. (33), the solution series

$$f(\xi, \eta) = f_0(\xi, \eta) + \sum_{n=1}^{+\infty} f_n(\xi, \eta) \quad (37)$$

$$g(\xi, \eta) = g_0(\xi, \eta) + \sum_{n=1}^{+\infty} g_n(\xi, \eta) \quad (38)$$

The above series relates the initial guesses $f_0(\xi, \eta)$ and $g_0(\xi, \eta)$ with the exact solution $f(\xi, \eta)$ and $g(\xi, \eta)$ by means of the unknown terms $f_n(\xi, \eta)$ and $g_n(\xi, \eta)$, where $n=1, 2, 3, \dots$. According to the fundamental theorem in calculus, the Taylor series (34) and (35) are *unique*, and are completely determined by the zero-order deformation Eqs. (31a)–(31d). Thus, the governing equations and boundary/initial conditions of $f_n(\xi, \eta)$ and $g_n(\xi, \eta)$ can be deduced directly from the zero-order deformation Eqs. (31a)–(31d). Based on the definition (36) and the solution series (37) and (38), Liao [7–11] provided a rather general approach to obtain the equations governing $f_n(\xi, \eta)$ and $g_n(\xi, \eta)$. For the sake of simplicity, define the vectors

$$\vec{f}_m = \{f_0, f_1, f_2, \dots, f_m\} \quad (39)$$

$$\vec{g}_m = \{g_0, g_1, g_2, \dots, g_m\} \quad (40)$$

Differentiating the zero-order deformation Eqs. (31a)–(31d) n times with respect to the embedding parameter q , then setting $q=0$, and finally dividing by $n!$, we have the so-called n th-order deformation equation

$$\mathcal{L}_f[f_n(\xi, \eta) - \chi_n f_{n-1}(\xi, \eta)] = \hbar R_n^f(\vec{f}_{n-1}, \vec{g}_{n-1}, \xi, \eta) \quad (41a)$$

$$\mathcal{L}_g[g_n(\xi, \eta) - \chi_n g_{n-1}(\xi, \eta)] = \hbar R_n^g(\vec{f}_{n-1}, \vec{g}_{n-1}, \xi, \eta) \quad (41b)$$

subject to the boundary/initial conditions

$$f_n(\xi, 0) = 0, \quad \left. \frac{\partial f_n(\xi, \eta)}{\partial \eta} \right|_{\eta=0} = 0, \quad \left. \frac{\partial f_n(\xi, \eta)}{\partial \eta} \right|_{\eta=\infty} = 0 \quad (41c)$$

and

$$g_n(\xi, 0) = 0, \quad g_n(\xi, \infty) = 0 \quad (41d)$$

where

$$R_n^f(\vec{f}_{n-1}, \vec{g}_{n-1}, \xi, \eta) = \frac{\partial^3 f_{n-1}}{\partial \eta^3} + \frac{1}{2}(1-\xi)\eta \frac{\partial^2 f_{n-1}}{\partial \eta^2} - \xi(1-\xi) \frac{\partial^2 f_{n-1}}{\partial \xi \partial \eta} + \xi \left[\sum_{j=0}^{n-1} \left(f_{n-1-j} \frac{\partial^2 f_j}{\partial \eta^2} - \frac{\partial f_{n-1-j}}{\partial \eta} \frac{\partial f_j}{\partial \eta} \right) + 2\lambda g_{n-1} \right] \quad (41e)$$

and

$$R_n^g(\vec{f}_{n-1}, \vec{g}_{n-1}, \xi, \eta) = \frac{\partial^2 g_{n-1}}{\partial \eta^2} + \frac{1}{2}(1-\xi)\eta \frac{\partial g_{n-1}}{\partial \eta} - \xi(1-\xi) \frac{\partial g_{n-1}}{\partial \xi} + \xi \left[\sum_{j=0}^{n-1} \left(f_{n-1-j} \frac{\partial g_j}{\partial \eta} - g_{n-1-j} \frac{\partial f_j}{\partial \eta} \right) - 2\lambda \frac{\partial f_{n-1}}{\partial \eta} \right] \quad (41f)$$

under the definition

$$\chi_k = \begin{cases} 0, & k \leq 1 \\ 1, & k > 1 \end{cases} \quad (42)$$

As proved by Sajid et al. [13] and Hayat et al. [14], directly substituting the series (34) and (35) into zero-order deformation Eqs. (31a)–(31c), and then equating the coefficients of the like power of q , one can obtain exactly the *same* equations as Eqs. (41a)–(41f), no matter whether q is regarded as a small parameter or not. This is mainly because the Taylor series (34) and (35) are *unique*.

The original Eq. (4a) is third order with respect to the similarity variable η and first order with respect to the dimensionless time ξ , and Eq. (4b) is second order with respect to η and first order with respect to ξ , respectively. In general, it is more difficult to solve these kinds of combined initial-boundary-value problems than pure boundary-value ones, even if they are linear. Fortunately, the homotopy analysis method provides us with great freedom to choose the auxiliary linear operators $\mathcal{L}_f$ and $\mathcal{L}_g$: we can choose $\mathcal{L}_f$ and $\mathcal{L}_g$ in such a way that the high-order deformation equations are pure boundary-value ones. To do so, $\mathcal{L}_f$ may be a third-order linear differential operator with respect to η , and $\mathcal{L}_g$ a second-order linear operator with respect to η , respectively. Therefore, without loss of generality, we write

$$\mathcal{L}_f[\phi(\xi, \eta)] = \frac{\partial^3 \phi}{\partial \eta^3} + A_2(\eta) \frac{\partial^2 \phi}{\partial \eta^2} + A_1(\eta) \frac{\partial \phi}{\partial \eta} + A_0(\eta) \phi \quad (43)$$

and

$$\mathcal{L}_g[\phi(\xi, \eta)] = \frac{\partial^2 \phi}{\partial \eta^2} + B_1(\eta) \frac{\partial \phi}{\partial \eta} + B_0(\eta) \phi \quad (44)$$

where $A_0(\eta)$, $A_1(\eta)$, $A_2(\eta)$, $B_0(\eta)$, and $B_1(\eta)$ are real functions to be determined below. Let $f_n^*(\xi, \eta)$ and $g_n^*(\xi, \eta)$ denote the special solution of Eqs. (41a)–(41d). Its general solutions read

$$f_n(\xi, \eta) = f_n^*(\xi, \eta) + C_1(\xi) e_1^f(\eta) + C_2(\xi) e_2^f(\eta) + C_3(\xi) e_3^f(\eta) \quad (45)$$

$$g_n(\xi, \eta) = g_n^*(\xi, \eta) + C_4(\xi) e_1^g(\eta) + C_5(\xi) e_2^g(\eta) \quad (46)$$

where $C_1(\xi)$, $C_2(\xi)$, $C_3(\xi)$, $C_4(\xi)$, and $C_5(\xi)$ are integral constants determined by the boundary conditions (41c) and (41d), the real functions $e_1^f(\eta)$, $e_2^f(\eta)$, $e_3^f(\eta)$, $e_1^g(\eta)$, and $e_2^g(\eta)$ are the so-called

kernels of the two auxiliary linear operators, satisfying

$$\mathcal{L}_f[e_1^f(\eta)] = \mathcal{L}_f[e_2^f(\eta)] = \mathcal{L}_f[e_3^f(\eta)] = 0, \quad (47)$$

$$\mathcal{L}_g[e_1^g(\eta)] = \mathcal{L}_g[e_2^g(\eta)] = 0$$

According to the solution expressions (25) and (26), the kernels should belong to the basis functions. If we choose $e_1^f(\eta)$, $e_2^f(\eta)$, $e_3^f(\eta)$ as the first three simplest basis functions among (24), i.e.

$$e_1^f(\eta) = 1, \quad e_2^f(\eta) = \exp(-\eta), \quad e_3^f(\eta) = \exp(-2\eta)$$

then, the boundary condition at infinity is automatically satisfied and the coefficient $C_3(\xi)$ cannot be *uniquely* determined. To ensure that the high-order deformation equations have *unique* solutions, we should choose

$$e_1^f(\eta) = 1, \quad e_2^f(\eta) = \exp(-\eta), \quad e_3^f(\eta) = \exp(+\eta)$$

Then, to satisfy the boundary condition $f_n'(\xi, +\infty) = 0$, we have $C_3(\xi) = 0$, and $C_1(\xi)$, $C_2(\xi)$ are determined by the two boundary conditions $f_n(\xi, 0) = 0$ and $f_n'(\xi, 0) = 0$. Substituting the above kernels into Eq. (47), we have

$$A_0(\eta) = 0 \quad (48)$$

$$-[1 + A_1(\eta) - A_2(\eta)]\exp(-\eta) = 0 \quad (49)$$

$$[1 + A_1(\eta) + A_2(\eta)]\exp(+\eta) = 0 \quad (50)$$

which give

$$A_0(\eta) = A_2(\eta) = 0, \quad A_1(\eta) = -1$$

Similarly, choosing the kernels $e_1^g(\eta) = \exp(-\eta)$ and $e_2^g(\eta) = \exp(+\eta)$, we have

$$B_0(\eta) = -1, \quad B_1(\eta) = 0$$

Thus, we have the two auxiliary linear operators

$$\mathcal{L}_f\phi = \frac{\partial^3\phi}{\partial\eta^3} - \frac{\partial\phi}{\partial\eta} \quad (51)$$

$$\mathcal{L}_g\phi = \frac{\partial^2\phi}{\partial\eta^2} - \phi \quad (52)$$

which have the properties

$$\mathcal{L}_f[C_1(\xi)\exp(-\eta) + C_2(\xi)\exp(+\eta) + C_3(\xi)] = 0 \quad (53)$$

$$\mathcal{L}_g[C_4(\xi)\exp(-\eta) + C_5(\xi)\exp(+\eta)] = 0 \quad (54)$$

Note that, in contrast to perturbation approximations, our HAM series solutions do *not* contain the error function $\text{erf}(\eta/2)$. This is mainly because we have great freedom to choose the auxiliary linear operators $\mathcal{L}_f$ and $\mathcal{L}_g$, which are much simpler than $\mathcal{L}_f^p$ and $\mathcal{L}_g^p$ appeared in the high-order perturbation Eqs. (14a) and (14b), respectively.

The high-order deformation Eqs. (41a)–(41d) are linear boundary-value equations. Thus, according to Eqs. (37) and (38), the original *nonlinear*, combined initial-boundary-value problem is transferred into an infinite number of *linear* boundary-value problems. However, in contrast to perturbation techniques, this kind of transformation does *not* need any small parameters. Besides, in contrast to the perturbation method, the homotopy analysis method provides us with great *freedom* to choose the auxiliary linear operators $\mathcal{L}_f$ and $\mathcal{L}_g$. Using this kind of freedom, we can obtain results at rather high order of approximations by means of choosing the linear operators (51) and (52), which are simpler than (18) and (19), and whose kernels do *not* contain the error functions $\text{erf}(\eta/2)$.

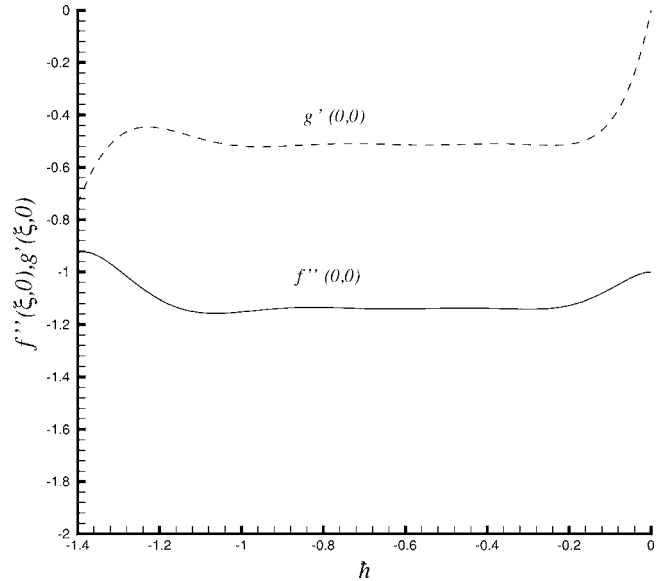


Fig. 1 The 15th-order HAM approximation of $f''(0,0)$ and $g'(0,0)$ in the case of $\lambda = 1/2$

5 Result Analysis

Liao [7] proved in general that, as long as a solution series given by the homotopy analysis method converges, it must be one of the solutions of the equation considered. Thus, it is important to ensure that the HAM solution series are convergent. Note that the solution series (37) and (38) contain one auxiliary parameter $\hbar$. As shown by Liao [7–11] and others [13–16,22], it is the auxiliary parameter $\hbar$ that provides us with a simple way to *adjust* and *control* the convergence region of the solution series. In general, by means of choosing a proper value of the auxiliary parameter $\hbar$, one can always ensure the convergence of the solution series given by the homotopy analysis method. For example, let us consider the case $\lambda = 1/2$ for the nonlinear unsteady problem at hand. First of all, we investigate the convergence of $f''(0,0)$ and $g'(0,0)$ in the case of $\lambda = 1/2$ by means of regarding $\hbar$ as a variable. Physically, $f''(0,0)$ and $g'(0,0)$ are *independent* of the auxiliary parameter $\hbar$, and thus are *unique*. This is mainly because the auxiliary parameter $\hbar$ has only mathematical meanings. Thus, mathematically, as long as the series of $f''(0,0)$ and $g'(0,0)$ are convergent, they must be convergent to the *same* values, respectively. As shown in Fig. 1, all points $(\hbar, f''(0,0))$ with different values of $\hbar$ but the same value of $f''(0,0)$ create a horizontal line segment of the curve $f''(0,0) \sim \hbar$ in a region of $-0.8 \leq \hbar \leq -0.2$, as does the curve $g'(0,0) \sim \hbar$. Therefore, if we choose $-0.8 \leq \hbar \leq -0.2$, we can ensure the convergence of the series of $f''(0,0)$ and $g'(0,0)$. In fact, we indeed get convergent results of $f''(0,0)$ and $g'(0,0)$ by choosing $\hbar = -1/2$. In general, by means of plotting such kinds of $\hbar$ curves, we can always choose a proper value of the auxiliary parameter $\hbar$ to get accurate HAM approximations, as suggested by Liao [7–11]. Similarly, one can investigate the convergence of the series solution at $\xi = 0$ in the whole region $0 \leq \eta < +\infty$. It is found that, when $\xi = 0$ and $\hbar = -1/2$, the solution series of $f(0, \eta)$ and $g(0, \eta)$ converge to the exact initial solutions (7) and (8), as shown in Fig. 2. Furthermore, it is found that when $\hbar = -1/2$, the series solution of $f(\xi, \eta)$ and $g(\xi, \eta)$ are even convergent in the *whole* region $0 \leq \xi \leq 1$ and $0 \leq \eta < +\infty$. When $\lambda = 1/2$, the 15th-order approximations of coefficient of skin friction

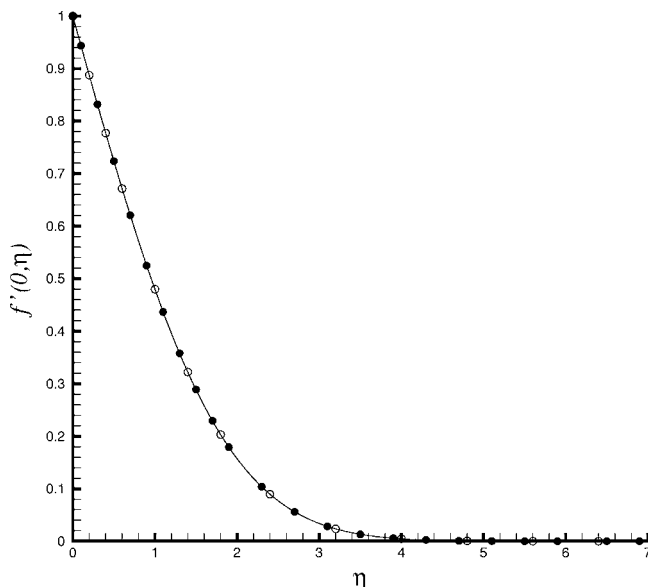


Fig. 2 Comparison of $f'(0, \eta)$ with the exact solution: solid line exact solution; filled circle: 15th-order HAM approximation; open circle: 20th-order HAM approximation

$$C_f^x \text{Re}_x^{1/2} = \xi^{-1/2} (-0.5744017005 - 0.4493808859\xi - 0.06481391332\xi^2 - 0.01742724136\xi^3 - 0.009296051716\xi^4 - 0.005899366598\xi^5 - 0.004065430821\xi^6 - 0.002955448193\xi^7 - 0.001917447834\xi^8 - 0.009525724822\xi^9 + 0.07823876097\xi^{10} - 0.3945931407\xi^{11} + 1.034178447\xi^{12} - 1.478649518\xi^{13} + 1.079883330\xi^{14} - 0.3178853418\xi^{15}) \quad (55)$$

and

$$C_f^y \text{Re}_x^{1/2} = \xi^{-1/2} (-0.5524842929\xi + 0.008804549832\xi^2 + 0.004367680454\xi^3 + 0.002135276175\xi^4 + 0.001719399534\xi^5 + 0.001533164270\xi^6 + 0.001399929708\xi^7 + 0.001031699397\xi^8 + 0.009003153919\xi^9 - 0.09321414938\xi^{10} + 0.5469861363\xi^{11} - 1.668879409\xi^{12} + 2.763217719\xi^{13} - 2.335134057\xi^{14} + 0.7942135326\xi^{15}) \quad (56)$$

agree well with the numerical results for *all* time $\xi \in [0, 1]$, corresponding to $0 \leq \tau < +\infty$, as shown in Fig. 3.

In general, for any given value of λ , we can always choose a proper value of the auxiliary parameter $\hbar$ in a similar way to ensure that the solution series converge for *all* time $0 \leq \tau < +\infty$ in the *whole* spatial region $0 \leq \eta < +\infty$. For example, in cases of $\lambda = 0.5$ and $\lambda = 1$, our HAM results of the coefficient of skin friction $C_f^x \text{Re}_x^{1/2}$ and $C_f^y \text{Re}_x^{1/2}$ agree well with the numerical solutions, as shown in Figs. 4 and 5. As shown in Table 1, the 25th-order HAM approximations have errors less than 0.5%. Note that for a larger value of λ , the perturbation solution becomes worse, and its average error is about 40%. However, by choosing a proper value of the auxiliary parameter $\hbar$, our HAM approximations always converge to the numerical solutions for any values of λ . Thus, the

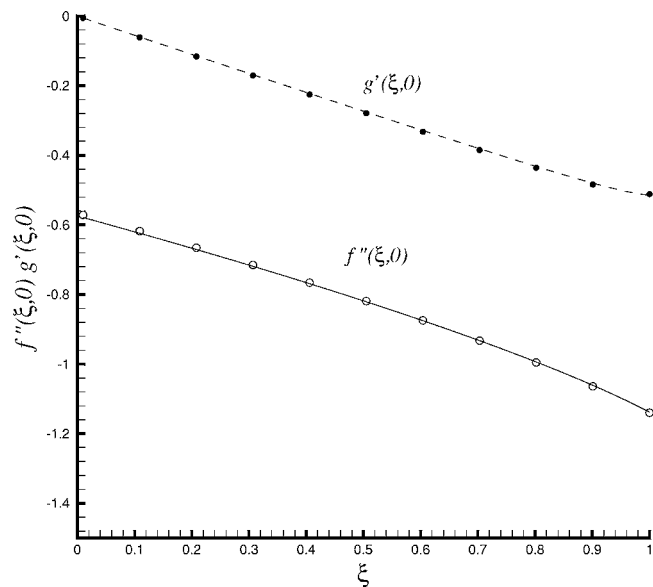


Fig. 3 Comparison of numerical solutions with HAM results of $f'(\xi, 0)$ and $g'(\xi, 0)$ in the case of $\lambda = 1/2$ by means of $\hbar = -1/2$; solid line: numerical result of $f'(\xi, 0)$; open circle: 15th-order HAM approximation of $f'(\xi, 0)$; dash line: numerical result of $g'(\xi, 0)$; filled circle: 15th-order HAM approximation of $g'(\xi, 0)$

homotopy analysis method is indeed a powerful analytic tool for nonlinear problems with strong nonlinearity.

Indeed, the homotopy analysis method provides us with a simple way to ensure the convergence of solution series by means of choosing a proper value of the auxiliary parameter $\hbar$. This is an advantage of the homotopy analysis method over all other analytic techniques. By the way, as proved by Hayat et al. [13] and Sajid et al. [14], and illustrated by Abbasbandy [15,16], the approximations given by the so-called “homotopy perturbation method” [17], which was proposed 7 years later than the homotopy analysis method [6], are only special cases for those given by the homotopy analysis method when $\hbar = -1$. Like other traditional analytic techniques, the “homotopy perturbation method” [17] cannot provide such a simple way to adjust and control the convergence of the solution series [13–16]. For example, in the cases under consideration, the results given by the “homotopy perturbation method” (HPM) are valid only for *small* time, as shown in Figs. 4 and 5.

6 Conclusion

In this paper, the three-dimensional unsteady viscous flows due to the impulsively stretching surface of the incompressible rotating fluid, governed by the Navier–Stokes equations, are solved by means of one analytic technique for strongly nonlinear problems, namely the homotopy analysis method [6–11]. In contrast to the corresponding perturbation approximations which have 40% average errors, our series solutions are uniformly valid for all time $0 \leq \tau < +\infty$ in the whole spatial region $0 \leq \eta < +\infty$, but with only less than 0.5% error, as shown in Table 1. Explicit analytic expressions of coefficients of skin friction are given, which are useful in engineering and for validation of numerical simulations. All of these verify the validity and potential of our approach for complicated viscous flows.

The homotopy analysis method has some advantages over other traditional ones. First, it provides us with great freedom to choose the auxiliary linear operators. Using this kind of freedom, we transfer the original nonlinear, combined initial-boundary-value problem at hand into an infinite number of linear boundary-value subproblems, which are so easy to solve that we can get results at rather high orders of approximations. Second, contrary to all other

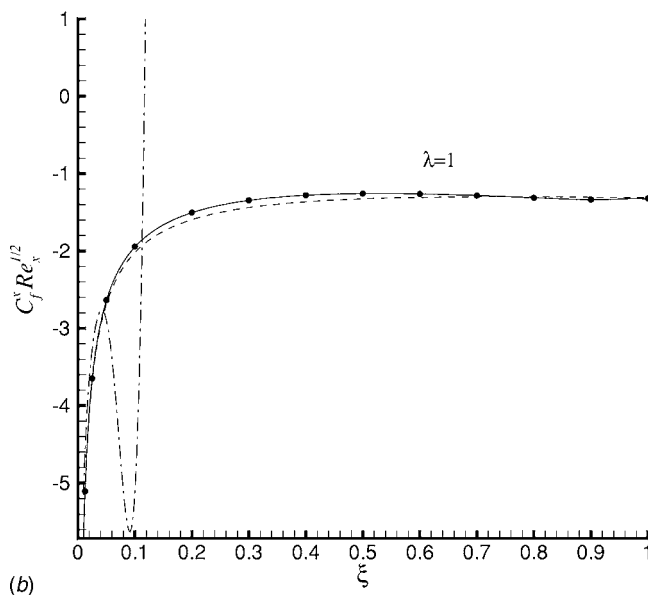
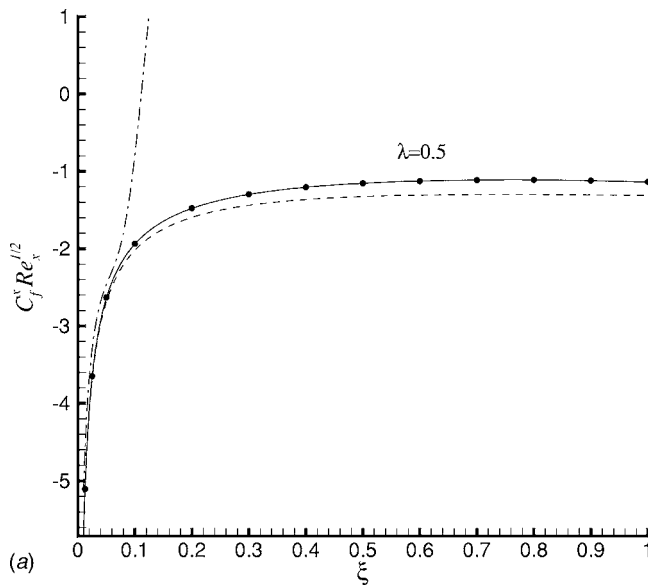


Fig. 4 Comparison of numerical results [4] of $C_f^x \sqrt{Re_x}$ with analytic approximations; symbols: numerical result; solid line: 15th-order HAM approximation: (a) $h=-0.5$; (b) $h=-0.25$; dash line: perturbation approximation given by Nazar et al. [4]; dash-dotted line: 15th-order HPM approximation ($h=-1$)

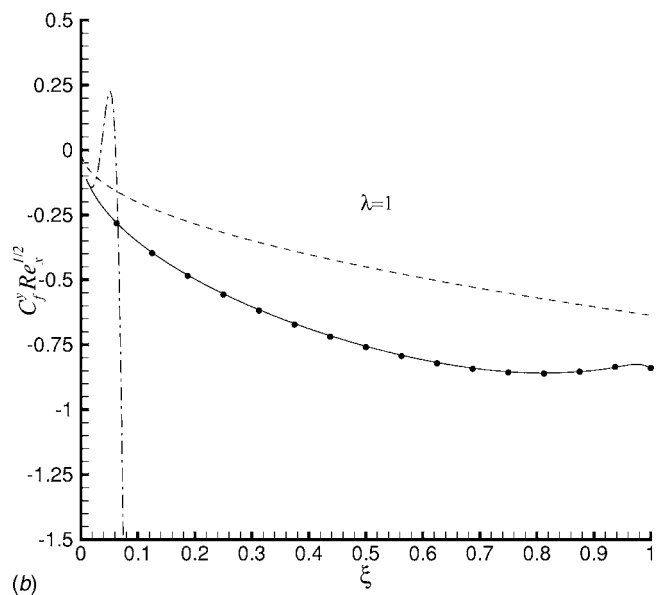
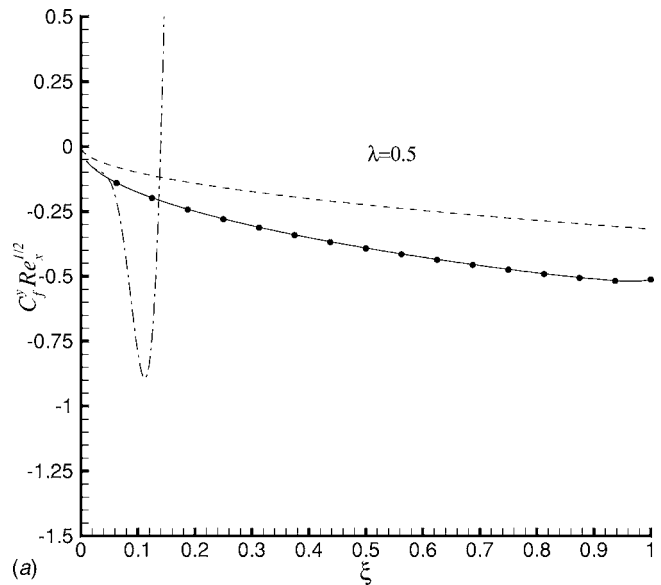


Fig. 5 Comparison of numerical results [4] of $C_f^y \sqrt{Re_x}$ with analytic approximations; symbols: numerical result; solid line: 15th-order HAM approximation (a) $h=-0.5$; (b) $h=-0.25$; dash line: perturbation approximation given by Nazar et al. [4]; dash-dotted line: 15th-order HPM approximation ($h=-1$)

analytic techniques, the homotopy analysis method provides us with a simple way to ensure convergence of solution series. Thus, we can always get accurate enough approximations by means of the homotopy analysis method. Third, in contrast to perturbation techniques, the homotopy analysis method is independent of small/large parameters. Thus, it is suitable for more nonlinear problems. Finally, the homotopy analysis method logically contains other nonperturbation techniques such as the Lyapunov's small parameter method, the δ -expansion method, and the Adomian's decomposition method, as proved by Liao [7]. Currently, Hayat et al. [13], Sajid et al. [14], and Abbasbandy [15,16] pointed out that the so-called "homotopy perturbation method" [17] proposed in 1999 is also a special case of the homotopy analysis method [6,7] propounded in 1992. Thus, the homotopy analysis method is rather general.

There exist numerous three-dimensional unsteady viscous flows and heat transfer problems, which are often rather complicated. The proposed homotopy analysis method provides us with a new

approach to get accurate and convergent series solutions of unsteady three-dimensional Navier–Stokes equations, which are uniformly valid in the whole time $0 \leq \tau < +\infty$.

Table 1 Comparisons of HAM approximations of $f''(1,0)$ and $g'(1,0)$ with numerical results given by Wang [1] and Nazar et al. [4]

λ	$f''(1,0)$			$g'(1,0)$		
	Wang	Nazar et al.	HAM (25th)	Wang	Nazar et al.	HAM (25th)
0.5	-1.138	-1.138	-1.138	-0.513	-0.513	-0.511
1.0	-1.325	-1.325	-1.323	-0.837	-0.837	-0.830

Acknowledgment

Thanks go to the anonymous reviewers for their valuable comments. This work is supported by National Natural Science Foundation of China (Contract No. 10572095).

Nomenclature

C_f^x	= skin friction coefficient in the x direction
C_f^y	= skin friction coefficient in the y direction
f, g	= real functions
F, G	= real functions
$\mathcal{L}_f, \mathcal{L}_g$	= auxiliary linear operators
$\mathcal{N}_f, \mathcal{N}_g$	= nonlinear operators
p	= pressure
q	= embedding parameter
Re_x	= local Reynolds number, $Re_x = ax^2/\nu$
t	= time
u, v, w	= velocities in the x, y, z direction
x, y, z	= spatial coordinates
η	= similarity variable
λ	= non-dimensional parameter defined by $\lambda = \Omega/a$
ξ	= dimensionless time defined by $\xi = 1 - e^{-\tau}$
ρ	= density
τ	= dimensionless time defined by $\tau = at$
ν	= kinematic viscosity
Ω	= angular velocity in the z direction
∇^2	= Laplace operator
$\hbar$	= nonzero auxiliary parameter

References

- [1] Wang, C. Y., 1988, "Stretching a Surface in a Rotating Fluid," *ZAMP*, **39**, pp. 177–185.
- [2] Wang, C. Y., Dut, Q., Miklavcic, M., and Chang, C. C., 1997, "Impulsive Stretching of a Surface in a Viscous Fluid," *SIAM J. Appl. Math.*, **57**, pp. 1–14.
- [3] Seshadri, R., Sreeshylan, N., and Nath, G., 2002, "Unsteady Mixed Convection Flow in the Stagnation Region a Heated Vertical Plate due to Impulsive Motion," *Int. J. Heat Mass Transfer*, **45**, pp. 1345–1352.
- [4] Nazar, N., Amin, N., and Pop, I., 2002, "Unsteady Boundary Layer Flow due to Stretching Surface in a Rotating Fluid," *Mech. Res. Commun.*, **31**, pp. 121–128.
- [5] Cebeci, T., and Bradshaw, P., 1984, *Physical and Computational Aspects of Convective Heat Transfer*, Springer, New York.
- [6] Liao, S. J., 1992, "The Proposed Homotopy Analysis Technique for the Solution of Nonlinear Problems," Ph.D. thesis, Shanghai Jiao Tong University, Shanghai, China.
- [7] Liao, S. J., 2003, *Beyond Perturbation: Introduction to Homotopy Analysis Method*, Chapman & Hall/CRC Press, Boca Raton, FL.
- [8] Liao, S. J., 1999, "An Explicit, Totally Analytic Approximation of Blasius Viscous Flow Problems," *Int. J. Non-Linear Mech.*, **34**(4), pp. 759–778.
- [9] Liao, S. J., 2004, "On the Homotopy Analysis Method for Nonlinear Problems," *Appl. Math. Comput.*, **147**, pp. 499–513.
- [10] Liao, S. J., 2006, "Series Solutions of Unsteady Boundary-Layer Flows Over a Stretching Flat Plate," *Stud. Appl. Math.*, **117**(3), pp. 2529–2539.
- [11] Liao, S. J., and Magyari, E., 2006, "Exponentially Decaying Boundary Layers as Limiting Cases of Families of Algebraically Decaying Ones," *Z. Angew. Math. Phys.*, **57**(5), pp. 777–792.
- [12] Allan, F. M., 2007, "Derivation of the Adomian Decomposition Method Using the Homotopy Analysis Method," *Appl. Math. Comput.*, in press.
- [13] Hayat, T., and Sajid, M., 2007, "On Analytic Solution for Thin Film Flow of a Fourth Grade Fluid Down a Vertical Cylinder," *Phys. Lett. A*, **361**, pp. 316–322.
- [14] Sajid, M., Hayat, T., and Asghar, S., 2007, "Comparison Between the HAM and HPM Solutions of Thin Film Flows of Non-Newtonian Fluids on a Moving Belt," *Nonlinear Dyn.*, in press.
- [15] Abbasbandy, S., 2006, "The Application of the Homotopy Analysis Method to Nonlinear Equations Arising in Heat Transfer," *Phys. Lett. A*, **360**, pp. 109–113.
- [16] Abbasbandy, S., 2007, "The Application of the Homotopy Analysis Method to Solve a Generalized Hirota–Satsuma Coupled KdV Equation," *Phys. Lett. A*, **361**, pp. 478–483.
- [17] He, J. H., 1999, "Homotopy Perturbation Technique," *Comput. Methods Appl. Mech. Eng.*, **178**, pp. 257–262.
- [18] Liao, S. J., 1999, "A Uniformly Valid Analytic Solution of 2D Viscous Flow Past a Semi-infinite Flat Plate," *J. Fluid Mech.*, **385**, pp. 101–128.
- [19] Liao, S. J., and Campo, A., 2002, "Analytic Solutions of the Temperature Distribution in Blasius Viscous Flow Problems," *J. Fluid Mech.*, **453**, pp. 411–425.
- [20] Liao, S. J., 2003, "On the Analytic Solution of Magnetohydrodynamic Flows of Non-Newtonian Fluids Over a Stretching Sheet," *J. Fluid Mech.*, **488**, pp. 189–212.
- [21] Song, Y., Zheng, L. C., and Zhang, X. X., 2006, "On the Homotopy Analysis Method for Solving the Boundary Layer Flow Problem Over a Stretching Surface With Suction and Injection," *Journal of University of Science and Technology Beijing*, **28**, pp. 782–784, in Chinese.
- [22] Shi, Y. R., Xu, J. X., and Wu, Z. X., (2006), "Application of the Homotopy Analysis Method to Solving Nonlinear Evaluation Equations," *Acta Phys. Sin.*, **55**(4), pp. 1555–1560.
- [23] Asghar, S., Gulzar, M. M., and Ayub, M., 2006, "Effect of Partial Slip on Flow of a Third Grade Fluid," *Acta Mech. Sin.*, **22**, pp. 195–198.
- [24] Hayat, T., and Khan, M., 2005, "Homotopy Solutions for a Generalized Second Grade Fluid Past a Porous Plate," *Nonlinear Dyn.*, **42**, pp. 395–405.
- [25] Zhu, S. P., 2006, "A Closed-Form Analytical Solution for the Valuation of Convertible Bonds With Constant Dividend Yield," *ANZIAM J.*, **47**, pp. 477–494.
- [26] Zhu, S. P., 2006, "An Exact and Explicit Solution for the Valuation of American Put Options," *Quant. Finance*, **6**, pp. 229–242.
- [27] Liao, S. J., Su, J., and Chwang, A. T., 2006, "Series Solutions for a Nonlinear Model of Combined Convective and Radiative Cooling of a Spherical Body," *Int. J. Heat Mass Transfer*, **49**, pp. 2437–2445.
- [28] Williams, J. C., and Rhyne, T. H., 1980, "Boundary Layer Development on a Wedge Impulsively Set into Motion," *SIAM J. Appl. Math.*, **38**, pp. 215–224.

Stoneley Wave Generation in Joined Materials With and Without Thermal Relaxation Due to Thermal Mismatch

L. M. Brock

Fellow ASME
Mechanical Engineering,
University of Kentucky,
Lexington, KY 40506
e-mail: brock@engr.uky.edu

Two perfectly bonded, thermoelastic half-spaces differ only in their thermal parameters. Their governing equations include as special cases the Fourier heat conduction model and models with either one or two thermal relaxation times. An exact solution in transform space for the problem of line loads applied to the interface is obtained. Even though the elastic properties of the half-spaces are identical, a Stoneley function arises, and conditions for the existence of roots are more restrictive than for the isothermal case of two elastically dissimilar half-spaces. Moreover, roots may be either real or imaginary. An exact expression for the time transform of the Stoneley residue contribution to interface temperature change is derived. Asymptotic results for the inverse that, valid for either very short or very long times after load application, is obtained and show that, for long times, residue contributions for all three special cases obey Fourier heat conduction. Short-time results are sensitive to case differences. In particular, a time step load produces a propagating step in temperature for the Fourier and double-relaxation time models, but a propagating impulse for the single-relaxation time model.
[DOI: 10.1115/1.2723826]

Introduction

Joined dissimilar elastic materials occur in geological formations [1] and as structural elements [2]. Transient analyses [1,3] show that dynamic loading of these can produce, in addition to dilatational and rotational waves, interface (Stoneley) waves. Such waves are similar to Rayleigh surface waves [4] and so may be important in assessing interface integrity.

The studies [1,3] concern isothermal materials. If, however, coupled thermoelasticity [5] based on Fourier heat flow [6] governs, then material dissimilarity is characterized not only by mass density ρ , shear modulus μ , and Poisson's ratio ν , but also by thermal conductivity k , coefficient of volumetric expansion α_v , and specific heat at constant volume c_v . Indeed, if coupled thermoelasticity allows thermal relaxation [7], then thermal relaxation time(s) τ arise and may also differ. To illustrate the effects of thermal parameters (k, α_v, c_v, τ) this article considers the hypothetical limit of two perfectly bonded materials whose properties (ρ, μ, ν) are identical, i.e., isothermally there is a single material, but whose properties (k, α_v, c_v, τ) differ. For generality, a coupled elasticity formulation that includes Fourier [5] and thermal relaxation [7] models as special cases is employed.

The materials are treated as half-spaces, originally at rest at a uniform temperature and then subjected to thermal-mechanical line loads at the interface. The study begins with construction of an exact solution in transform space. The solution exhibits a Stoneley function that has no counterpart in the isothermal limit. Conditions for the existence of Stoneley roots are determined, and it is found that they can be more restrictive than in the isothermal two-material case. Expressions for the roots, analytic to within a single integration, are developed and found to give both real and imaginary values, again in contrast to the isothermal two-material case.

An exact formula for the time transform of the change in inter-

face temperature is developed. Analytical expressions for the change itself, valid for either very long or very short times after load application are then obtained for each of three models of coupled thermoelasticity. The first is the classical formulation by Chadwick [5] and the second and third are, respectively, single- and double-relaxation formulations due to Lord and Shulman [8] and Green and Lindsay [9]. Consistent with previous observation [10], the long-time results for the Stoneley contribution to interface temperature change all have the character of the Fourier model, and describe a temperature wave. The short-time results, however, are waves that are sensitive to the particular model considered.

Statement of General Problem and Governing Equations

In terms of Cartesian coordinates (x, y, z) two isotropic, homogeneous, linear thermoelastic half-spaces are perfectly bonded along the plane $y=0$. The thermal properties of the half-spaces are dissimilar, but their elastic properties are identical. For time $t \leq 0$ both are at rest at the uniform ambient (absolute) temperature T_0 when, at $t=0$ thermal-mechanical disturbances are introduced at the interface along the line $(x, y)=0$. The disturbances may be time-dependent, but do not vary along the line, so that a state of plane strain is generated. In half-space 1 ($y > 0$) and half-space 2 ($y < 0$) the field equations for $t > 0$ are

$$\left(\nabla^2 - s_r^2 \frac{\partial^2}{\partial t^2} \right) (u_{xn}, u_{yn}) + \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y} \right) (m \Delta_n - \alpha_{vn} D_n^{\text{II}} \theta_n) = 0 \quad (1a)$$

$$h_n \nabla^2 \theta_n - s_r \frac{\partial}{\partial t} \left(\frac{\varepsilon_n}{\alpha_{vn}} D_n \Delta_n - D_n^{\text{I}} \theta_n \right) = 0 \quad (1b)$$

$$\frac{1}{\mu} (\sigma_{xn}, \sigma_{yn}, \sigma_{zn}) = (m-1) \Delta_n - \alpha_{vn} D_n^{\text{II}} \theta_n + 2 \left(\frac{\partial u_{xn}}{\partial x}, \frac{\partial u_{yn}}{\partial y}, 0 \right) \quad (1c)$$

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received September 15, 2006; final manuscript received February 4, 2007. Review conducted by Sridhar Krishnaswamy.

$$\frac{1}{\mu} \sigma_{xy} = \frac{\partial u_{xn}}{\partial y} + \frac{\partial u_{yn}}{\partial x} \quad (1d)$$

In (1) $n=(1,2)$, $(u_{xn}, u_{yn}, \Delta_n, \theta_n)$ are, respectively, displacement components, dilatation and change in temperature from T_0 , and $(\sigma_{xn}, \sigma_{yn}, \sigma_{zn}, \sigma_{xyn})$ are stress components. These vary with (x, y, t) . For the Fourier Model F [5] and single- and double-relaxation time Model I [8] and II [9], respectively,

$$F: (D_n, D_n^I, D_n^{II}) = 1 \quad (2a)$$

$$I: D_n^{II} = 1, (D_n, D_n^I) = 1 + \tau_n^I \frac{\partial}{\partial t} \quad (2b)$$

$$II: (D_n, D_n^{II}) = 1 + \tau_n^{II} \frac{\partial}{\partial t}, D_n^I = 1 + \tau_n^I \frac{\partial}{\partial t} \quad (2c)$$

Here $n=(1,2)$, $\tau_n^I > \tau_n^{II} \geq 0$ are thermal relaxation times, and Model II is seen to introduce thermal relaxation explicitly in constitutive Eq. (1c). In (1)

$$m = \frac{1}{1-2\nu} \quad a = 2 \frac{1-\nu}{1-2\nu} \quad \varepsilon_n = \frac{\mu T_0}{\rho c_{vn}} \alpha_{vn}^2 \quad h_n = v_r \tau_n^h \quad s_r = \frac{1}{v_r} \quad (3a)$$

$$\tau_n^h = \frac{k_n}{\mu c_{vn}} \quad v_r = \sqrt{\frac{\mu}{\rho}} \quad (3b)$$

In (1) and (3), $n=(1,2)$ and $(\varepsilon_n, h_n, s_r, v_r, \tau_n^h)$ are, respectively, the thermal coupling constant, thermoelastic characteristic length, rotational wave slowness, rotational wave speed, and thermoelastic characteristic time. Data [5,11–13] suggests that

$$v_r \approx O(10^3) \text{ m/s} \quad m \geq 2 \quad \varepsilon_n \approx O(10^{-2})$$

$$h_n \approx O(10^{-9}) \text{ m} \quad (\tau_n^I, \tau_n^{II}) \approx O(10^{-13}) \text{ s} \quad (4)$$

These values indicate that $\tau_n^h \gg \tau_n^I > \tau_n^{II}$.

For $y \neq 0$ the initial ($t \leq 0$) conditions are

$$(u_{xn}, u_{yn}, \theta_n) \equiv 0, \quad n = (1,2) \quad (5)$$

For $t > 0$ the interface ($y=0$) conditions are

$$u_{x1} - u_{x2} = 0 \quad u_{y1} - u_{y2} = 0 \quad \theta_1 - \theta_2 = 0 \quad (6a)$$

$$\sigma_{xy1} - \sigma_{xy2} = F_x \delta(x) \quad \sigma_{y1} - \sigma_{y2} = F_y \delta(x)$$

$$k_1 \frac{\partial \theta_1}{\partial y} - k_2 \frac{\partial \theta_2}{\partial y} = F_T \delta(x) \quad (6b)$$

Here (F_x, F_y, F_T) are the time-dependent thermal-mechanical load magnitudes and δ is the Dirac function. For $t > 0$ $(u_{x1}, u_{y1}, \theta_1)$ and $(u_{x2}, u_{y2}, \theta_2)$ should vanish as $y \rightarrow \infty$ and $y \rightarrow -\infty$, respectively, and singular behavior is expected at $(x, y) = 0$.

Transform Solution for the General Problem

Unilateral and bilateral [14] Laplace transforms over (t, x) can be defined as

$$\hat{F}(x) = \int_0^\infty F(x, t) \exp(-pt) dt \quad \tilde{F} = \int_{-\infty}^\infty \hat{F}(x) \exp(-pqx) dq \quad (7)$$

Here p is positive real, and q is imaginary. Decomposition of (1a) and (1b) in view of (5) gives

$$\nabla^2 (a \Delta_n - \alpha_{vn} D_n^{II} \theta_n) - s_r^2 \frac{\partial^2 \Delta_n}{\partial t^2} = 0 \quad \left(\nabla^2 - s_r^2 \frac{\partial^2}{\partial t^2} \right) r_{xyn} = 0 \quad (t > 0) \quad (8a)$$

$$(\Delta_n, \theta_n, r_{xyn}) \equiv 0 \quad (t \leq 0) \quad (8b)$$

Here $n=(1,2)$, $y \neq 0$, and r_{xyn} is rotation in plane strain. Application of (7) to (8) gives eigenfunctions and eigenvalues

$$\exp(\pm p A_n^+ y) \quad \exp(\pm p A_n^- y) \quad \exp(\pm p B y) \quad (9a)$$

$$A_n^+(q^2) = \sqrt{s_n^{+2} - q^2} \quad A_n^-(q^2) = \sqrt{s_n^{-2} - q^2} \quad B(q^2) = \sqrt{s_r^2 - q^2} \quad (9b)$$

In (9) $n=(1,2)$ and the branch points are defined by (3) and

$$s_n^\pm = k_n^\pm s_d \quad s_d = \frac{s_r}{\sqrt{a}}$$

$$2k_n^\pm = \sqrt{\left(1 + \sqrt{\frac{a d_n^I}{\tau_n^h p}}\right)^2 + \frac{\varepsilon_n d_n}{\tau_n^h p}} \pm \sqrt{\left(1 - \sqrt{\frac{a d_n^I}{\tau_n^h p}}\right)^2 + \frac{\varepsilon_n d_n}{\tau_n^h p}} \quad (10)$$

In (10) s_d is the isothermal dilatational wave slowness, and from (2), (5), and (7),

$$F: (d_n, d_n^I) = 1 \quad I: (d_n, d_n^I) = 1 + \tau_n^I p \quad II: (d_n, d_n^{II}) = 1 + \tau_n^{II} p \quad (11)$$

It can be shown in view of (4) for all three models that $k_n^+ > 1 > k_n^- > 0$ and thus $(s_n^+, s_r) > s_n^-$ for all positive real p . The inequality $s_n^+ > s_r$ ($k_n^+ > \sqrt{a}$) also holds when

$$I: p < \frac{m + \varepsilon_n}{m \tau_n^h - (m + \varepsilon_n) \tau_n^I} \quad II: p < \frac{m + \varepsilon_n}{m(\tau_n^h - \tau_n^I) - \varepsilon_n \tau_n^{II}} \quad F: p < 1 + \frac{\varepsilon_n}{m} \quad (12)$$

Application of (7) to (1a) and (1b) in light of (5) and using (10) and (12) gives transforms $(\tilde{u}_{1x}, \tilde{u}_{1y}, \tilde{\theta}_1)$ ($y > 0$) and $(\tilde{u}_{2x}, \tilde{u}_{2y}, \tilde{\theta}_2)$ ($y < 0$) as linear combinations of (9a). Operating on (1c), (1d), (5), and (6) with (7) gives the equations necessary to find the coefficients. For present purposes it is sufficient to display results for half-space 2:

$$\begin{bmatrix} \tilde{u}_{2x} \\ \tilde{u}_{2y} \\ v_r^2 \tilde{\theta}_2 \end{bmatrix} = \begin{bmatrix} q & q & 1 \\ A_2^+ & A_2^- & -q \\ \omega_2 \eta_2^+ & \omega_2 \eta_2^- & 0 \end{bmatrix} \begin{bmatrix} C_+ \exp(p A_2^+ y) \\ C_- \exp(p A_2^- y) \\ C_B \exp(p B y) \end{bmatrix} \quad (13a)$$

$$\begin{bmatrix} C_+ \\ C_- \\ C_B \end{bmatrix} = \frac{1}{\rho p S} \begin{bmatrix} M_+^+ & M_+^- & 0 \\ M_-^+ & M_-^- & 0 \\ 0 & 0 & S \end{bmatrix} \begin{bmatrix} F_+ \\ F_- \\ F_B \end{bmatrix} \quad (13b)$$

For $n=(1,2)$ in view of (11) and (12),

$$\omega_n = \frac{p}{\alpha_{vn} d_n} \quad \eta_n^\pm = 1 - k_n^{\pm 2} \quad (14a)$$

$$\eta_n^+ \eta_n^- = - \frac{\varepsilon_n d_n}{\tau_n^h p}$$

$$\eta_n^- - \eta_n^+ = \eta_n = \sqrt{\left[1 + \frac{1}{\tau_n^h p} (a d_n^I + \varepsilon_n d_n)\right]^2 - 4 \sqrt{\frac{a d_n^I}{\tau_n^h p}}} \quad (14b)$$

For ω_n parameter d_n is defined by

$$I, F: d_n = 1, II: (d_n, d_n^{II}) = 1 + \tau_n^{II} p \quad (15)$$

Equation (11) governs (14b), however. Introduction of branch cuts $\text{Im}(q)=0, |\text{Re}(q)| > s_n^\pm$ and $\text{Im}(q)=0, |\text{Re}(q)| > s_r$ for $(A_n^\pm, B)$, such that $\text{Re}(A_n^\pm, B) \geq 0$ in the cut q -plane guarantees that (13a) is bounded as $y \rightarrow -\infty$ for positive real p . In (13b)

$$F_+ = \omega_1 \eta_1^- (q \hat{F}_x + A_1^- \hat{F}_y) - \frac{v_r^2}{k_1} \hat{F}_T \quad (16a)$$

$$F_- = \omega_1 \eta_1^+ (q \hat{F}_x + A_1^+ \hat{F}_y) - \frac{v_r^2}{k_1} \hat{F}_T \quad (16b)$$

$$F_B = \frac{1}{2} \left(\frac{q}{B} \hat{F}_y - \hat{F}_x \right) \quad (16c)$$

The matrix coefficients in (13a) are defined by

$$M_+^+ = (\omega_1 \eta_1^+ - \omega_2 \eta_2^-) (k_1 A_1^- + k_2 A_2^-) \quad (17a)$$

$$M_-^+ = (\omega_2 \eta_2^- - \omega_1 \eta_1^-) (k_1 A_1^+ + k_2 A_2^+) \quad (17b)$$

$$S = - \left(k_1 \omega_1^2 \frac{\varepsilon_1 d_1}{t_1^h p} + k_2 \omega_2^2 \frac{\varepsilon_2 d_2}{t_2^h p} \right) (A_1^+ - A_1^-) (A_2^+ - A_2^-) \\ + \omega_1 \omega_2 [\eta_1^+ \eta_2^+ (k_1 A_1^+ + k_2 A_2^+) (A_1^- + A_2^-) + \eta_1^- \eta_2^- (k_1 A_1^- + k_2 A_2^-) (A_1^+ + A_2^+)] \\ - \omega_1 \omega_2 [\eta_1^+ \eta_2^- (k_1 A_1^+ + k_2 A_2^-) (A_1^- + A_2^+) + \eta_1^- \eta_2^+ (k_1 A_1^- + k_2 A_2^+) (A_1^+ + A_2^-)] \quad (17c)$$

Equation (11) defines (d_1, d_2) in (17c). Quantity S is the Stoneley function for the present thermoelastic case, and some discussion of its behavior is now given.

Stoneley Function

For positive real p , S has branch cuts $\text{Im}(q)=0, s_* < |\text{Re}(q)| < s^*$, where in view of (11),

$$s_* = \min(s_1^-, s_2^-) \quad s^* = \max(s_1^+, s_2^+) \quad (18)$$

Study of (17c) shows that

$$S(q) \approx - (k_1 + k_2) (\omega_1 \eta_1) (\omega_2 \eta_2) q^2 \quad |q| \rightarrow \infty \quad (19a)$$

$$S(0) = s_d^2 (M_{12} \omega_1 \omega_2 - M_1 \omega_1^2 - M_2 \omega_2^2) \quad (19b)$$

In (19b), (M_1, M_2, M_{12}) are defined by

$$(M_1, M_2) = \left(\frac{k_1 \varepsilon_1 d_1}{t_1^h p}, \frac{k_2 \varepsilon_2 d_2}{t_2^h p} \right) (k_1^+ - k_1^-) (k_2^+ - k_2^-) \quad (20a)$$

$$M_{12} = \eta_1^+ \eta_2^+ (k_1 k_1^+ + k_2 k_2^+) (k_1^- + k_2^-) + \eta_1^- \eta_2^- (k_1 k_1^- + k_2 k_2^-) (k_1^+ + k_2^+) \\ - \eta_1^+ \eta_2^- (k_1 k_1^+ + k_2 k_2^-) (k_1^- + k_2^+) - \eta_1^- \eta_2^+ (k_1 k_1^- + k_2 k_2^+) (k_1^+ + k_2^-) \quad (20b)$$

Equation (11) governs (20a) and in view of (12), quantities $(M_1, M_2, M_{12}) > 0$ for positive real p . As outlined in Appendix A, the sign of $S(0)$ depends on parameter P given by (A3) and the dimensionless ratio ω_1/ω_2 . Use of (17c), (19), and (20), the fact that S has branch cuts $\text{Im}(q)=0, s_* < |\text{Re}(q)| < s^*$, and argument theory [15] in the manner of [16] show that three cases arise,

$$\text{Case A: } S(0) > 0, \quad \frac{S(\pm s^*)}{S(q)} \rightarrow 0 -, \quad \text{Im}(q) = 0, \quad |q| \rightarrow \infty \quad (21a)$$

$$\text{Case B: } S(0) > 0, \quad \frac{S(\pm s^*)}{S(q)} \rightarrow 0 +, \quad \text{Im}(q) = 0, \quad |q| \rightarrow \infty \quad (21b)$$

$$\text{Case C: } S(0) < 0 \quad (21c)$$

For Case A, S exhibits roots $q = \pm s_0, s_0 > 0$. For Case B no roots arise in the cut q -plane. For Case C, S exhibits roots $q = \pm i\tau_0, \tau_0 > 0$.

Following [17,18] an expression for s_0 that is analytic to within a single integration can be obtained. We introduce the function

$$G(q) = \frac{S(q)}{(k_1 + k_2) (\omega_1 \eta_1) (\omega_2 \eta_2) s_0^2 - q^2} \quad (22)$$

It has branch cuts $\text{Im}(q)=0, s_* < |\text{Re}(q)| < s^*$, approaches unity as $|q| \rightarrow \infty$, and has no roots or zeros in the cut q -plane. After Noble [19], it factors as the product of functions $G_\pm$ that are analytic in the overlapping strips $\text{Re}(q) > -s_*$ and $\text{Re}(q) < s_*$, respectively, and are given by

$$\ln G_\pm(q) = \frac{1}{\pi} \int_{s_*}^{s^*} \tan^{-1} \frac{\text{Im } S(u + i0)}{\text{Re } S(u + i0)} \frac{du}{u \pm q} \quad (23)$$

Setting $G = G_+ G_-$ in (22) and evaluating it at $q=0$ gives the formula

$$s_0 = \frac{s_d}{G_\pm(0) \sqrt{\eta_1 \eta_2 (k_1 + k_2)}} \sqrt{M_{12} - M_1 \frac{\omega_1}{\omega_2} - M_2 \frac{\omega_2}{\omega_1}} \quad (24)$$

Replacing s_0^2 by the term $-\tau_0^2$ in (22) gives (23) again, but (24) is replaced by

$$\tau_0 = \frac{s_d}{G_\pm(0) \sqrt{\eta_1 \eta_2 (k_1 + k_2)}} \sqrt{M_1 \frac{\omega_1}{\omega_2} + M_2 \frac{\omega_2}{\omega_1} - M_{12}} \quad (25)$$

It is noted that in (23) both G_+ and G_- are analytic at $q = \pm(s_* - 0)$ and $q = \pm(s^* + 0)$. Thus evaluating (22) at these locations shows by way of a check that $S(0)$ and $S(\pm s_*)$ have the same sign, and that the limit in (21b) occurs whenever $S(0) < 0$. Because Case A and B are analogous to the isothermal problem, the results obtained so far are used to study Stoneley effects in interface temperature for these cases.

Interface Temperature Change

It can then be shown that the temperature change on the interface exhibits, in view of (13a), (16), and (17) the transform

$$(\tilde{\theta}_1, \tilde{\theta}_2) = \tilde{\theta}_{12} = \frac{\omega_1 \omega_2}{pS} \left(q M_x \frac{\hat{F}_x}{\mu p} + M_y \frac{\hat{F}_y}{\mu p} + M_T \frac{k_1 \hat{F}_T}{\rho} \right) \quad (26)$$

Here S is given by (17c) and

$$M_x = \omega_2 k_1 \frac{\varepsilon_1 d_1}{t_1^h} (A_1^+ - A_1^-) (\eta_2^+ - \eta_2^-) + \omega_1 k_2 \frac{\varepsilon_2 d_2}{t_2^h} (A_2^+ - A_2^-) (\eta_1^+ - \eta_1^-) \quad (27a)$$

$$M_y = \omega_2 k_1 \frac{\varepsilon_1 d_1}{t_1^h} (A_1^+ - A_1^-) (\eta_2^- A_2^+ - \eta_2^+ A_2^-) \\ + \omega_1 k_2 \frac{\varepsilon_2 d_2}{t_2^h} (A_2^+ - A_2^-) (\eta_1^- A_1^+ - \eta_1^+ A_1^-) \quad (27b)$$

$$M_T = \eta_1^+ \eta_2^- (A_1^- + A_2^+) + \eta_1^- \eta_2^+ (A_1^+ + A_2^-) - \eta_1^+ \eta_2^+ (A_1^- + A_2^-) \\ - \eta_1^- \eta_2^- (A_1^+ + A_2^+) \quad (27c)$$

Equation (11) governs d_n in (27a) and (27b) for $n=(1,2)$. The inverse of the second operation in (7) can be written as

$$\hat{F}(x) = \frac{p}{2\pi i} \int \tilde{F} \exp(pqx) dq \quad (28)$$

Integration is over a Bromwich contour, which for Case A can be taken as the entire $\text{Im}(q)$ -axis. However, (21a) and (29) show that the integrands that result upon substituting (26) in (28) vanish as

$|q| \rightarrow \infty$ for all $x(M_x, M_y)$ and $x \neq 0$ (M_T) because

$$S \approx O(q^2) \quad M_x \approx O(1/\sqrt{-q^2}) \quad M_y \approx O(1)$$

$$M_T \approx O(\sqrt{-q^2}) \quad |q| \rightarrow \infty \quad (29)$$

The (M_x, M_T) integration can therefore be performed in the Cauchy principal value sense about $\text{Im}(q)=0$, $\text{Re}(q) < -s_*(x>0)$ or $\text{Im}(q)=0$, $\text{Re}(q) > s_*(x<0)$. Similarly, the M_y -term gives an integral around $\text{Im}(q)=0$, $-s^* < \text{Re}(q) < -s_*(x>0)$ or $\text{Im}(q)=0$, $s^* < \text{Re}(q) < s_*(x<0)$ and the pole residue

$$\hat{\theta}_{12}^S = \frac{\hat{F}_y}{2\mu s_0 p} \frac{N_y}{\eta_1 \eta_2 (k_1 + k_2) G_0} \exp(-ps_0|x|) \quad (30a)$$

$$\ln G_0 = \frac{2}{\pi} \int_{s_*}^{s^*} \tan^{-1} \frac{\text{Im } S(u+i0)}{\text{Re } S(u+i0)} \frac{udu}{u^2 - s_0^2} \quad (30b)$$

$$N_y = \omega_2 k_1 \frac{\varepsilon_1 d_1}{\tau_1^h} (\alpha_1^+ - \alpha_1^-) (\eta_2^+ \alpha_2^- - \eta_2^- \alpha_2^+) - \omega_1 k_2 \frac{\varepsilon_2 d_2}{\tau_2^h} (\alpha_2^+ - \alpha_2^-) (\eta_1^+ \alpha_1^- - \eta_1^- \alpha_1^+) \quad (30c)$$

$$\alpha_n^\pm = \sqrt{s_0^2 - s_n^{\pm 2}}, \quad n = (1, 2) \quad (30d)$$

Equation (11) again governs d_n , and (9a), (10), (11), (19), and (30) show that $\hat{\theta}_{12}^S$ appropriately vanishes when thermal properties of the two half-spaces are also identical. For Case B a term such as (30a) does not arise. Inversion of (30a) is now sought for Case A for the three models. To allow more insight into behavior, analytical results are achieved with asymptotic versions of the transforms that are valid for very long or very short times after the line loads are applied.

Inversion for Long Times

A robust asymptotic result for long times, here defined for all three models as

$$t \gg \max(\tau_1^h, \tau_2^h) \quad (31)$$

is obtained by inverting an approximate transform valid for $\max(\tau_1^h p, \tau_2^h p) \ll 1$. It is noted that all D -operators (and thus the corresponding d -factors) become unity, i.e., all three models reduce to the Fourier Model F. For $n=(1, 2)$, (10)–(12) yield

$$k_n^+ \approx \sqrt{\frac{a_n^\varepsilon}{\tau_n^h p}} \quad \eta_n^+ \approx -\frac{a_n^\varepsilon}{\tau_n^h p} \quad k_n^- \approx \sqrt{\frac{a_n^\varepsilon}{a}} \quad \eta_n^- \approx \frac{\varepsilon_n}{a} \quad a_n^\varepsilon = a + \varepsilon_n \quad (32a)$$

$$s_n^+ \approx \lambda_n^\varepsilon \sqrt{\frac{s_r}{p}} \quad s_n^- \approx \frac{s_r}{\sqrt{a_n^\varepsilon}} = s_n^\varepsilon = \frac{1}{v_n^\varepsilon} \quad \omega_n \approx \frac{p}{\alpha_{nv}} \quad \lambda_n^\varepsilon = \sqrt{\frac{a_n^\varepsilon}{ah_n}} \quad (32b)$$

Here $(v_n^\varepsilon, s_n^\varepsilon)$ are the thermoelastic dilatational wave speed and slowness [20]. Equations (10)–(12), (19a), and (32b) and the results of Appendix A indicate that conditions for Case A are always met, so that (26) and (30) give

$$s_0 \approx \frac{\sqrt{k_1 \lambda_1^\varepsilon + k_2 \lambda_2^\varepsilon}}{\sqrt{k_1 + k_2} \sqrt{\lambda_1^\varepsilon + \lambda_2^\varepsilon}} (\sqrt{a_1^\varepsilon} + \sqrt{a_2^\varepsilon}) \frac{s_r}{a} > s_r \quad (33a)$$

$$\theta_{12}^S \approx \frac{N_F}{\mu(k_1 + k_2)} \frac{\sqrt{\alpha_1^0}}{\sqrt{\alpha_1^0 + \alpha_2^0}} F_y(t - s_0|x|) H(t - s_0|x|) \quad (33b)$$

$$N_F = \frac{k_1 \varepsilon_1 a_1^\varepsilon \alpha_1^0}{\tau_1^h \alpha_{v2}} \left(1 - \frac{\alpha_2^0}{s_0}\right) - \frac{k_2 \varepsilon_2 a_2^\varepsilon \alpha_2^0}{\tau_2^h \alpha_{v1}} \left(1 - \frac{\alpha_1^0}{s_0}\right) \quad (33c)$$

$$\alpha_n^0 = \sqrt{s_0^2 - s_n^{\varepsilon 2}}, \quad n = (1, 2) \quad (33d)$$

Here H is the Heaviside function.

Inversion for Short Times: Model F

The short time range for Fourier Model F is defined as

$$t \ll \min(\tau_1^h, \tau_2^h) \quad (34)$$

A robust asymptotic result can therefore be obtained from a transform approximation valid for $\min(\tau_1^h p, \tau_2^h p) \gg 1$. It can be shown for $n=(1, 2)$ that

$$k_n^+ \approx 1 \quad \eta_n^+ \approx \frac{-\varepsilon_n}{\tau_n^h p} \quad k_n^- \approx \sqrt{\frac{a}{\tau_n^h p}} \quad \eta_n^- \approx 1 \quad (35a)$$

$$s_n^+ \approx s_d \quad s_n^- \approx \frac{1}{\sqrt{h_n p}} \quad \omega_n = \frac{p}{\alpha_{vn}} \quad (35b)$$

In light of (12) and (35) materials are chosen so that

$$s_1^- < s_2^- \ll s_d < s_r \quad (36)$$

Appendix A and (19) show that Case A always arises, and (26) and (30) yield

$$s_0 \approx \frac{\sqrt{2} \sqrt{k_1 \sqrt{h_2} + k_2 \sqrt{h_1}}}{\sqrt{k_1 + k_2} \sqrt{\sqrt{h_1} + \sqrt{h_2}}} s_d > s_r \quad (37a)$$

$$\theta_{12}^S \approx \frac{N_F}{\mu \sqrt{2} (k_1 + k_2)} F_y(t - s_0|x|) du H(t - s_0|x|) \quad (37b)$$

$$N_F = \left(\frac{k_1 \varepsilon_1}{\tau_1^h \alpha_{v2}} - \frac{k_2 \varepsilon_2}{\tau_2^h \alpha_{v1}} \right) a \alpha_0 \left(1 - \frac{\alpha_0}{s_0}\right) \quad \alpha_0 = \sqrt{s_0^2 - s_d^2} \quad (37c)$$

Inversion for Short Times: Model 1

For the single-relaxation time model, valid results are obtained for

$$t \ll \min(\tau_1^l, \tau_2^l) \quad (38)$$

with approximate transforms valid for $\max(\tau_1^l p, \tau_2^l p) \gg 1$. Then for $n=(1, 2)$

$$2k_n^\pm \approx \sqrt{(1 + \sqrt{al_n^l})^2 + \varepsilon_n l_n^l} \pm \sqrt{(1 - \sqrt{al_n^l})^2 + \varepsilon_n l_n^l} \quad l_n^l = \frac{\tau_n^l}{\tau_n^h} \ll 1 \quad (39a)$$

$$\omega_n \approx \frac{p}{\alpha_{vn}} \quad \eta_n^+ \eta_n^- \approx -\varepsilon_n \ell_n^l \quad (39b)$$

It is noted that l_n^l is a dimensionless ratio of characteristic times, and the s -quantities are constants, i.e., wave slowness. In light of (12) one can consider the situation

$$s_1^- < s_2^- < s_1^+ < s_2^+ < s_r \quad (40)$$

Use of Appendix A, (19) and (39) shows that Case A arises only if

$$z_- < \frac{\alpha_{v2}}{\alpha_{v1}} < z_+ \quad M_1 < 0 \quad (41)$$

Parameters $z_\pm$ are given by (A5) in Appendix A, with (39) understood and

$$(M_1, M_2) \approx (k_1 \varepsilon_1 \ell_1^l, k_2 \varepsilon_2 \ell_2^l) (k_1^+ - k_1^-) (k_2^+ - k_2^-) \quad (42)$$

Parameter M_1 is defined as

$$M_I = \eta_1^+ \eta_2^-(k_1 \alpha_1^+ + k_2 \alpha_2^-)(\alpha_1^- + \alpha_2^+) + \eta_1^- \eta_2^+(k_1 \alpha_1^- + k_2 \alpha_2^+)(\alpha_1^+ + \alpha_2^-) - \eta_1^+ \eta_2^+(k_1 \alpha_1^+ + k_2 \alpha_2^+)(\alpha_1^- + \alpha_2^-) + \eta_1^- \eta_2^-(k_1 \alpha_1^- + k_2 \alpha_2^-)(\alpha_1^+ + \alpha_2^+) \quad (43)$$

For Case A (24) is valid, with

$$s_* = s_1^- \quad s^* = s_2^+ \quad G_{\pm}(0) \approx \exp \Psi_I(0) \quad (44)$$

Inversion of (30a) then produces

$$\theta_{12}^S \approx \frac{N_I}{\mu s_0 \eta_1 \eta_2 (k_1 + k_2)} \exp(-2\Psi_I(s_0)) \dot{F}_y(t - s_0|x|) H(t - s_0|x|) \quad (45a)$$

$$N_I = \frac{k_1 \varepsilon_1 l_1^I}{\alpha_{v2}} (\alpha_1^+ - \alpha_1^-) (\eta_2^+ \alpha_2^- - \eta_2^- \alpha_2^+) - \frac{k_2 \varepsilon_2 l_2^I}{\alpha_{v1}} (\alpha_2^+ - \alpha_2^-) (\eta_1^+ \alpha_1^- - \eta_1^- \alpha_1^+) \quad (45b)$$

The superposed dot signifies time differentiation, Ψ_I is defined by (B1a) in Appendix B and (39) holds.

Inversion for Short Times: Model 2

For the double-relaxation time model, valid results for

$$t < \min(\tau_1^I, \tau_2^I) \quad (46)$$

are obtained by examining approximate transforms valid for $\min(\tau_1^I p, \tau_2^I p) \gg 1$. For both materials asymptotic results are

$$2k_n^{\pm} \approx \sqrt{(1 + \sqrt{a l_n^I})^2 + \varepsilon_n l_n^I} \pm \sqrt{(1 - \sqrt{a l_n^I})^2 + \varepsilon_n l_n^I} \quad (47a)$$

$$l_n^I = \frac{\tau_n^I}{\tau_n^I} < l_n^I \ll 1$$

$$\omega_n \approx \frac{1}{\alpha_{nv} \tau_n^I} \quad \eta_n^+ \eta_n^- \approx -\varepsilon_n l_n^I \quad (47b)$$

As with Model 1 each s -parameter is wave slowness, and situation (40) can again be considered. Use of Appendix A, (19) and (47) shows that Case A arises only when

$$z_- < \frac{\alpha_{v2} \tau_2^I}{\alpha_{v1} \tau_1^I} < z_+ \quad M_{II} < 0 \quad (48)$$

Again (A5) in Appendix A holds, but now

$$(M_1, M_2) \approx (k_1 \varepsilon_1 l_1^I, k_2 \varepsilon_2 l_2^I) (k_1^+ - k_1^-) (k_2^+ - k_2^-) \quad (49a)$$

$$M_{II} = M_I - (\alpha_1^+ - \alpha_1^-) (\alpha_2^+ - \alpha_2^-) \Omega_{II} \quad \Omega_{II} = k_2 \varepsilon_2 l_2^I \frac{\alpha_{v1}}{\alpha_{v2}} + k_1 \varepsilon_1 l_1^I \frac{\alpha_{v2}}{\alpha_{v1}} \quad (49b)$$

It is understood that (47), not (39), holds for all quantities, including M_I . If (48) is satisfied then (24) holds, with

$$s_* = s_1^- \quad s^* = s_2^+ \quad G_{\pm}(0) \approx \exp \Psi_{II}(0) \quad (50)$$

Inversion of (30a) then gives

$$\theta_{12}^S \approx \frac{N_{II}}{\mu s_0 \eta_1 \eta_2 (k_1 + k_2)} \exp(-2\Psi_{II}(s_0)) F_y(t - s_0|x|) H(t - s_0|x|) \quad (51a)$$

$$N_{II} = \frac{k_1 \varepsilon_1 l_1^I}{\alpha_{v2} \tau_2^I} (\alpha_1^+ - \alpha_1^-) (\eta_2^+ \alpha_2^- - \eta_2^- \alpha_2^+) - \frac{k_2 \varepsilon_2 l_2^I}{\alpha_{v1} \tau_1^I} (\alpha_2^+ - \alpha_2^-) (\eta_1^+ \alpha_1^- - \eta_1^- \alpha_1^+) \quad (51b)$$

Function Ψ_{II} is defined in Appendix C.

Some Observations

Equation (17c) shows that a Stoneley function arises in transform space for a transient study of perfectly bonded thermoelastic half-spaces. Condition (21) for the existence of Stoneley roots is similar to those for the isothermal case, but is also more limiting. Expressions (24) and (25) for the roots, analytic to within a single integration, may depend on the unilateral Laplace (time) transform variable p , i.e., not correspond to, as in the isothermal case, a constant Stoneley wave slowness. Moreover, the root can for positive real p be either real (24) or imaginary (25).

It is found that the line load force that acts normal to the interface produces, from the residue of the real root, contribution (30a) to the time transform of the interface temperature change. The contribution has an analytical form, and asymptotic versions of this, valid for long or short times after the line load is applied, are inverted analytically.

Inversion (33b) shows that the residue contribution behaves for long times as if the half-spaces obey classical Fourier theory [5] even when thermal relaxation [8,9] is present. Conditions for existence of the Stoneley root (in asymptotic form) are always met, and the root given in (33a) is constant, i.e., a Stoneley wave slowness. Equation (33b) for the contribution, therefore, describes a temperature wave.

For short times, (37a), (24), (44), and (50) show that each model can exhibit its own constant real root, so that the temperature wave contribution of the residue for the three models are not identical. Existence conditions for the Fourier model are always met, whereas restriction (41) and (48) hold for, respectively, the single- and double-relaxation time models. Moreover, contributions given in (37b) and (51a) for the Fourier and double-relaxation time models are proportional to the line load function F_y . The contribution given in (45a) for the single-relaxation time model is proportional to the time derivative of F_y .

The observation that $\tau^h \gg \tau^I > \tau^{II}$, made in connection with (4), implies in view of (31), (34), (38), and (46) that asymptotic result (33b) for the Stoneley wave contribution to interface temperature is the most robust. Nevertheless, work [21] in a fluid shows that response for very short times after load application can be distinctive. Indeed, as the comments made above imply, a loading function F_y that is a step (Heaviside) function in time generates for long times a propagating step function that is identical in form for all three models. For short times, however, the Fourier (F) and double-relaxation (II) time model contributions are propagating step functions that differ from each other, and that for the single-relaxation (I) model is a propagating impulse.

In summary, the present results show that Stoneley waves can be generated by thermal mismatch alone, for either coupled thermoelasticity based on Fourier heat flow [5] or thermal relaxation with one [8] or two [9] relaxation times. While the dissimilarity in perfectly bonded half-spaces treated here is hypothetical, the possibility that one or more thermal or mechanical properties of two different materials may match exists [12]. It is hoped in any event that results given here prove useful in the transient study of solids that consist of dissimilar thermoelastic materials. It is noted in closing that the general case of complete mismatch based on the coupled thermoelastic model used here has been treated [22]. The present results serve as a check.

Appendix A

Because (ω_1, ω_2) are positive for positive real p , (19a) is studied in terms of quadratic

$$M_{12}z - M_1 z^2 - M_2 \quad z = \frac{\omega_1}{\omega_2} > 0 \quad (A1)$$

Its discriminant and the location of its maximum value are

$$M_{12}^2 - 4M_1M_2 \quad z = \frac{M_{12}}{2M_1} \quad (M_1, M_2, M_{12}) > 0 \quad (A2)$$

The former can be rewritten in the form

$$(k_1^+ - k_1^-)^2(k_2^+ - k_2^-)^2P \quad P = (C_1 + C_2)^2 - 4C_3 \quad (A3)$$

Factor P is quadratic in (k_1, k_2) ,

$$C_1 = k_1C_{11} + k_2C_{12} \quad C_2 = k_1C_{21} + k_2C_{22} \quad C_3 = k_1k_2C_{33} \quad (A4a)$$

$$C_{11} = (k_1^{+2} + k_1^+k_1^- + k_1^{-2})(1 + k_2^+k_2^-) \quad (A4b)$$

$$C_{22} = (k_2^{+2} + k_2^+k_2^- + k_2^{-2})(1 + k_1^+k_1^-) \quad (A4b)$$

$$(C_{12}, C_{21}) = (k_2^+k_2^-, k_1^+k_1^-)(k_1^+ + k_1^-)(k_2^+ + k_2^-) \quad (A4c)$$

$$C_{33} = (1 - k_1^{+2})(1 - k_1^{-2})(1 - k_2^{+2})(1 - k_2^{-2}) \quad (A4d)$$

In view of (11), (12), and (15) it is seen that $(C_1, C_2, C_3) > 0$, but particular combinations of (k_1, k_2) and $(k_1^\pm, k_2^\pm)$ determine the sign of P and thus, the sign of (A2). It follows that

$$P > 0: S(0) > 0 \left(z_- < \frac{\omega_1}{\omega_2} < z_+ \right), \quad (A5a)$$

$$S(0) < 0 \left(0 < \frac{\omega_1}{\omega_2} < z_-, \frac{\omega_1}{\omega_2} > z_+ \right) \quad (A5a)$$

$$P < 0: S(0) < 0 \quad (A5b)$$

In (A5b) the limit terms

$$z_\pm = \frac{1}{2M_1}(M_{12} \pm \sqrt{M_{12}^2 - 4M_1M_2}) \quad (A6)$$

It can be shown that the sign of the discriminant of P follows from the sign of

$$C_{33} - (C_{11} + C_{21})(C_{22} + C_{12}) \quad (A7)$$

Thus if $(k_1^\pm, k_2^\pm)$ have values for positive real p such that (A7) is negative, then P does not change sign with (k_1, k_2) .

Appendix B

Function Ψ_I that appears in (44) and (45a) is defined by

$$\Psi_I(q) = \frac{1}{\pi} \left(\int_{s_1^-}^{s_2^-} \frac{\psi_1 u du}{u^2 - q^2} + \int_{s_2^-}^{s_1^+} \frac{\psi_2 u du}{u^2 - q^2} + \int_{s_1^+}^{s_2^+} \frac{\psi_3 u du}{u^2 - q^2} \right) \quad (B1a)$$

$$\psi_1 = \tan^{-1} \alpha_1^- \frac{N_1}{D_1} \quad \psi_2 = \tan^{-1} \frac{1}{D_2} (\alpha_1^- N_{21} + \alpha_2^- N_{22})$$

$$\psi_3 = \tan^{-1} A_2^+ \frac{N_3}{D_3} \quad (B1b)$$

In (B1b) the quantities

$$N_1 = k_1[A_1^+(\eta_1^+ - \eta_1^-)(\eta_2^+ - \eta_2^-) + \eta_1^-(\eta_2^- A_2^+ - \eta_2^+ A_2^-)]$$

$$+ k_2 \eta_1^+(\eta_2^+ A_2^+ - \eta_2^- A_2^-) \quad (B2a)$$

$$D_1 = k_1 \eta_1^+ A_1^+ (\eta_2^- A_2^+ - \eta_2^+ A_2^-) + k_2 \eta_2^+ A_2^+ [\eta_1^-(A_1^+ + A_2^-) - \eta_1^+ A_2^-]$$

$$+ k_2 \eta_2^- A_2^- [\eta_1^+ A_2^+ - \eta_1^-(A_1^+ + A_2^-)] \quad (B2b)$$

$$N_{21} = k_1[A_1^+(\eta_1^+ - \eta_1^-)(\eta_2^+ - \eta_2^-) + \eta_1^-(\eta_2^- A_2^+)] + k_2 \eta_1^+ \eta_2^+ A_2^+ \quad (B3a)$$

$$N_{22} = k_2[A_2^+(\eta_1^+ - \eta_1^-)(\eta_2^+ - \eta_2^-) + \eta_1^-(\eta_2^- A_2^+)] + k_1 \eta_1^+ \eta_2^+ A_1^+ \quad (B3b)$$

$$D_2 = A_1^+ A_2^+ (k_1 \eta_1^+ \eta_2^- + k_2 \eta_1^- \eta_2^+) - \alpha_1^- \alpha_2^- (k_1 \eta_1^- \eta_2^+ + k_2 \eta_1^+ \eta_2^-) \quad (B3c)$$

$$N_3 = k_2 \eta_2^+ [\eta_1^+(\alpha_1^- + \alpha_2^-) - \eta_1^-(\alpha_1^+ + \alpha_2^-)] + \eta_2^- [\eta_1^- \alpha_1^+ (k_1 \alpha_1^- + k_2 \alpha_2^-) - \eta_1^+ (k_1 \alpha_1^+ + k_2 \alpha_2^-)] \quad (B4a)$$

$$D_3 = k_1 \eta_2^+ [\eta_1^+ \alpha_1^+ (\alpha_1^- + \alpha_2^-) - \eta_1^- \alpha_1^- (\alpha_1^+ + \alpha_2^-)]$$

$$+ \eta_2^- [\eta_1^- \alpha_1^+ (k_1 \alpha_1^- + k_2 \alpha_2^-) - \eta_1^+ \alpha_1^- (k_1 \alpha_1^+ + k_2 \alpha_2^-)] \quad (B4b)$$

In (B2)–(B4), Eqs. (9b), (14a), and (39) hold with argument u^2 , and

$$\alpha_1^\pm = \sqrt{u^2 - s_1^{\pm 2}} \quad \alpha_2^\pm = \sqrt{u^2 - s_2^{\pm 2}} \quad (B5)$$

Appendix C

The function Ψ_{II} that appears in (50) and (51a) has the same form as function Ψ_I defined by (B1a) in Appendix B. However, (B1b) in Appendix B is modified,

$$\psi_1 = \tan^{-1} \alpha_1^- \frac{N_1 + \Omega_{II}(A_2^+ - A_2^-)}{D_1 + \Omega_{II}A_1^+(A_2^+ - A_2^-)} \quad (C1a)$$

$$\psi_2 = \tan^{-1} \frac{\alpha_1^-(N_{21} + \Omega_{II}A_2^+) + \alpha_2^-(N_{22} + \Omega_{II}A_1^+)}{D_2 + \Omega_{II}(A_1^+A_2^+ - \alpha_1^-\alpha_2^-)} \quad (C1b)$$

$$\psi_3 = \tan^{-1} A_2^+ \frac{N_3 - \Omega_{II}\alpha_2^-(\alpha_1^+ - \alpha_1^-)}{D_3 + \Omega_{II}A_2^+(\alpha_1^+ - \alpha_1^-)} \quad (C1c)$$

In Eqs. (B2)–(B5) in Appendix B, (C1), (9b), and (14a), Eq. (49b) defines Ω_{II} and (47) holds.

References

- [1] Cagniard, L., 1962, *Reflection and Refraction of Progressive Seismic Waves*, E. A. Flinn and C. H. Dix, translators, McGraw-Hill, New York.
- [2] Jones, R. M., 1999, *Mechanics of Composite Materials*, Taylor and Francis, New York.
- [3] Stoneley, R., 1924, "Elastic Waves at the Surface of Separation of Two Solids," *Proc. R. Soc. London, Ser. A*, **106**, pp. 416–428.
- [4] Lamb, H., 1904, "On the Propagation of Tremors over the Surface of an Elastic Solid," *Philos. Trans. R. Soc. London, Ser. A*, **A204**, pp. 1–12.
- [5] Chadwick, P., 1960, "Thermoelasticity, the Dynamical Theory," in *Progress in Solid Mechanics*, I. N. Sneddon and R. Hill, eds., North-Holland, Amsterdam, Vol. 1, pp. 265–230.
- [6] Carrier, G. F., and Pearson, C. E., 1988, *Partial Differential Equations*, Academic, New York.
- [7] Chandrasekharia, D. S., 1986, "Thermoelasticity with Second Sound," *Appl. Mech. Rev.*, **39**, pp. 355–376.
- [8] Lord, H. W., and Shulman, Y., 1967, "Generalized Dynamical Theory of Thermoelasticity," *J. Mech. Phys. Solids*, **15**, pp. 297–309.
- [9] Green, A. E., and Lindsay, K. A., 1972, "Thermoelasticity," *J. Elast.*, **2**, pp. 1–7.
- [10] Brock, L. M., 2004, "Dynamic Contact and Fracture: Some Results for Transversely Isotropic Solids," *Proceedings of the 7th National Congress on Mechanics*, Chania, Greece, A. Kounadis, C. Provodakis, and G. Exadaktylos, eds.
- [11] Achenbach, J. D., 1973, *Wave Propagation in Elastic Solids*, North-Holland, Amsterdam.
- [12] 1998, *Metals Handbook Desk Edition*, J. R. Davis, ed., ASM International, Metals Park, OH.
- [13] Sharma, J. N., and Sharma, P. K., 2002, "Free Vibrations of Homogeneous Transversely Isotropic Thermoelastic Cylindrical Panel," *J. Therm. Stresses*, **25**, pp. 169–182.
- [14] Sneddon, I. N., 1972, *The Use of Integral Transforms*, McGraw-Hill, New York.
- [15] Hille, E., 1959, *Analytic Function Theory*, Ginn, Waltham, MA, Vol. 1.
- [16] Brock, L. M., 1997, "Transient Three-Dimensional Rayleigh and Stoneley Signal Effects in Thermoelastic Solids," *Int. J. Solids Struct.*, **34**, pp. 1463–1478.
- [17] Norris, A. N., and Achenbach, J. D., 1984, "Elastic Wave Diffraction by a Semi-Infinite Crack in a Transversely Isotropic Material," *Q. J. Mech. Appl. Math.*, **37**, pp. 565–580.
- [18] Brock, L. M., 1998, "Analytical Results for Roots of Two Irrational Functions

- in Elastic Wave Propagation," J. Austral. Math. Soc., **B40**, pp. 72–79.
- [19] Noble, B., 1958, *Methods Based on the Wiener-Hopf Technique*, Pergamon, New York.
- [20] Brock, L. M., and Georgiadis, H. G., 1999, "Response of Welded Thermoelastic Solids to the Rapid Motion of Thermomechanical Sources Parallel to the Interface," Int. J. Solids Struct., **36**, pp. 1503–1521.
- [21] Fan, Q.-M., and Lu, W.-Q., 2002, "A New Numerical Method to Stimulate the Non-Fourier Heat Conduction in a Single-Phase Medium," Int. J. Heat Mass Transfer, **45**, pp. 2815–2821.
- [22] Brock, L. M., 2007, "Analysis of Stoneley Signals in Perfectly Bonded Dissimilar Thermoelastic Half-Spaces with and without Thermal Relaxation," J. Mech. Materials Struct. (to appear).

Modal Analysis of Nonviscously Damped Beams

S. Adhikari¹

e-mail: s.adhikari@bristol.ac.uk

M. I. Friswell

Sir George White Professor of Aerospace Engineering

Department of Aerospace Engineering,
Queens Building,
University of Bristol,
Queens Walk,
Bristol BS8 1TR, UK

Y. Lei

Associate Professor
College of Aerospace and Material Engineering,
National University of Defense Technology,
Changsha 410073, PRC

Linear dynamics of Euler–Bernoulli beams with nonviscous non-local damping is considered. It is assumed that the damping force at a given point in the beam depends on the past history of velocities at different points via convolution integrals over exponentially decaying kernel functions. Conventional viscous and viscoelastic damping models can be obtained as special cases of this general damping model. The equation of motion of the beam with such a general damping model results in a linear partial integro-differential equation. Exact closed-form equations of the natural frequencies and mode shapes of the beam are derived. Numerical examples are provided to illustrate the new results.
[DOI: 10.1115/1.2712315]

1 Introduction

Viscous damping is the most common damping model used for linear dynamic systems. However, within the scope of linear theory, more general nonviscous models have been used in the recent past [1–5]. Nonviscous damping models in general have more parameters and therefore are more likely to have a better match with experimental measurements. A linear damped continuous dynamic system in which the displacement variable $u(\mathbf{r}, t)$,

where $\mathbf{r}$ is the spatial position vector and t is time, specified in some domain $\mathcal{D}$, is governed by a linear partial differential equation

$$\rho(\mathbf{r})\ddot{u}(\mathbf{r}, t) + \mathcal{L}_1\dot{u}(\mathbf{r}, t) + \mathcal{L}_2u(\mathbf{r}, t) = p(\mathbf{r}, t); \quad \mathbf{r} \in \mathcal{D}, \quad t \in [0, T] \quad (1)$$

with homogeneous linear boundary conditions of the form

$$\mathcal{M}_1u(\mathbf{r}, t) = 0; \quad \mathbf{r} \in \Gamma_1 \quad \text{and} \quad \mathcal{M}_2\dot{u}(\mathbf{r}, t) = 0; \quad \mathbf{r} \in \Gamma_2 \quad (2)$$

specified on some boundary surfaces Γ_1 and Γ_2 . In the above equation $\rho(\mathbf{r})$ is the mass distribution of the system: $p(\mathbf{r}, t)$ is the distributed time-varying forcing function; and $\mathcal{L}_2$ is the spatial self-adjoint stiffness operator and $\mathcal{M}_1$ and $\mathcal{M}_2$ are linear operators acting on the boundary. For external (or foundation) damping the operator $\mathcal{L}_1$ can be written in the form

$$\mathcal{L}_1\dot{u}(\mathbf{r}, t) = \int_{\mathcal{D}} \int_{-\infty}^t C_1(\mathbf{r}, \xi, t - \tau) \dot{u}(\xi, \tau) d\tau d\xi \quad (3)$$

where $C_1(\mathbf{r}, \xi, t)$ is the kernel function. The velocities $\dot{u}(\xi, \tau)$ at different time instants and spatial locations are coupled through this kernel function. Lei et al. [5] considered both external and internal damping, and although internal damping is not considered further in this paper, the proposed method may be extended to this case. Kernel functions that serve similar purposes have been described by different names in different subjects (for example, retardation functions, heredity functions, after-effect functions, relaxation functions), and different models have been used to describe them. Equation (1) together with Eq. (3) represents a continuous dynamic system with general linear damping. It may be noted that if $\mathcal{L}_1 = 0$ in Eq. (1), i.e., an undamped system, or if the system satisfies the criteria given by Caughey and O’Kelly [6], then the system will possess classical normal modes. However, due to the general nature of the operator $\mathcal{L}_1$ as described by Eq. (3), there is no definite reason why the system should have classical normal modes. Thus the mode shapes and natural frequencies of such systems in general will be complex in nature. In this context we wish to note that the system expressed by Eq. (1) and the damping operator defined in Eq. (3) represents a partial integro-differential equation with the boundary conditions given in Eq. (2). In this technical brief we are interested in the natural frequencies and mode shapes of the system. Exact closed-form expressions of such quantities for the general case are difficult to obtain. We make the following general assumptions:

1. The mass and stiffness distributions are homogeneous, that is, they do not vary with the position vector $\mathbf{r}$; and
2. The damping kernel function is separable in space and time so that

$$C_1(\mathbf{r}, \xi, t - \tau) = C(\mathbf{r})c(\mathbf{r} - \xi)g(t - \tau) \quad (4)$$

Depending on the nature of the functions $c(\bullet)$ and $g(\bullet)$, several special cases, starting from the simple viscous model to the more general nonviscous model, may arise [5]. For example, if both $c(\bullet)$

¹Corresponding author.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received February 8, 2006; final manuscript received August 15, 2006. Review conducted by Oliver M. O’Reilly.



Fig. 1 Euler–Bernoulli beam with nonviscous damping patch

and $g(\bullet)$ are delta functions then the result is the “locally reacting” viscous damping model. If only $c(\bullet)$ is a delta function, then the resulting model becomes a locally reacting viscoelastic damping which is also known as the time hysteresis model. For the case when only $g(\bullet)$ is a delta function, the resulting model becomes nonlocal viscous damping or the spatial hysteresis model. Here the general case, that is, when none of these two functions are the delta functions, is considered. Based on the transfer matrix method [7], we propose a new method for modal analysis of a Euler–Bernoulli beam with general linear damping given by Eq. (4).

2 Governing Equation of Motion

The Euler–Bernoulli beam considered in this study is shown in Fig. 1. The left and right coordinates of the beam are denoted by x_L and x_R , respectively. The beam has a nonlocal viscoelastic damping patch between x_1 and x_2 . It is assumed that the elastic properties of the beam are uniformly distributed with bending rigidity EI and mass density ρA .

In order to formulate and solve the equation of motion, it is necessary to use some kind of plausible functional form of the kernel functions in space and time. In this study we choose the following functional forms [5]

$$g(t) = g_\infty \mu \exp(-\mu t) \quad (5a)$$

so that

$$G(s) = \frac{g_\infty \mu}{s + \mu}, \quad g_\infty, \mu \geq 0 \quad (5b)$$

and

$$c(x - \xi) = \frac{\alpha}{2} \exp(-\alpha|x - \xi|), \quad C(x) = 1 \quad (6)$$

These models imply that the correlations in both space and time decay exponentially. If $\alpha \rightarrow \infty, \mu \rightarrow \infty$ one obtains the standard viscous model; if $\alpha \rightarrow \infty$ and μ is finite one obtains the time hysteresis model; and if α is finite but $\mu \rightarrow \infty$ one obtains the spatial hysteresis model.

The equation of motion of the part with the damping patch can be expressed by

$$\begin{aligned} EI \frac{\partial^4 w(x,t)}{\partial x^4} + \rho A \frac{\partial^2 w(x,t)}{\partial t^2} + \int_{x_1}^{x_2} \int_{-\infty}^t \frac{\alpha}{2} \exp(-\alpha|x - \xi|) \\ \times g_\infty \mu \exp[-\mu(t - \tau)] \frac{\partial w(\xi, \tau)}{\partial t} \bigg|_{\tau=\tau} d\xi d\tau \\ = 0 \quad \text{when } x \in [x_1, x_2] \end{aligned} \quad (7)$$

The equation of motion of the part outside the damping patch can be expressed by

$$\begin{aligned} EI \frac{\partial^4 w(x,t)}{\partial x^4} + \rho A \frac{\partial^2 w(x,t)}{\partial t^2} + C_0 \frac{\partial w(x,t)}{\partial t} \\ = 0 \quad \text{when } x \in (x_L, x_1) \cup (x_2, x_R) \end{aligned} \quad (8)$$

Appropriate boundary conditions must be satisfied at $x=x_L$ and at $x=x_R$. In addition to this we need to satisfy relevant continuity conditions at the internal points x_1 and x_2 . No forcing is assumed because the central interest in this study is to obtain the eigensolutions. The function $w(x,t)$ in Eqs. (7) and (8) is smooth and

continuously differentiable up to fourth order with respect to x . Here we assume that $w(x,t)$ is continuously differentiable up to fifth order. In what follows, Eqs. (7) and (8) are solved separately and the solutions are combined to obtain the eigensolutions. We begin with the solution of Eq. (7).

3 Solution for the Section With Nonlocal Viscoelastic Damping

Assuming zero initial conditions, the Laplace transform of the displacement (with no external force) satisfies

$$\begin{aligned} EI W^{IV}(x,s) + s^2 \rho A W(x,s) + \frac{\alpha}{2} s G(s) \int_{x_1}^{x_2} \exp(-\alpha|x - \xi|) W(\xi,s) d\xi \\ = 0, \quad x \in (x_1, x_2) \end{aligned} \quad (9)$$

Here s is the complex Laplace parameter; and $W(x,s)$ is the Laplace transform of $w(x,t)$. The roman superscripts, for example $(\bullet)^{IV}$, denote the order of derivative with respect to the spatial variable x . It is useful to separate the contribution arising from the term $|x - \xi|$ in Eq. (9) as

$$\begin{aligned} EI W^{IV}(x,s) + s^2 \rho A W(x,s) + \frac{\alpha}{2} s G(s) \int_{x_1}^x \exp[-\alpha(x - \xi)] W(\xi,s) d\xi \\ + \frac{\alpha}{2} s G(s) \int_x^{x_2} \exp[\alpha(x - \xi)] W(\xi,s) d\xi = 0 \end{aligned} \quad (10)$$

The function $W(x,s)$ is continuously differentiable up to fifth order with respect to x because $w(x,t)$ is assumed to be continuously differentiable up to fifth order. Differentiating Eq. (10) with respect to the spatial variable x one obtains

$$\begin{aligned} EI W^V(x,s) + s^2 \rho A W^I(x,s) - \frac{\alpha^2}{2} s G(s) \int_{x_1}^x \exp[-\alpha(x - \xi)] W(\xi,s) d\xi \\ + \frac{\alpha^2}{2} s G(s) \int_x^{x_2} \exp[\alpha(x - \xi)] W(\xi,s) d\xi = 0 \end{aligned} \quad (11)$$

Differentiating again gives

$$\begin{aligned} EI W^{VI}(x,s) + s^2 \rho A W^{II}(x,s) + \frac{\alpha^3}{2} s G(s) \int_{x_1}^x \exp[-\alpha(x - \xi)] W(\xi,s) d\xi \\ + \frac{\alpha^3}{2} s G(s) \int_x^{x_2} \exp[\alpha(x - \xi)] W(\xi,s) d\xi - 2 \frac{\alpha^2}{2} s G(s) W(x,s) = 0 \end{aligned} \quad (12)$$

Using Eqs. (10) and (12) we have

$$\begin{aligned} EI W^{VI}(x,s) + s^2 \rho A W^{II}(x,s) - \alpha^2 [EI W^{IV}(x,s) + s^2 \rho A W(x,s)] \\ - \alpha^2 s G(s) W(x,s) = 0 \end{aligned} \quad (13)$$

Equation (13) is a sixth-order ordinary differential equation which can be solved by transforming into the first-order form. Recall that the solution of a sixth-order ordinary differential equation requires six boundary conditions corresponding to $W(x,s), \dots, W^{VI}(x,s)$. However, the compatibility with the other beam solutions only gives four boundary conditions. The other boundary conditions are implicit in the integro-differential equation. For example, at $x=x_1$, Eq. (10) becomes

$$\begin{aligned} EI W^{IV}(x_1,s) + s^2 \rho A W(x_1,s) + \frac{\alpha}{2} s G(s) \int_{x_1}^{x_2} \exp[\alpha(x - \xi)] W(\xi,s) d\xi \\ = 0 \end{aligned} \quad (14)$$

and Eq. (11) becomes

$$EIW^V(x_1, s) + s^2 \rho A W^I(x_1, s) + \frac{\alpha^2}{2} s G(s) \int_{x_1}^{x_2} \exp[\alpha(x - \xi)] W(\xi, s) d\xi = 0 \quad (15)$$

Notice the integrals are the same in Eqs. (14) and (15) and hence combining these equations gives

$$EIW^V(x_1, s) - \alpha EIW^{IV}(x_1, s) + s^2 \rho A W^I(x_1, s) - \alpha s^2 \rho A W(x_1, s) = 0 \quad (16)$$

A similar procedure for $x=x_2$ gives

$$EIW^V(x_2, s) + \alpha EIW^{IV}(x_2, s) + s^2 \rho A W^I(x_2, s) + \alpha s^2 \rho A W(x_2, s) = 0 \quad (17)$$

To find the eigenvalues using the transfer matrix approach, we need to relate the boundary conditions at the ends of the beam segments. For the beam with the damping layer, we define the state vector $\boldsymbol{\eta}(x, s)$ and partition it as

$$\boldsymbol{\eta}(x, s) = \begin{bmatrix} \mathbf{u}(x, s) \\ \mathbf{v}(x, s) \end{bmatrix} \in \mathbb{C}^6 \quad (18)$$

where

$$\mathbf{u}(x, s) = [W(x, s), W^I(x, s), W^{II}(x, s), W^{III}(x, s)]^T \in \mathbb{C}^4 \quad (19)$$

and

$$\mathbf{v}(x, s) = [W^{IV}(x, s), W^V(x, s)]^T \in \mathbb{C}^2 \quad (20)$$

Using the state vector in Eq. (18), Eq. (13) can be cast in matrix form as

$$\frac{d}{dx} \boldsymbol{\eta}(x, s) = \boldsymbol{\Phi}(s) \boldsymbol{\eta}(x, s), \quad x \in (x_1, x_2) \quad (21)$$

where

$$\boldsymbol{\Phi}(s) = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ \frac{\alpha^2 s G(s) + \alpha^2 s^2 \rho A}{EI} & 0 & -\frac{s^2 \rho A}{EI} & 0 & \alpha^2 & 0 \end{bmatrix} \quad (22)$$

The solution of Eq. (21) can be expressed as

$$\boldsymbol{\eta}(x, s) = \exp[\boldsymbol{\Phi}(s)(x - x_1)] \boldsymbol{\eta}_1(s), \quad x \in (x_1, x_2) \quad (23)$$

where $\boldsymbol{\eta}_1(s) = \boldsymbol{\eta}(x_1, s)$. In particular

$$\boldsymbol{\eta}_2(s) = \exp[\boldsymbol{\Phi}(s)(x_2 - x_1)] \boldsymbol{\eta}_1(s) = \boldsymbol{\Psi}(s) \boldsymbol{\eta}_1(s) \quad (24)$$

where $\boldsymbol{\eta}_2(s) = \boldsymbol{\eta}(x_2, s)$ and $\boldsymbol{\Psi}(s) = \exp[\boldsymbol{\Phi}(s)(x_2 - x_1)]$. We also have from Eqs. (16) and (17) that

$$\mathbf{b}_1(s)^T \boldsymbol{\eta}_1(s) = 0 \quad (25a)$$

and

$$\mathbf{b}_2(s)^T \boldsymbol{\eta}_2(s) = 0 \quad (25b)$$

where

$$\mathbf{b}_1(s) = \begin{bmatrix} -\alpha s^2 \rho A \\ s^2 \rho A \\ 0 \\ 0 \\ -\alpha EI \\ EI \end{bmatrix} \quad (26a)$$

and

$$\mathbf{b}_2(s) = \begin{bmatrix} \alpha s^2 \rho A \\ s^2 \rho A \\ 0 \\ 0 \\ \alpha EI \\ EI \end{bmatrix} \quad (26b)$$

From Eq. (24) and Eq. (25b)

$$\mathbf{b}_2(s)^T \boldsymbol{\eta}_2(s) = \mathbf{b}_2(s)^T \boldsymbol{\Psi}(s) \boldsymbol{\eta}_1(s) = 0 \quad (27)$$

Combining Eqs. (25a) and (27) we have

$$\mathbf{E}(s) \boldsymbol{\eta}_1(s) = \mathbf{0} \quad (28)$$

where

$$\mathbf{E}(s) = \begin{bmatrix} \mathbf{b}_1(s)^T \\ \mathbf{b}_2(s)^T \boldsymbol{\Psi}(s) \end{bmatrix}$$

We partition $\mathbf{E}(s)$ as

$$\mathbf{E}(s) = [\mathbf{E}_1(s) \quad \mathbf{E}_2(s)] \quad (29)$$

where $\mathbf{E}_1(s)$ is 2×4 and $\mathbf{E}_2(s)$ is 2×2 . From Eq. (28) we have

$$\mathbf{E}_1(s) \mathbf{u}(x_1, s) + \mathbf{E}_2(s) \mathbf{v}(x_1, s) = \mathbf{0} \quad (30)$$

or

$$\mathbf{v}(x_1, s) = -\mathbf{E}_2(s)^{-1} \mathbf{E}_1(s) \mathbf{u}(x_1, s) \quad (31)$$

Using Eq. (24) we finally have

$$\mathbf{u}(x_2, s) = \mathbf{T}(s) \mathbf{u}(x_1, s) \quad (32)$$

where

$$\mathbf{T}(s) = [\mathbf{I}_{4 \times 4} \quad \mathbf{0}_{4 \times 2}] \boldsymbol{\Psi}(s) \begin{bmatrix} \mathbf{I}_{4 \times 4} \\ -\mathbf{E}_2(s)^{-1} \mathbf{E}_1(s) \end{bmatrix} \quad (33)$$

4 Eigensolutions of the Complete Beam

Applying the Laplace transform to the equation of motion for the viscously damped segments Eq. (8), and considering the state vector $\mathbf{u}(x, s)$ we can obtain

$$\frac{d}{dx} \mathbf{u}(x, s) = \bar{\boldsymbol{\Phi}}(s) \mathbf{u}(x, s), \quad x \in (x_L, x_1) \cup (x_2, x_R) \quad (34)$$

where

$$\bar{\boldsymbol{\Phi}}(s) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -\frac{s^2 \rho A + s C_0}{EI} & 0 & 0 & 0 \end{bmatrix} \quad (35)$$

The transfer function matrices can be obtained in the usual manner from Eq. (34) as discussed before. The boundary conditions of the beam can be expressed as

$$\mathbf{M}(s) \mathbf{u}(x_L, s) + \mathbf{N}(s) \mathbf{u}(x_R, s) = \mathbf{0} \quad (36)$$

where $\mathbf{M} \in \mathbb{C}^{4 \times 4}$ and $\mathbf{N} \in \mathbb{C}^{4 \times 4}$ are the boundary matrices (see, for example, Ref. [7]). Using the transfer matrices corresponding to the three parts, $\mathbf{u}(x_R, s)$ can be expressed in terms of $\mathbf{u}(x_L, s)$ as

$$\mathbf{u}(x_R, s) = \mathbf{T}_R(s) \mathbf{T}(s) \mathbf{T}_L(s) \mathbf{u}(x_L, s) \quad (37)$$

where $\mathbf{T}(s)$ is defined in Eq. (33) and

$$\mathbf{T}_R(s) = \exp[\bar{\boldsymbol{\Phi}}(s)(x_R - x_2)] \quad (38a)$$

and

Table 1 The first five eigenvalues of the damped pinned-pinned beam

j	$\lambda_j \times 10^{-3}$			
	Proportional viscous	$\mu=100, \alpha=10$	$\mu=100, \alpha=0.1$	$\mu=1, \alpha=0.01$
1	$-0.1161 \pm 0.0725i$	$-0.0604 \pm 0.1369i$	$-0.0214 \pm 0.0702i$	$-0.0214 \pm 0.0698i$
2	$-0.1161 \pm 0.2901i$	$-0.0622 \pm 0.2949i$	$-0.0585 \pm 0.2853i$	$-0.0585 \pm 0.2853i$
3	$-0.1161 \pm 0.6528i$	$-0.0696 \pm 0.6490i$	$-0.0703 \pm 0.6453i$	$-0.0703 \pm 0.6453i$
4	$-0.1161 \pm 1.1606i$	$-0.0574 \pm 1.1575i$	$-0.0576 \pm 1.1550i$	$-0.0576 \pm 1.1550i$
5	$-0.1161 \pm 1.8134i$	$-0.0505 \pm 1.8124i$	$-0.0505 \pm 1.8113i$	$-0.0505 \pm 1.8113i$

$$\mathbf{T}_L(s) = \exp[\bar{\Phi}(s)(x_1 - x_L)] \quad (38b)$$

Substituting Eq. (37) into Eq. (36), one concludes that the eigenvalues are the roots of characteristic equation

$$\det[\mathbf{M}(s) + \mathbf{N}(s)\mathbf{T}_R(s)\mathbf{T}(s)\mathbf{T}_L(s)] = 0 \quad (39)$$

Assuming λ_j are the eigenvalues, the corresponding mode shapes can be determined by

$$\begin{aligned} \psi_j(x) &= \mathbf{u}(x, \lambda_j) \\ &= \begin{cases} \exp[\bar{\Phi}(\lambda_j)(x - x_L)]\mathbf{u}_0(\lambda_j), & x_L \leq x \leq x_1 \\ \mathbf{T}_m(x, \lambda_j)\mathbf{T}_L(\lambda_j)\mathbf{u}_0(\lambda_j), & x_1 \leq x \leq x_2 \\ \exp[\bar{\Phi}(\lambda_j)(x - x_2)]\mathbf{T}(\lambda_j)\mathbf{T}_L(\lambda_j)\mathbf{u}_0(\lambda_j), & x_2 \leq x \leq x_R \end{cases} \end{aligned} \quad (40)$$

Here

$$\mathbf{T}_m(x, \lambda_j) = [\mathbf{I}_{4 \times 4} \quad \mathbf{0}_{4 \times 2}] \exp[\bar{\Phi}(\lambda_j)(x - x_1)] \begin{bmatrix} \mathbf{I}_{4 \times 4} \\ -\mathbf{E}_2(\lambda_j)^{-1}\mathbf{E}_1(\lambda_j) \end{bmatrix} \quad (41)$$

and $\mathbf{u}_0(\lambda_j) \forall j$ is a vector in the null space of the matrix in Eq. (39) evaluated at $s = \lambda_j$.

5 Numerical Example

A damped pinned-pinned beam similar to that shown in Fig. 1 is used to illustrate the proposed method. The numerical values used are as follows: $x_R = 0$ m, $x_L = 1$ m, $x_1 = 0.25$ m, and $x_2 = 0.75$ m, $E = 70$ GN/m², $\rho = 2700$ kg/m³, $C_0 = g_\infty = 15.667$ Ns/m, and the cross section is 5×5 mm². The first five eigenvalues for different values of the relaxation parameter, including the case when the whole beam is uniformly viscously damped with parameter C_0 , is shown in Table 1.

The real parts of the first four modes are shown in Fig. 2 for the four sets of parameter values given in Table 1.

The corresponding imaginary parts of the first four modes are shown in Fig. 3.

Note that, unlike the eigenvalues, μ and α do significantly affect the eigenvectors.

6 Conclusions

The increasing use of advanced composite materials and active control mechanisms demand sophisticated treatment of damping forces within a distributed parameter system. This technical brief proposes a new method to obtain the natural frequencies and mode shapes of Euler-Bernoulli beams with general linear damping models. It was assumed that the damping force at a given point in the beam depends on the past history of velocities at different points via convolution integrals over exponentially decaying kernel functions. Due to the general nature of the damping forces, the equation of motion becomes an integro-differential equation which couples the deflections at different time instants and spatial locations. The conventional transfer matrix method was extended to such integro-differential equations. Commonly used viscous and viscoelastically damped systems can be consid-

ered as special cases of the general formulation derived in the paper. The method was applied to a uniform beam with a nonlocal viscoelastic damping patch. Future work will discuss computational issues and forced vibration problems.

Acknowledgment

Sondipon Adhikari gratefully acknowledges the support of the Engineering and Physical Sciences Research Council through the award of an Advanced Research Fellowship. Michael Friswell gratefully acknowledges the support of the Royal Society through

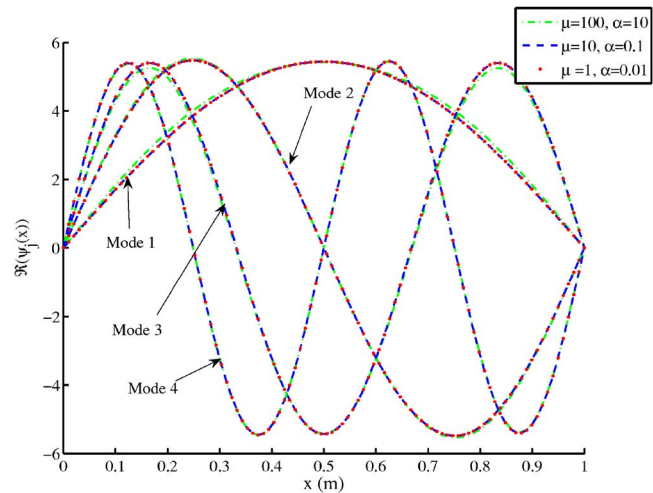


Fig. 2 The real parts of the first four modes

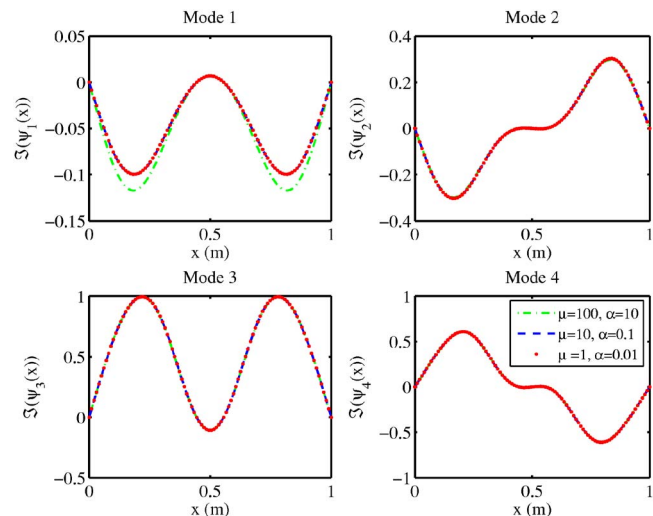


Fig. 3 The imaginary parts of the first four modes

a Royal Society-Wolfson Research Merit Award. Y. Lei gratefully acknowledges the support of China Scholarship Council through a Scholarship Fund Award.

References

- [1] Banks, H. T., and Inman, D. J., 1991, "On Damping Mechanisms in Beams," *ASME J. Appl. Mech.*, **58**, pp. 716–723.
- [2] Woodhouse, J., 1998, "Linear Damping Models for Structural Vibration," *J. Sound Vib.*, **215**(3), pp. 547–569.
- [3] Adhikari, S., 2002, "Dynamics of Non-Viscously Damped Linear Systems," *J. Eng. Mech.*, **128**(3), pp. 328–339.
- [4] Wagner, N., and Adhikari, S., 2003, "Symmetric State-Space Formulation for a Class of Non-viscously Damped Systems," *AIAA J.*, **41**(5), pp. 951–956.
- [5] Lei, Y., Friswell, M. I., and Adhikari, S., 2006, "A Galerkin Method for Distributed Systems with Non-Local Damping," *Int. J. Solids Struct.*, **43**(11–12), pp. 3381–3400.
- [6] Caughey, T. K., and O'Kelly, M. E. J., 1965, "Classical Normal Modes in Damped Linear Dynamic Systems," *ASME J. Appl. Mech.*, **32**, pp. 583–588.
- [7] Yang, B., and Tan, C. A., 1992, "Transfer Functions of One-Dimensional Distributed Parameter Systems," *Trans. ASME, J. Appl. Mech.*, **59**(4), pp. 1009–1014.

Alternative Approaches for the Derivation of Discontinuous Galerkin Methods for Nonlinear Mechanics

L. Noels¹

Department of Aeronautics and Astronautics,
Massachusetts Institute of Technology,
77 Massachusetts Ave.,
Cambridge, MA 02139-4307;
LTAS-Milieux Continus & Thermomécanique,
University of Liège,
Chemin des Chevreuils 1,
B-4000 Liège, Belgium
e-mail: noels@mit.edu, l.noels@ulg.ac.be

R. Radovitzky²

Department of Aeronautics and Astronautics,
Massachusetts Institute of Technology,
77 Massachusetts Ave.,
Cambridge, MA 02139-4307
e-mail: rapa@mit.edu

Discontinuous Galerkin methods are commonly derived by seeking a weak statement of the governing differential equations via a weighted-average approach allowing for discontinuous fields at the element interfaces of the discretization. In order to ensure consistency and stability of the formulation, this approach requires the definition of a numerical flux and a stabilization term. Discontinuous Galerkin methods may also be formulated from a linear combination of the governing and compatibility equations weighted by suitable operators. A third approach based on a variational statement of a generalized energy functional has been proposed recently for finite elasticity. This alternative approach naturally leads to an expression of the numerical flux and the stabilization terms in the context of large deformation mechanics problems. This paper compares these three approaches and establishes the conditions under which identical formulations are obtained. [DOI: 10.1115/1.2712228]

Keywords: discontinuous Galerkin, elliptic equation, hyperelasticity

1 Introduction

In the last several years, discontinuous Galerkin methods have been an increasing focus of attention in the area of computational fluid mechanics [1–3] and, more recently, in solid mechanics [4–8]. The main appeal of the method lies in its ability to enforce compatibility conditions in a weak fashion, while preserving a high order of accuracy.

The most common approach for deriving discontinuous Galerkin methods is to seek a weak statement of the continuum problem via a weighted-average approach, but allowing the test and shape functions to be discontinuous across element interfaces. This leads to undefined interelement terms or numerical fluxes, which are chosen to enforce the consistency of the numerical

scheme. A common problem encountered in the formulation of discontinuous Galerkin methods is the non-uniqueness of the discrete problem solution and the appearance of spurious energy at element interfaces, a problem that was identified in the early contribution by Nitsche [9], who introduced a stabilization term on the boundary to enforce weakly the homogeneous Dirichlet boundary condition. Stabilization is now frequently achieved through a quadratic boundary term. The formulation of alternative numerical fluxes and stabilization terms is subject to the requirements of consistency and stability, but is otherwise unrestricted. This strategy for deriving discontinuous Galerkin formulations has been widely used in fluid mechanics and comprehensively discussed by Cockburn in [3]. Arnold et al. [10] have analyzed the different numerical fluxes and stabilization terms that have been proposed for the Poisson equation. Brezzi et al. [11] have provided a rigorous derivation of a consistent and stable formulation resulting in average numerical fluxes as proposed by Bassi and Rebay [12].

More recently, Brezzi et al. [13] proposed a new framework to derive these terms. The basic idea is to start from a continuous weighted averaging inside each element and to weigh the equations for the jumps at the element interfaces by suitable operators. The choice of these operators is constrained by conditions to ensure uniqueness of the solution and stability of the method. The resulting linear combination of these equations constitutes the weak formulation of the problem.

An alternative strategy for formulating discontinuous Galerkin methods for elliptic problems is to start from an energy functional leading to a variational statement of the problem. This approach was first presented by Lew et al. [5] in the context of linear elasticity and recently extended to nonlinear elasticity in [6,7]. The fact that the numerical fluxes and the stabilization terms arise naturally from the variational approach is particularly advantageous in large-deformation elasticity since it automatically ensures the consistency.

It is thus timely and appropriate to establish the parallel between these alternative formulation strategies and to identify and investigate their limitations and advantages.

2 Variational Approach

The variational formulation of discontinuous Galerkin methods for large deformations closely follows the work presented in [6]. The starting point of this approach is a three-field functional allowing for interelement discontinuities.

Let us consider $B_0 \subset \mathbb{R}^3$ the region of space occupied by a body in its reference configuration. For simplicity, we assume that this body is not subjected to body forces. Its surface boundary ∂B_0 is the union of a part $\partial_D B_0$ constrained by imposed displacements $\bar{\varphi}$ and a part $\partial_N B_0$ subject to surface tractions $\bar{\mathbf{T}}$.

Let $\mathbf{N}$ be the unit surface normal in the reference configuration, $\mathbf{P}$ the first Piola-Kirchhoff stress tensor, ∇_0 the gradient operator with respect to the reference frame, and let φ be the deformation mapping. The continuum problem is governed by the following equations stated in material form

$$\nabla_0 \mathbf{P} = \mathbf{0} \text{ in } B_0 \quad (1)$$

$$\varphi = \bar{\varphi} \text{ on } \partial_D B_0 \quad (2)$$

$$\mathbf{P} \cdot \mathbf{N} = \bar{\mathbf{T}} \text{ on } \partial_N B_0 \quad (3)$$

Let us assume $\mathbf{P} = \partial W / \partial \mathbf{F}$, where $\mathbf{F} = \nabla_0 \varphi$ are the deformation gradients. Therefore, Eqs. (1)–(3) constitute the Euler-Lagrange equations corresponding to the three-field Hu-Washizu-de Veubeke [14–16] energy functional $I(\varphi, \mathbf{F}, \mathbf{P}) : H^1(B_0) \times [H^0(B_0)]^2 \times [H^0(B_0)]^2 \rightarrow \mathbb{R}$

¹Postdoctoral Scholar at the Belgian National Fund for Scientific Research (FNRS).

²Corresponding author.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received June 28, 2006; final manuscript received July 17, 2006. Review conducted by Robert M. McMeeking.

$$I(\boldsymbol{\varphi}, \mathbf{F}, \mathbf{P}) = \int_{B_0} [W(\mathbf{F}) + \mathbf{P} : (\nabla_0 \boldsymbol{\varphi} - \mathbf{F})] dV_0 - \int_{\partial_D B_0} (\boldsymbol{\varphi} - \bar{\boldsymbol{\varphi}}) \cdot \mathbf{P} \cdot \mathbf{N} dS_0 - \int_{\partial_N B_0} \bar{\mathbf{T}} \cdot \boldsymbol{\varphi} dS_0 \quad (4)$$

where $H^1(B_0)$, $[H^0(B_0)]^2$ are the appropriate Sobolev spaces.

Let us consider a finite element discretization $B_{0h} = \cup_{e=1}^E \Omega_0^e$ approximating the reference configuration B_0 . Subscript I denotes the boundary between the elements. Then, $\partial \Omega_0^e = \partial_D \Omega_0^e \cup \partial_N \Omega_0^e \cup \partial_I \Omega_0^e$ and $\partial_I B_{0h} = [\cup_{e=1}^E \partial \Omega_0^e] \setminus \partial B_{0h}$. A finite-dimensional piecewise polynomial approximation $\boldsymbol{\varphi}_h, \mathbf{F}_h, \mathbf{P}_h$ of the solution is defined in the spaces

$$\begin{aligned} X_h^k &= \{ \boldsymbol{\varphi}_h \in L^2(B_{0h}) \mid [\boldsymbol{\varphi}_h]_{\Omega_0^e} \in \mathbb{P}^k(\Omega_0^e) \forall \Omega_0^e \in B_{0h} \} \quad \text{with} \\ X_{hc}^k &= \{ \delta \boldsymbol{\varphi}_h \in X_h^k \mid [\delta \boldsymbol{\varphi}_h]_{\partial_D B_{0h}} = 0 \} \end{aligned} \quad (5)$$

$$\begin{aligned} E_h^k &= \{ \mathbf{F}_h \in [L^2(B_{0h})]^2 \mid [\mathbf{F}_h]_{\Omega_0^e} \in \mathbb{P}^k(\Omega_0^e) \forall \Omega_0^e \in B_{0h} \} \quad \text{with} \\ E_{hc}^k &= \{ \delta \mathbf{F}_h \in E_h^k \mid [\delta \mathbf{F}_h]_{\partial_N B_{0h}} = 0 \} \end{aligned} \quad (6)$$

$$\begin{aligned} S_h^k &= \{ \mathbf{P}_h \in [L^2(B_{0h})]^2 \mid [\mathbf{P}_h]_{\Omega_0^e} \in \mathbb{P}^k(\Omega_0^e) \forall \Omega_0^e \in B_{0h} \} \quad \text{with} \\ S_{hc}^k &= \{ \delta \mathbf{P}_h \in S_h^k \mid [\delta \mathbf{P}_h]_{\partial_N B_{0h}} = 0 \} \end{aligned} \quad (7)$$

where we have also expressed the respective constrained spaces. The discrete approximation $I_h(\boldsymbol{\varphi}_h, \mathbf{F}_h, \mathbf{P}_h)$ of the energy functional (4) may then be derived from the approximate fields (5)–(7) and the domain discretization B_{0h}

$$\begin{aligned} I_h : X_h^k \times E_h^k \times S_h^k &\rightarrow \mathbb{R} : I_h(\boldsymbol{\varphi}_h, \mathbf{F}_h, \mathbf{P}_h) \\ &= \sum_e \int_{\Omega_0^e} W(\mathbf{F}_h) + \mathbf{P}_h : (\nabla_0 \boldsymbol{\varphi}_h - \mathbf{F}_h) dV_0 - \sum_e \int_{\partial_N \Omega_0^e} \bar{\mathbf{T}} \cdot \boldsymbol{\varphi}_h dS_0 \\ &\quad - \sum_e \int_{\partial_D \Omega_0^e} (\boldsymbol{\varphi}_h - \bar{\boldsymbol{\varphi}}_h) \cdot \mathbf{P}_h \cdot \mathbf{N} dS_0 \\ &\quad - \frac{1}{2} \sum_e \int_{\partial_I \Omega_0^e} (\boldsymbol{\varphi}_h - \boldsymbol{\varphi}_h^{\text{ext}}) \cdot \mathbf{P}_h \cdot \mathbf{N} dS_0 \end{aligned} \quad (8)$$

The last term of Eq. (8) enforces weakly the interelement compatibility, where $\boldsymbol{\varphi}_h^{\text{ext}}$ represents the deformation of neighboring elements. The factor $\frac{1}{2}$ has been introduced in order to avoid duplication of the contribution of this term to the total energy. Let us now introduce the jump $[\![\bullet]\!]$ and mean $\langle \bullet \rangle$ operators defined on the space of the trace of functions which can possibly adopt multiple values on the interior boundary $\text{TR}(\partial_I B_{0h}) = \Pi_{e=1}^E [L^2(\partial_I \Omega_0^e)]$

$$\begin{aligned} [\![\bullet]\!], \langle \bullet \rangle : [\text{TR}(\partial_I B_{0h})]^1 \text{ or } 2 &\rightarrow [L^2(\partial_I B_{0h})]^1 \text{ or } 2 : [\![\bullet]\!] \\ &= \bullet^+ - \bullet^-, \langle \bullet \rangle = \frac{1}{2} [\bullet^+ + \bullet^-] \end{aligned} \quad (9)$$

In these expressions, the bullet represents a generic field

$$\bullet^\pm = \lim_{\varepsilon \rightarrow 0_+} \bullet(\mathbf{X} \pm \varepsilon \mathbf{N}^-) \quad \forall \mathbf{X} \in \partial_I B_{0h} \quad (10)$$

and $\mathbf{N}^-$ is conventionally defined as the reference outward unit normal of $\partial \Omega_0^e$. Equation (8) can then be rewritten as

$$\begin{aligned} I_h : X_h^k \times E_h^k \times S_h^k &\rightarrow \mathbb{R} : I_h(\boldsymbol{\varphi}_h, \mathbf{F}_h, \mathbf{P}_h) \\ &= \int_{B_{0h}} [W(\mathbf{F}_h) + \mathbf{P}_h : (\nabla_0 \boldsymbol{\varphi}_h - \mathbf{F}_h)] dV_0 - \int_{\partial_N B_{0h}} \bar{\mathbf{T}} \cdot \boldsymbol{\varphi}_h dS_0 \\ &\quad - \int_{\partial_D B_{0h}} (\boldsymbol{\varphi}_h - \bar{\boldsymbol{\varphi}}_h) \cdot \mathbf{P}_h \cdot \mathbf{N} dS_0 + \int_{\partial_I B_{0h}} [\![\boldsymbol{\varphi}_h]\!] \cdot \langle \mathbf{P}_h \rangle \cdot \mathbf{N}^- dS_0 \end{aligned} \quad (11)$$

The weak formulation of the equations of equilibrium, compatibility, and constitutive behavior are obtained by taking the first variations of I_h with respect to its arguments

$$\begin{aligned} 0 &= \int_{B_{0h}} \mathbf{P}_h : \nabla_0 \delta \boldsymbol{\varphi}_h dV_0 - \int_{\partial_N B_{0h}} \bar{\mathbf{T}} \cdot \delta \boldsymbol{\varphi}_h dS_0 \\ &\quad + \int_{\partial_I B_{0h}} [\![\delta \boldsymbol{\varphi}_h]\!] \cdot \langle \mathbf{P}_h \rangle \cdot \mathbf{N}^- dS_0 \quad \forall \delta \boldsymbol{\varphi}_h \in X_{hc}^k \end{aligned} \quad (12)$$

$$\begin{aligned} 0 &= \int_{B_{0h}} \delta \mathbf{P}_h : (\nabla_0 \boldsymbol{\varphi}_h - \mathbf{F}_h) dV_0 - \int_{\partial_D B_{0h}} (\boldsymbol{\varphi}_h - \bar{\boldsymbol{\varphi}}_h) \cdot \delta \mathbf{P}_h \cdot \mathbf{N} dS_0 \\ &\quad + \int_{\partial_I B_{0h}} [\![\boldsymbol{\varphi}_h]\!] \cdot \langle \delta \mathbf{P}_h \rangle \cdot \mathbf{N}^- dS_0 \quad \forall \delta \mathbf{P}_h \in S_{hc}^k \end{aligned} \quad (13)$$

$$0 = \int_{B_{0h}} \left(2\mathbf{F}_h \frac{\partial W(\mathbf{C}_h)}{\partial \mathbf{C}} - \mathbf{P}_h \right) : \delta \mathbf{F}_h dV_0 \quad \forall \delta \mathbf{F}_h \in E_{hc}^k \quad (14)$$

where $2\partial W(\mathbf{C})/\partial \mathbf{C} = \mathbf{S}$ is the second Piola-Kirchhoff stress tensor and $\mathbf{C} = \mathbf{F}^T \mathbf{F}$ is the right Cauchy-Green deformation tensor. The constitutive law (14) may clearly be enforced strongly, leading to the definition $\mathbf{P}(\mathbf{F}_h) = \mathbf{F}_h \mathbf{S}(\mathbf{C}_h) = \mathbf{F}_h 2\partial W(\mathbf{C}_h)/\partial \mathbf{C}$. An assumed form of the deformation gradients is adopted as proposed by Brezzi et al. [11] (see also [5])

$$\mathbf{F}_h = \nabla_0 \boldsymbol{\varphi}_h + \mathbf{R}_{\bar{\boldsymbol{\varphi}}_h}([\![\boldsymbol{\varphi}_h]\!]) \quad (15)$$

where the tensorial operator $\mathbf{R}_{\bar{\boldsymbol{\varphi}}_h} : L^2(\partial_I B_{0h}) \rightarrow E_h^k$ is defined as

$$\begin{aligned} \int_{B_{0h}} \mathbf{R}_{\bar{\boldsymbol{\varphi}}_h}([\![\mathbf{v}]\!]) : \boldsymbol{\tau} dV_0 &= \int_{\partial_I B_{0h}} [\![\mathbf{v}]\!] \cdot \langle \boldsymbol{\tau} \rangle \cdot \mathbf{N}^- dS_0 \\ &\quad + \int_{\partial_D B_{0h}} [\bar{\boldsymbol{\varphi}}_h - \mathbf{v}] \cdot \boldsymbol{\tau} \cdot \mathbf{N} dS_0 \quad \forall \boldsymbol{\tau} \in S_h^k, \end{aligned} \quad (16)$$

With this definition of the discrete deformation gradients the compatibility condition (13) is satisfied automatically. The non-local character of the average stresses $\langle \mathbf{P}(\mathbf{F}_h) \rangle$ resulting from $\mathbf{F}_h$, may be circumvented by following Brezzi et al. [11], who proposed the use of the lifting operator $\mathbf{r}_{s\bar{\boldsymbol{\varphi}}_h} : L^2(\partial_I B_{0h}) \rightarrow E_h^k$

$$\begin{aligned} \int_{B_{0h}} \mathbf{r}_{s\bar{\boldsymbol{\varphi}}_h}([\![\mathbf{v}]\!]) : \boldsymbol{\tau} dV_0 &= \begin{cases} \int_s [\![\mathbf{v}]\!] \cdot \langle \boldsymbol{\tau} \rangle \cdot \mathbf{N}^- dS_0 & \forall \boldsymbol{\tau} \in S_h^k \text{ and } \forall s \in \partial_I B_{0h} \\ \int_s [\bar{\boldsymbol{\varphi}}_h \cdot \boldsymbol{\tau} \cdot \mathbf{N} - \mathbf{v} \cdot \boldsymbol{\tau} \cdot \mathbf{N}] dS_0 & \forall \boldsymbol{\tau} \in S_h^k \text{ and } \forall s \in \partial_D B_{0h} \\ 0 & \forall \boldsymbol{\tau} \in S_h^k \text{ and } \forall s \in \partial_N B_{0h} \end{cases} \end{aligned} \quad (17)$$

where “s” is an arbitrary element side. In [6], the following form of the deformation mapping was proposed:

$$\mathbf{F}_h = \nabla_0 \boldsymbol{\varphi}_h + \sum_s^{N_s} \mathbf{r}_{s\bar{\varphi}_h}(\llbracket \boldsymbol{\varphi}_h \rrbracket) \quad \text{and} \quad \mathbf{C}_h = \mathbf{F}_h^T \mathbf{F}_h \quad \text{in } \Omega_0^e \quad (18)$$

$$\mathbf{F}_s = \nabla_0 \boldsymbol{\varphi}_h + \beta \mathbf{r}_{s\bar{\varphi}_h}(\llbracket \boldsymbol{\varphi}_h \rrbracket) \quad \text{and} \quad \mathbf{C}_s = \mathbf{F}_s^T \mathbf{F}_s \quad \text{on } s \in \partial\Omega_0^e \quad (19)$$

where N_s is the number of element sides and β is a parameter responsible of enforcing the stability of the numerical scheme. Here and subsequently, it will be assumed without loss of generality that Dirichlet boundary conditions are enforced strongly as in conventional finite element approaches, i.e.

$$\boldsymbol{\varphi}_h = \bar{\boldsymbol{\varphi}} \quad \forall \mathbf{X} \in \partial_D B_{0h} \quad \text{which implies} \quad \mathbf{r}_{s\bar{\varphi}_h} = 0 \quad \forall s \in \partial_D B_{0h} \quad (20)$$

as the main interest is in the treatment of interelement discontinuities.

Following the approach presented in [6], it is assumed that the displacement jumps at element interfaces are small compared to the deformation and the evaluations of the constitutive rules $\mathbf{P}(\mathbf{F}_h)$ and $\mathbf{P}(\mathbf{F}_s)$ are linearized:

$$\begin{aligned} \mathbf{P}(\mathbf{F}_h) &\approx \mathbf{P}(\nabla_0 \boldsymbol{\varphi}_h) + \mathbf{C}(\nabla_0 \boldsymbol{\varphi}_h) : (\mathbf{F}_h - \nabla_0 \boldsymbol{\varphi}_h) \\ \mathbf{P}(\mathbf{F}_s) &\approx \mathbf{P}(\nabla_0 \boldsymbol{\varphi}_h) + \mathbf{C}(\nabla_0 \boldsymbol{\varphi}_h) : (\mathbf{F}_s - \nabla_0 \boldsymbol{\varphi}_h) \end{aligned} \quad (21)$$

where $\mathbf{C} = \partial \mathbf{P} / \partial \mathbf{F}$ is the Lagrangian tangent modulus. It should be carefully noted that the linearization is done with respect to the displacement jumps and not the bulk deformation gradients. Therefore, the nonlinear response is unaffected [6]. Using Eqs. (17) and (21), and denoting $\bar{\mathbf{F}}_h = \nabla_0 \boldsymbol{\varphi}_h$ and all tensorial quantities evaluated at $\bar{\mathbf{F}}_h$ with the convention $\bar{\bullet} = \bullet(\bar{\mathbf{F}}_h)$, the volume integral in Eq. (12) can be rewritten as

$$\begin{aligned} \int_{B_{0h}} \mathbf{P}(\mathbf{F}_h) : \nabla_0 \delta \boldsymbol{\varphi}_h dV_0 &= \int_{B_{0h}} \bar{\mathbf{P}} : \nabla_0 \delta \boldsymbol{\varphi}_h dV_0 \\ &+ \int_{\partial_I B_{0h}} \llbracket \boldsymbol{\varphi}_h \rrbracket \cdot \langle \bar{\mathbf{C}} : \nabla_0 \delta \boldsymbol{\varphi}_h \rangle \cdot \mathbf{N}^- dS_0 \end{aligned} \quad (22)$$

while the term concerning the internal boundary can be rewritten as

$$\begin{aligned} \sum_s \int_{s \in \partial_I B_{0h}} \llbracket \delta \boldsymbol{\varphi}_h \rrbracket \cdot \langle \mathbf{P}(\mathbf{F}_s) \rangle \cdot \mathbf{N}^- dS_0 \\ = \int_{\partial_I B_{0h}} \llbracket \delta \boldsymbol{\varphi}_h \rrbracket \cdot \langle \bar{\mathbf{P}} \rangle \cdot \mathbf{N}^- dS_0 \\ + \sum_s \int_{s \in \partial_I B_{0h}} \llbracket \delta \boldsymbol{\varphi}_h \rrbracket \cdot \langle \beta \bar{\mathbf{C}} : \mathbf{r}_{s\bar{\varphi}_h}(\llbracket \boldsymbol{\varphi}_h \rrbracket) \rangle \cdot \mathbf{N}^- dS_0 \end{aligned} \quad (23)$$

By virtue of Eq. (17), the last term in (23) can be rewritten as

$$\begin{aligned} \sum_s \int_{s \in \partial_I B_{0h}} \llbracket \delta \boldsymbol{\varphi}_h \rrbracket \cdot \langle \beta \bar{\mathbf{C}} : \mathbf{r}_{s\bar{\varphi}_h}(\llbracket \boldsymbol{\varphi}_h \rrbracket) \rangle \cdot \mathbf{N}^- dS_0 \\ = \sum_s \int_{B_{0h}} \mathbf{r}_{s\bar{\varphi}_h}(\llbracket \delta \boldsymbol{\varphi}_h \rrbracket) : \beta \bar{\mathbf{C}} : \mathbf{r}_{s\bar{\varphi}_h}(\llbracket \boldsymbol{\varphi}_h \rrbracket) dV_0 \\ = \sum_s \int_{s \in \partial_I B_{0h}} \llbracket \boldsymbol{\varphi}_h \rrbracket \cdot \langle \beta \bar{\mathbf{C}} : \mathbf{r}_{s\bar{\varphi}_h}(\llbracket \delta \boldsymbol{\varphi}_h \rrbracket) \rangle \cdot \mathbf{N}^- dS_0 \end{aligned} \quad (24)$$

which exposes the symmetry of the stabilization term with respect to $\llbracket \boldsymbol{\varphi}_h \rrbracket$ and $\llbracket \delta \boldsymbol{\varphi}_h \rrbracket$. Recalling that $\mathbf{r}_{s\bar{\varphi}_h}$ is dimensionless, on the boundary we can assume that

$$\langle \beta \bar{\mathbf{C}} : \mathbf{r}_{s\bar{\varphi}_h}(\llbracket \boldsymbol{\varphi}_h \rrbracket) \rangle = \mathbf{N}^- \cdot \langle \beta \bar{\mathbf{C}} \rangle \cdot \frac{\llbracket \boldsymbol{\varphi}_h \rrbracket}{h_s} \quad \forall \mathbf{X} \in \partial_I B_{0h} \quad (25)$$

where h_s is a characteristic length of the mesh and where β includes a multiplicative constant. Using this last result, the final stabilized weak formulation is obtained from Eqs. (12), (22), and (23) and consists in finding $\boldsymbol{\varphi} \in X_h^k$ such that $\forall \delta \boldsymbol{\varphi}_h \in X_{h_c}^k$

$$\begin{aligned} 0 &= \int_{B_{0h}} \bar{\mathbf{P}} : \nabla_0 \delta \boldsymbol{\varphi}_h dV_0 + \int_{\partial_I B_{0h}} \llbracket \boldsymbol{\varphi}_h \rrbracket \cdot \langle \bar{\mathbf{C}} : \nabla_0 \delta \boldsymbol{\varphi}_h \rangle \cdot \mathbf{N}^- dS_0 \\ &- \int_{\partial_N B_{0h}} \bar{\mathbf{T}} \cdot \delta \boldsymbol{\varphi}_h dS_0 + \int_{\partial_I B_{0h}} \llbracket \delta \boldsymbol{\varphi}_h \rrbracket \cdot \langle \bar{\mathbf{P}} \rangle \cdot \mathbf{N}^- dS_0 \\ &+ \sum_s \int_{s \in \partial_I B_{0h}} \llbracket \delta \boldsymbol{\varphi}_h \rrbracket \otimes \mathbf{N}^- \cdot \left\langle \frac{\beta \bar{\mathbf{C}}}{h_s} \right\rangle \cdot \llbracket \boldsymbol{\varphi}_h \rrbracket \otimes \mathbf{N}^- dS_0 \end{aligned} \quad (26)$$

The consistency and stability of this formulation were demonstrated theoretically and verified numerically in [6]. It was found that the symmetrization term $\int_{\partial_I B_{0h}} \llbracket \boldsymbol{\varphi}_h \rrbracket \cdot \langle \bar{\mathbf{C}} : \nabla_0 \delta \boldsymbol{\varphi}_h \rangle \cdot \mathbf{N}^- dS_0$ does not contribute to ensuring stability and, thus, can be neglected for ease of implementation. Numerical stability is guaranteed provided that $\beta \geq C^k$, where $C^k > 0$ is a constant depending only on the polynomial degree k of the approximate solution. The convergence rate of the error with the mesh size is of the order k .

An important aspect of the variational approach for formulating discontinuous Galerkin methods, is that the numerical flux and the stabilization term arise naturally from the variational principle. The particular choice of functional in the formulation proposed has led to the same average numerical fluxes as the ones proposed by Bassi and Rebay [12], see also Lew et al. [5].

In the next section we investigate under what conditions the same formulation can be obtained via a weak statement of the differential equations governing the large deformations of solids.

3 Approach Based on a Weak Statement of the Continuum Equations

The formulation of discontinuous Galerkin methods for nonlinear mechanics may also be based on a weak statement of the continuum equations by the standard approach of multiplying the equilibrium equations (1) by a suitable test function followed by integration in the domain. However, in the discontinuous Galerkin case, the test functions are allowed to be discontinuous across domain boundaries and the integration is carried out in each sub-domain Ω_0^e in a piecewise fashion. The weak solution is obtained by finding $\mathbf{P}_h \in S_h^k$ such that

$$\begin{aligned} \sum_e \int_{\Omega_0^e} \nabla_0 \cdot \mathbf{P}_h \cdot \delta \boldsymbol{\varphi}_h dV_0 \\ = \sum_e \int_{\Omega_0^e} -\mathbf{P}_h : \nabla_0 \delta \boldsymbol{\varphi}_h dV_0 + \sum_e \int_{\partial\Omega_0^e} \delta \boldsymbol{\varphi}_h \cdot \mathbf{P}_h \cdot \mathbf{N} dS_0 \\ = \sum_e \int_{\Omega_0^e} -\mathbf{P}_h : \nabla_0 \delta \boldsymbol{\varphi}_h dV_0 \\ + \sum_e \int_{\partial\Omega_0^e \cap \partial_I B_{0h}} \delta \boldsymbol{\varphi}_h \cdot \bar{\mathbf{T}} dS_0 \\ + \sum_e \int_{\partial\Omega_0^e \cap \partial_I B_{0h}} \delta \boldsymbol{\varphi}_h \cdot \mathbf{P}_h \cdot \mathbf{N} dS_0 \\ = 0 \quad \forall \delta \boldsymbol{\varphi} \in X_{h_c}^k \end{aligned} \quad (27)$$

where we have used the divergence theorem, the equilibrium of the traction on the Neumann boundary (3) and the definition of the

constrained space X_{hc}^k on the Dirichlet boundary (7).

The defining characteristic of discontinuous Galerkin methods is that both the test functions $\delta\varphi_h$ and the stress tensor $\mathbf{P}_h$ are allowed to have finite jumps on the interelement boundary $\partial_l B_{0h}$. The weak enforcement of equilibrium and compatibility on $\partial_l B_{0h}$ is achieved by introducing a numerical flux $\mathbf{h}(\mathbf{P}_h^+, \mathbf{P}_h^-, \mathbf{N}^-)$ dependent on the limit values on the surface of the neighboring elements, such that

$$\sum_e \int_{\partial\Omega_0^e \cap \partial_l B_{0h}} \delta\varphi_h \cdot \mathbf{P}_h \cdot \mathbf{N} dS_0 \rightarrow - \int_{\partial_l B_{0h}} \llbracket \delta\varphi_h \rrbracket \cdot \mathbf{h}(\mathbf{P}_h^-, \mathbf{P}_h^+, \mathbf{N}^-) dS_0 \quad (28)$$

Although there is, in principle, significant freedom in the choice of $\mathbf{h}$, consistent formulations require

$$\mathbf{h}(\mathbf{P}, \mathbf{P}, \mathbf{N}) = \mathbf{P} \cdot \mathbf{N} \quad \text{and} \quad \mathbf{h}(\mathbf{P}_h^-, \mathbf{P}_h^+, \mathbf{N}^-) = -\mathbf{h}(\mathbf{P}_h^+, \mathbf{P}_h^-, \mathbf{N}^+) \quad (29)$$

where $\mathbf{P}$ is the exact solution. The interelement boundary term may be rewritten using (9) and (10) as

$$\begin{aligned} \sum_e \int_{\partial\Omega_0^e \cap \partial_l B_{0h}} \delta\varphi_h \cdot \mathbf{P}_h \cdot \mathbf{N} dS_0 \\ = - \int_{\partial_l B_{0h}} \llbracket \delta\varphi_h \rrbracket \cdot \mathbf{P}_h \cdot \mathbf{N}^- dS_0 \\ = - \int_{\partial_l B_{0h}} \llbracket \delta\varphi_h \rrbracket \cdot \langle \mathbf{P}_h \rangle \cdot \mathbf{N}^- dS_0 - \int_{\partial_l B_{0h}} \langle \delta\varphi_h \rangle \cdot \llbracket \mathbf{P}_h \rrbracket \cdot \mathbf{N}^- dS_0 \end{aligned} \quad (30)$$

Comparing this relation with Eq. (28), an obvious choice for $\mathbf{h}$ is

$$\mathbf{h}(\mathbf{P}_h^-, \mathbf{P}_h^+, \mathbf{N}^-) = \langle \mathbf{P}_h \rangle \cdot \mathbf{N}^- \quad (31)$$

The last term of Eq. (30) may be neglected because only compatibility of the displacements needs to be enforced. This form of the numerical flux was proposed for the scalar elliptic equation by Bassi and Rebay [12]. A complete discussion on alternative numerical flux formulations has been presented by Arnold et al. [10]. The weak statement of equilibrium consisting of finding $\mathbf{P}_h \in S_h^k$

$$\begin{aligned} \int_{B_{0h}} \mathbf{P}_h \cdot \nabla_0 \delta\varphi_h dV_0 - \int_{\partial_N B_{0h}} \delta\varphi_h \cdot \bar{\mathbf{T}} dS_0 \\ + \int_{\partial_l B_{0h}} \llbracket \delta\varphi_h \rrbracket \cdot \mathbf{h}(\mathbf{P}_h^-, \mathbf{P}_h^+, \mathbf{N}^-) dS_0 = 0 \quad \forall \delta\varphi_h \in X_{hc}^k \end{aligned} \quad (32)$$

must be accompanied by a (weak) enforcement of displacement compatibility and must be enforced weakly on $\partial_l B_{0h}$, which also ensures numerical stability. To this end, the compatibility equation $\varphi_h^- - \varphi_h^+ = 0$ on $\partial_l B_{0h}$ is multiplied by a test function $\delta\mathbf{P}_h$ which, by similar arguments used in the weak formulation of the equilibrium equation, leads to the definition of a new flux $\mathbf{g}$ such that

$$\int_{\partial_l B_{0h}} \llbracket \varphi_h \rrbracket \cdot \mathbf{g}(\delta\mathbf{P}^-, \delta\mathbf{P}^+, \mathbf{N}^-) dS_0 = 0 \quad \forall \delta\mathbf{P} \in S_{hc}^k \quad (33)$$

The constitutive law is enforced strongly i.e.

$$\mathbf{P}_h = \bar{\mathbf{P}} = \frac{\partial W}{\partial \mathbf{F}}(\nabla_0 \varphi_h) \quad (34)$$

$$\delta\mathbf{P}_h = \delta\bar{\mathbf{P}} = \bar{\mathbf{C}} \nabla_0 \delta\varphi_h \quad (35)$$

where $\bar{\mathbf{C}} = (\partial\mathbf{P}/\partial\mathbf{F})(\mathbf{P}_h)$, and it is assumed that $\nabla_0 \delta\varphi_h \in S_{hc}^k$. It will be shown below that this judicious choice of a subset of the admissible stress space S_{hc}^k avoids the linearization with respect to

the displacement jump, by contrast to the variational formulation of Sec. 2.

The weak form of the nonlinear elasticity problem resulting from the addition of Eqs. (32) and (33) is thus to find $\varphi_h \in X_{hc}^k$ such that

$$\begin{aligned} \int_{B_{0h}} \bar{\mathbf{P}} \cdot \nabla_0 \delta\varphi_h dV_0 - \int_{\partial_N B_{0h}} \delta\varphi_h \cdot \bar{\mathbf{T}} dS_0 \\ + \int_{\partial_l B_{0h}} \llbracket \delta\varphi_h \rrbracket \cdot \mathbf{h}(\bar{\mathbf{P}}^-, \bar{\mathbf{P}}^+, \mathbf{N}^-) dS_0 \\ + \int_{\partial_l B_{0h}} \llbracket \varphi_h \rrbracket \cdot \mathbf{g}(\bar{\mathbf{C}}^- \nabla_0 \delta\varphi_h^-, \bar{\mathbf{C}}^+ \nabla_0 \delta\varphi_h^+, \mathbf{N}^-) dS_0 \\ = 0 \quad \forall \delta\varphi_h \in X_{hc}^k \end{aligned} \quad (36)$$

Unfortunately, this formulation is numerically unstable, even for consistent numerical fluxes. A common solution to this problem in this type of weak formulation approach is to introduce a quadratic term in $\llbracket \varphi_h \rrbracket$, $\llbracket \delta\varphi_h \rrbracket$. Whereas in scalar problems this can be achieved by simply adding a term proportional to the scalar product $\llbracket \varphi_h \rrbracket \cdot \llbracket \delta\varphi_h \rrbracket$, an appropriate term in the context of nonlinear elasticity must take into account the tensorial character of the virtual work at the interelement boundary and stabilize general displacement discontinuities. A simple way to achieve this is to choose a stabilization term proportional to $\llbracket \varphi_h \rrbracket \cdot \bar{\mathbf{C}} \cdot \llbracket \delta\varphi_h \rrbracket$. The final weak formulation of the problem is to find $\varphi_h \in X_{hc}^k$ such that

$$\begin{aligned} \int_{B_{0h}} \bar{\mathbf{P}} \cdot \nabla_0 \delta\varphi_h dV_0 - \int_{\partial_N B_{0h}} \delta\varphi_h \cdot \bar{\mathbf{T}} dS_0 \\ + \int_{\partial_l B_{0h}} \llbracket \delta\varphi_h \rrbracket \cdot \mathbf{h}(\bar{\mathbf{P}}^-, \bar{\mathbf{P}}^+, \mathbf{N}^-) dS_0 \\ + \int_{\partial_l B_{0h}} \llbracket \varphi_h \rrbracket \cdot \mathbf{g}(\bar{\mathbf{C}}^- \nabla_0 \delta\varphi_h^-, \bar{\mathbf{C}}^+ \nabla_0 \delta\varphi_h^+, \mathbf{N}^-) dS_0 \\ + \int_{\partial_l B_{0h}} \left\{ \llbracket \delta\varphi_h \rrbracket \otimes \mathbf{N}^- : \left\langle \frac{\beta \bar{\mathbf{C}}}{h_s} \right\rangle : \llbracket \varphi_h \rrbracket \otimes \mathbf{N}^- \right\} dS_0 \\ = 0 \quad \forall \delta\varphi_h \in X_{hc}^k \end{aligned} \quad (37)$$

This approach to the stabilization of discontinuous Galerkin formulation in which β plays the role of a penalty parameter is commonly referred to as Interior Penalty method [17]. For numerical fluxes of the form (31), expression (26) is recovered, which demonstrates that the assumptions (34) and (35) and the introduction of a quadratic term correspond to the linearization of the stress tensor in Eq. (21).

4 Formulation From the Framework of Brezzi et al. [13]

In a recent paper, Brezzi et al. [13] have proposed an alternative approach for deriving discontinuous Galerkin methods for the scalar elliptic equation. The main idea is to use as starting point a weak formulation of the continuum equations inside the elements as an individual problem. In addition, interelement compatibility and equilibrium are enforced weakly by the introduction of suitable operators. The resulting weak formulation of the problem becomes finding $\varphi_h \in X_{hc}^k$ such that

$$\begin{aligned} \sum_e \int_{\Omega_0^e} \nabla_0 \cdot \mathbf{P}(\nabla_0 \varphi_h) \cdot \mathbf{B}_0(\delta \varphi_h) dV_0 + \int_{\partial_T B_{0h}} [\![\varphi_h]\!] \cdot \mathbf{B}_1(\delta \varphi_h) \cdot \mathbf{N}^- dS_0 \\ + \int_{\partial_T B_{0h}} \mathbf{N}^- \cdot [\![\mathbf{P}(\nabla_0 \varphi_h)]\!] \cdot \mathbf{B}_2(\delta \varphi_h) dS_0 = 0 \quad \forall \delta \varphi_h \in X_{h_c}^k \end{aligned} \quad (38)$$

where $\mathbf{B}_0$, $\mathbf{B}_1$, and $\mathbf{B}_2$ are operators from $X_{h_c}^k$ to $L^2(B_{0h})$, to $[\text{TR}(\partial_T B_{0h})]^2$ and to $\text{TR}(\partial_T B_{0h})$ respectively. Conditions on these operators required to ensure uniqueness and stability of the solution are discussed in [13]. One possible choice for the operators is

$$\mathbf{B}_0(\delta \varphi_h) = -\delta \varphi_h \quad (39)$$

$$\mathbf{B}_1(\delta \varphi_h) = \frac{\beta}{h_s} \langle \bar{\mathbf{C}} \rangle : [\![\delta \varphi_h]\!] \otimes \mathbf{N}^- \quad (40)$$

$$\mathbf{B}_2(\delta \varphi_h) = -\langle \delta \varphi_h \rangle \quad (41)$$

Introducing these values in Equation (38) and integrating by parts leads to finding $\varphi_h \in X_h^k$ such that

$$\begin{aligned} - \sum_e \int_{\partial \Omega_0^e} \mathbf{N} \cdot \mathbf{P}_h \cdot \delta \varphi_h dS_0 + \sum_e \int_{\Omega_0^e} \mathbf{P}_h : \nabla_0 \delta \varphi_h dV_0 \\ - \int_{\partial_T B_{0h}} \mathbf{N}^- \cdot [\![\mathbf{P}(\nabla_0 \varphi_h)]\!] \cdot \langle \delta \varphi_h \rangle dS_0 + \int_{\partial_T B_{0h}} [\![\varphi_h]\!] \\ \otimes \mathbf{N}^- : \frac{\beta}{h_s} \langle \bar{\mathbf{C}} \rangle : [\![\delta \varphi_h]\!] \otimes \mathbf{N}^- dS_0 = 0 \quad \forall \delta \varphi_h \in X_{h_c}^k \end{aligned} \quad (42)$$

Using Eqs. (3) and (5), this last relation is rewritten as finding $\varphi_h \in X_h^k$

$$\begin{aligned} \int_{B_{0h}} \mathbf{P}_h : \nabla_0 \delta \varphi_h dV_0 + \int_{\partial_T B_{0h}} \mathbf{N}^- \cdot \langle \mathbf{P}(\nabla_0 \varphi_h) \rangle \cdot [\![\delta \varphi_h]\!] dS_0 \\ + \int_{\partial_T B_{0h}} [\![\varphi_h]\!] \otimes \mathbf{N}^- : \frac{\beta}{h_s} \langle \bar{\mathbf{C}} \rangle : [\![\delta \varphi_h]\!] \otimes \mathbf{N}^- dS_0 \\ = \int_{\partial_N B_0} \bar{\mathbf{T}} \cdot \delta \varphi_h dS_0 \quad \forall \delta \varphi_h \in X_{h_c}^k \end{aligned} \quad (43)$$

If the stress tensor is computed directly from the compatible deformation gradient, the kinematic compatibility inside the elements is strongly enforced and one has $\bar{\mathbf{P}} = \mathbf{P}_h = \mathbf{P}(\nabla_0 \varphi_h)$. The resulting formulation of the problem therefore corresponds to (26) without the symmetrization term. The choice

$$\mathbf{B}_1(\delta \varphi_h) = \frac{\beta}{h_s} \langle \bar{\mathbf{C}} \rangle : [\![\delta \varphi_h]\!] \otimes \mathbf{N}^- + \langle \bar{\mathbf{C}} : \nabla_0 \delta \varphi_h \rangle \quad (44)$$

leads to the recovery of the symmetrization term. It is clear that in this approach, there is no need to assume the form of the numerical fluxes, as they arise naturally by a specific suitable choice of interelement boundary operators. It also appears clearly from this formulation that the symmetrization term is not needed, as is also demonstrated in [6].

5 Conclusions

In this work, three different approaches for developing discontinuous Galerkin methods for nonlinear solid mechanics have been considered and compared.

In the first approach, the formulation results from a variational statement of a general energy functional combined with a linear-

ization of the stress tensor with respect to the displacement jumps. The specific forms of the numerical fluxes and stabilization term result naturally from the variational statement.

The second approach consists of a standard weighted-average formulation resulting in a weak statement of the continuum equations, but where both shape and weight functions are allowed to have finite discontinuities inside the domain. In this case, the interelement fluxes and stabilization terms are undefined, which opens the ground for proposing different numerical fluxes and stabilization terms. If, in particular, average numerical fluxes and a quadratic stabilization term are considered, and if the stresses in the element interiors are computed from the compatible deformation gradient, then the formulation obtained from the variational approach is recovered.

An alternative third strategy is to multiply the equilibrium and compatibility equations by suitable operators on the test functions. The combination of the resulting products leads to a new weak formulation, once the operators are selected. A suitable choice of the operators ensures consistency and stability of the formalism. It has been shown that the formulation obtained from the variational approach corresponds to a particular choice of these operators.

It may therefore be concluded that the variational formulation leads naturally to a well defined discontinuous Galerkin method but restricts the expression of the numerical fluxes and the stabilization term. The main advantage of the variational formulation lies in the automatic satisfaction of consistency and linearized stability in the nonlinear range as it was shown in [6].

Acknowledgment

This research was supported by the U.S. Army through the Institute for Soldier Nanotechnologies, under Contract DAAD-19-02-D-0002 with the U.S. Army Research Office. The content does not necessarily reflect the position of the Government, and no official endorsement should be inferred.

References

- [1] Cockburn, B., Hou, S., and Shu, C. W., 1990, "TVB Runge-Kutta Local Projection Discontinuous Galerkin Finite Element Method for Conservation Laws IV: The Multidimensional Case," *Math. Comput.*, **54**, pp. 545–581.
- [2] Cockburn, B., and Shu, C. W., 1998, "The Local Discontinuous Galerkin Method for Time Dependent Convection-Diffusion Systems," *SIAM (Soc. Ind. Appl. Math.) J. Numer. Anal.*, **36**(6), pp. 2440–2483.
- [3] Cockburn, B., 2003, "Discontinuous Galerkin Methods for Convection-Dominated Problems," in *High-Order Methods for Computational Physics*, A. Quateroni, ed., Lecture Notes in Computational Sciences and Engineering Vol. 11, Springer, Berlin, pp. 69–213.
- [4] Engel, G., Garikipati, K., Hughes, T. J. R., Larson, M. G., Mazzei, L., and Taylor, R. L., 2002, "Continuous/Discontinuous Finite Element Approximations of Fourth-Order Elliptic Problems in Structural and Continuum Mechanics with Applications to Thin Beams and Plates," *Comput. Methods Appl. Mech. Eng.*, **191**, pp. 3669–3750.
- [5] Lew, A., Neff, P., Sulsky, D., and Ortiz, M., 2004, "Optimal BV Estimates for a Discontinuous Galerkin Method for Linear Elasticity," *Appl. Math. Research eXpress*, **3**, pp. 73–106.
- [6] Noels, L., and Radovitzky, R., 2006, "A General Discontinuous Galerkin Method for Finite Hyperelasticity. Formulation and Numerical Applications," *Int. J. Numer. Methods Eng.*, **68**, pp. 64–97.
- [7] Ten Eyck, A., and Lew, A., 2006, "Discontinuous Galerkin Methods for Non-Linear Elasticity," *Int. J. Numer. Methods Eng.*, **67**, pp. 1204–1243.
- [8] Güzey, S., Stolarski, H. K., Cockburn, B., and Tamma, K. K., 2006, "Design and Development of a Discontinuous Galerkin Method for Shells," *Comput. Methods Appl. Mech. Eng.*, **195**, pp. 3528–3548.
- [9] Nitsche, J. A., 1971, "Über ein Variationsprinzip zur Lösung Dirichlet-Problemen bei Verwendung von Teilräumen, die Keinen Randbedingungen Unworfen Sind," *Abh. Math. Semin. Univ. Hambg.*, **36**, pp. 9–15.
- [10] Arnold, D. N., Brezzi, F., Cockburn, B., and Marini, L. D., 2002, "Unified Analysis of Discontinuous Galerkin Methods for Elliptic Problems," *SIAM (Soc. Ind. Appl. Math.) J. Numer. Anal.*, **39**(5), pp. 1749–1779.
- [11] Brezzi, F., Manzini, G., Marini, D., Pietra, P., and Russo, A., 2000, "Discontinuous Galerkin Approximations for Elliptic Problems," *Numer. Methods Partial Differ. Equ.*, **16**(4), pp. 365–378.
- [12] Bassi, F., and Rebay, S., 1997, "A High-Order Accurate Discontinuous Finite Element Method for the Numerical Solution of the Compressible Navier-

- Stokes Equations,” J. Comput. Phys., **131**, pp. 267–279.
- [13] Brezzi, F., Cockburn, B., Marini, L. D., and Süli, E., 2006, “Stabilization Mechanisms in Discontinuous Galerkin Finite Element Methods,” Comput. Methods Appl. Mech. Eng., **195**, pp. 25–26.
 - [14] Fraijs de Veubeke, B. M., 1951, “Diffusion des Inconnues Hyperstatiques dans les Voilures à Longerons Couplés,” *Bulletins du Service Technique de L’Aéronautique* Vol. 24, Marcel Hayez, Brussels, Belgium.
 - [15] Hu, H. C., 1955, “On Some Variational Methods on the Theory of Elasticity and Theory of Plasticity,” Sci. Sin., **4**, pp. 33–54.
 - [16] Washizu, K., 1955, “On the Variational Principles of Elasticity and Plasticity,” Technical Report No. 25–18. Aeroelastic and Structures Research Laboratory, Massachusetts Institute of Technology, Cambridge, MA.
 - [17] Cockburn, B., 2003, “Discontinuous Galerkin Methods,” Z. Angew. Math. Mech., **83**(11), pp. 731–754.

Effects of Temperature Changes on the Mechanical and Ballistic Responses in Biathlon Shooting

Christelle Grebot

Alain Burtheret¹

e-mail: alain.burtheret@univ-fcomte.fr

FEMTO-ST, UFR ST de Besançon,
Laboratoire de Mécanique Appliquée R. Chaleat,
UMR CNRS 6174,
Equipe Biomécanique et Mécanismes,
24 chemin de l'Épitaphe,
25000 Besançon, France

Background: Biathlon is a nordic sport that combines cross-country skiing with rifle marksmanship. It was reported that standing shooting was significantly affected because skiing exercise usually decreased the postural control of biathletes and increased the shooting time. Another hypothesis that may explain the decrease of one's shooting accuracy after a cession of cross-country skiing could be linked with mechanical factors. The goal of the present study was to examine the influence of negative temperatures on the trigger mechanism and on the ballistic responses of the bullet. Method of approach: In order to determine the possible variations of the force required for triggering, five biathlon rifles were equipped with strain gauges fixed on the trigger. A thermostat vessel was used to control the temperature changes at room temperature (+20°C) (+68°F) until -20°C (-4°F). Concerning the ballistic measurements, eight series of five shots were performed at +20°C (+68°F), at -3°C (+26.6°F), at -10°C (+14°F), and at -20°C (-4°F). The shooting precision was assessed by determining the group diameter (GD) and the shooting score (Sc). Results: The results showed that from +20°C (+68°F) until -8°C (+17.6°F), the triggering force was equal to 5 N (1.12 lb), whereas at -20°C (-4°F), a triggering force of 8 N (1.8 lb) was required. The increase of the triggering force that was found under -8°C (+17.6°F) could be caused by the difference between the coefficients of expansion of the different materials constituting the trigger mechanism. Concerning the ballistic measurements, GD at room temperature was significantly lower ($P < 0.05$) than -3°C (+26.6°F), -10°C (+14°F), and -20°C (-4°F). Furthermore, Sc was significantly better at +20°C (+68°F) ($P < 0.05$) compared to -3°C (+26.6°F), -10°C (+14°F), and -20°C (-4°F) conditions. Conclusion: It can be supposed that the degradation of GD and Sc could be due to the formation of frost in the barrel and by the difference of the expansion coefficient of the bullet-barrel materials. Consequently, both mechanical responses could partly explain the shooting accuracy impairment observed in negative temperature shooting conditions. [DOI: 10.1115/1.2712229]

Keywords: biathlon shooting, rifle mechanism, trigger, ballistic, temperature

¹Corresponding author.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received March 30, 2006; final manuscript received August 23, 2006. Review conducted by Thomas W. Shield.

1 Introduction

Biathlon is a nordic sport that combines cross-country skiing with rifle marksmanship. Biathlon competitions include cross-country skiing sequences of 2500 m (2285.97 yd) to 5000 m (4571.95 yd), which alternate with 2–4 periods of shooting. Shooting is performed in prone or standing position. Five shots are allowed to hit five targets positioned 50 m (45.19 yd) away from the firing line. In relays competitions, eight shots are allowed to hit the five targets. For prone shooting, the targets are 45 mm (1.77 in.) in diameter and in standing position the targets are 115 mm (4.52 in.) in diameter. At the end of the shooting period, each missed target incurs a 150 m penalty lap (137.15 yd) or a 1 min penalty time depending on the type competition. The lowest cumulative time (ski time, shooting time, and penalties) wins. The best performance in biathlon is obtained by maximizing the skiing velocity and by minimizing the time spent at the firing line and the penalty time. Therefore, the shooting performance is important in the final result.

Biathlon shooting is also a very complex motor activity requiring good postural stability and rapid execution [1]. Yet, the skills required for prone and standing shooting are not identical [2]. Shooting performance while prone requires the ability to discriminate (i.e., discrimination between perfect and approximate aiming), as well as a fine motor control (i.e., triggering action without hand or arm movement). In standing shooting, the stability of the body-rifle system is an important variable that characterizes elite biathletes [3]. However, in competition, it was reported that standing shooting was significantly affected because skiing exercise decreased the postural control of biathletes [4,5] and it increased the shooting time [6]. A significant correlation has been reported between the postural control and the shooting performance in standing position [3–7].

However, another hypothesis that may explain the decrease of the shooting accuracy that is observed after cross-country skiing could be linked with mechanical factors. Yet, biathletes and coaches often report that the negative temperatures impair the shooting precision [8]. According to the International Biathlon Union code of practice and for safety reasons, biathletes must exert a minimum 5 N (1.12 lb) force on the trigger in order to release the shot. Before competitions, this triggering value is tested with a dynamometer by biathlon referees. However, as the effects of negative temperatures on the triggering mechanisms and the shooting ballistic have never been clearly identified, it may be possible that the temperature changes observed during competition could impair the 5 N (1.12 lb) triggering force but also the ballistic responses of the bullet during shooting. Therefore, the purpose of the present study was to examine both the influence of the negative temperatures on the trigger mechanism and on the ballistic responses of the bullet.

2 Methods

2.1 Instrumentation. Five biathlon competition rifles (Anschütz Fortner Biathlon 1827, Germany) were tested in the present study. The triggers of each rifle were partially modified by making on the trigger a flat and rectangular beam (Fig. 1). The law of the strength of materials have been used to determine the section of the trigger. This section has a stiffness corresponding to the different values of the force applied on the trigger. Therefore, the elastic range of the material has been respected. Then, an active unidirectional strain gauge (CEA-06-125UN-350, Vishay MicroMeasure Society, USA) was stuck behind the trigger.

A scanner (model 5100 Vishay Instrument, Vishay MicroMeasure Society, USA) with 25 channels attached to a computer (Compaq) that is equipped with a software (System 5000.Stress analysis data system, Vishay MicroMeasure Society, USA) has been used to obtain data of the strain gauges. Scanner permits a mount in full bridge. The program of extensometric measures per-



Fig. 1 Trigger before (a) and after modification and sticking of the gauge (b)

mits to enter both characteristics of strain gauges and their graphs of temperature's compensations corresponding to each type of strain gauges in order to work out the correction of the apparent deformation. Therefore, if some measures are carried out at positive or negative temperatures, which are different from the reference condition at $+20^{\circ}\text{C}$ ($+68^{\circ}\text{F}$), then a correction of the apparent deformation is necessary. Each box of strain gauges is composed of a signalitic card. The graph of deformation depending on temperature is valid only for the type of strain gauges, indicated on the lot they belong to, as they are stuck on the used metal. The corrected deformation (ε) is obtained by subtracting the measured deformation (ε_i) from the apparent deformation ($\varepsilon_{\text{apparent}}$), i.e.: $\varepsilon = \varepsilon_i - \varepsilon_{\text{apparent}}$.

Thus, microdeformations due to the force applied on the trigger were obtained. The strain gauges were previously calibrated by using a tension-compression machine (Instron, UK) with a load cell of 10 N (2.24 lb). Then a conversion of microdeformations into force was calculated as

$$x\mu\Sigma = nF$$

where $\mu\Sigma$ corresponds to microdeformations and F corresponds to the force (N) (or lb). One microstrain corresponds to a force of 0.04 N (0.009 lb). In order to achieve the threshold of the trigger action, the necessary force in the middle part of the trigger is of $125\mu\Sigma$, i.e., $125 \times 0.04 = 5$ N (1.12 lb).

The ballistic measurements were performed with a test stand in a thermostat vessel, which permitted the reproduction of shooting conditions at negative temperatures. A frame, made of steel, with a rectangular and tubular section, was realized by mechanic welding and surrounded the thermostat vessel in order to secure the rigidity of the whole; it also permitted the fixing of the testing banc, thus securing the support of the rifle. The system that supported the rifle was made up of thread extensions (No. 1 of Fig. 2), which permitted the adjustment of the initial alignment of the rifle. Thus frame and testing banc could if necessary subsequently enable the mounting or the addition of accessories. The hanging system was realized by using extensions (No. 2 of Fig. 2) and some "Hasberg" tape (foils) (No. 3 of Fig. 2) which was of rect-

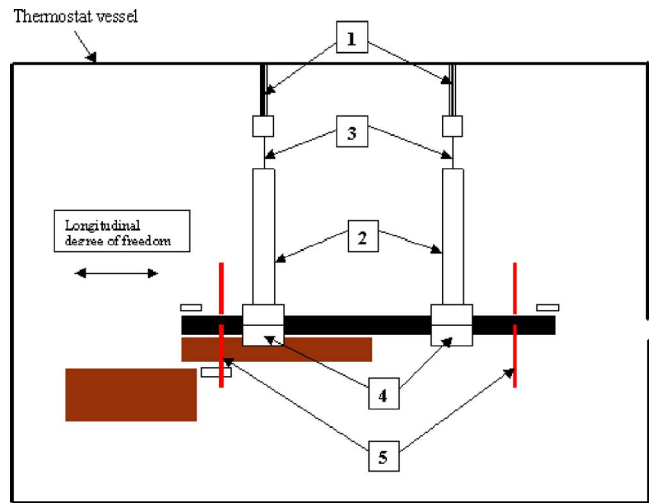


Fig. 2 Hanging system of the rifle

angular shape; it had a thickness of 0.1 mm and was 25 mm wide. Tapes significantly increased the radial stiffness and permitted the longitudinal degree of freedom of the rifle. In order to fix the rifle, a function modulus that would reduce the effects of mechanical tightening on the level of the rifle was realized. The junction to the rifle was made through two parallelepipedic blocks (No. 4 of Fig. 2), which had been machined in their middle in "V" shape. These two blocks were realized with "ertalon," because this material has got stable mechanical and thermic characteristics and an almost nonexistent thermic conductivity. In order to secure the radial support, two stiffeners—one at the back, the other in the front—were created and placed equally on each side of the rifle (No. 5 of Fig. 2). They were adjustable and were provided with some miniature accuracy ball bearing, which were placed in linear contact with the rifle. Therefore, this mount authorized only the longitudinal degree of freedom of the rifle. This system permitted a perfect stability of the rifle during the shooting. Each shooting of five bullets was carried out according to stages of temperatures obtained at a steady rate. This entity was in a position of absolute stability. Therefore, in this case, there was no contraction or thermic stress effects.

The closing lid of the thermostat vessel was made of several elements that could be dismantled simultaneously. It was also composed of portholes in order to visualize the loading manoeuvring. So the rifle was isolated inside the thermostat vessel during the shooting tests at negative temperatures.

The thermostat vessel with its testing banc were situated 50 m (45.71 yd) away from a paper target in a covered shooting range (Fig. 3). A small hole was performed in the vessel to allow the passage of the bullet. Only high quality bullets were used in the present study. All the bullets came from the same manufacturer, trademark and manufacturing batch.

2.2 Test Protocol. In order to measure the force required for triggering, the procedure was performed at room temperature ($+20^{\circ}\text{C}$) ($+68^{\circ}\text{F}$). Then a rifle was placed in a thermostat vessel and the triggering forces were assessed at $+20$, $+15$, $+10$, $+5$, 0 , -3 , -5 , -8 , -10 , -12 , -15 , -18 , and -20°C ($+68$, $+59$, $+50$, $+41$, $+32$, $+26.6$, $+23$, $+17.6$, $+14$, $+10.4$, $+5$, $+0.4$, and -4°F) because this last value corresponds to the extreme temperature authorized in biathlon competitions. According to the International Biathlon Union code of practice, the measures were performed in the middle part of the trigger. Bullets were placed in the same conditions of temperature as the rifle. Barrels were equipped with a probe of temperature in order to check that the expected temperature was attained. Therefore, the external temperature of the barrel

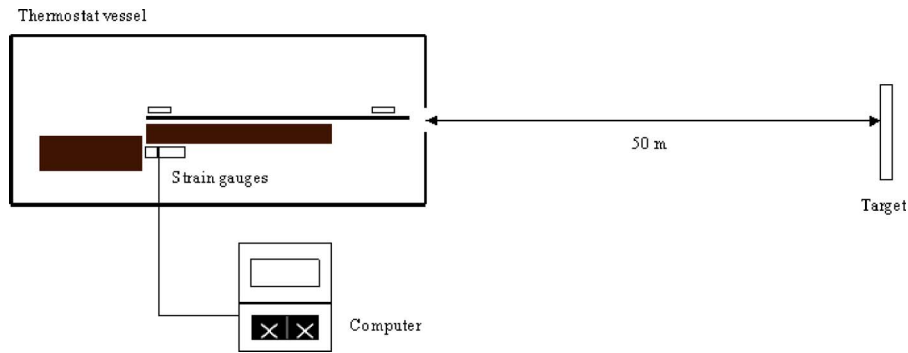


Fig. 3 Experimental mounting

was taken into consideration. Furthermore, the reproducibility of the measures was tested during eight measurements at every temperature. The same protocol was used for the four others rifles.

The ballistic measurements were performed by shooting eight series of five shots at $+20^{\circ}\text{C}$ ($+68^{\circ}\text{F}$), -3°C ($+26.6^{\circ}\text{F}$), -10°C ($+14^{\circ}\text{F}$), and -20°C (-4°F). The shooting precision was assessed by determining (in mm) (in in.) the GD corresponding to the smallest circle that encompassed each series of five rounds, and the Sc that corresponds to the distance (in mm) (in in.) of the mean of each shooting round series from the center of the best mean GD value measured at $+20^{\circ}\text{C}$ ($+68^{\circ}\text{F}$) (Fig. 4).

2.3 Statistics. The data are reported using mean and standard deviation (SD). A one-way analysis of variance with repeated measure and Tukey post hoc test were used to measure the effects of temperature on the ballistic responses. Statistical significance was accepted at the $P < 0.05$ level.

3 Results

3.1 Influence of the Temperature on the Mechanism (Fig. 5). From $+20^{\circ}\text{C}$ ($+68^{\circ}\text{F}$) until 0°C ($+32^{\circ}\text{F}$), the triggering force was equal to 5 N (1.12 lb) for each of the five triggers. Similarly from -3°C ($+26.6^{\circ}\text{F}$) until -8°C ($+17.6^{\circ}\text{F}$), the triggering force was about of 5 N (5.01 SD 0.02), (1.12 lb) (1.126 SD 0.0045) for the five triggers. However, under this temperature, the triggering force increased exponentially and attained 8 N (1.8 lb) at -20°C (-4°F) (8.02 SD 0.05), (1.80 SD 0.011). Note that from -10°C ($+14^{\circ}\text{F}$) some frost appeared on the trigger mechanism.

3.2 Influence of the Temperature on the Ballistic (Figs. 6(a) and 6(b)). A significant increase of group diameter was found ($F(3)=11.17, P < 0.0001$). Post hoc test revealed that at $+20^{\circ}\text{C}$ ($+68^{\circ}\text{F}$) (8.8 SD 0.8 mm) (0.34 SD 0.031 in.)

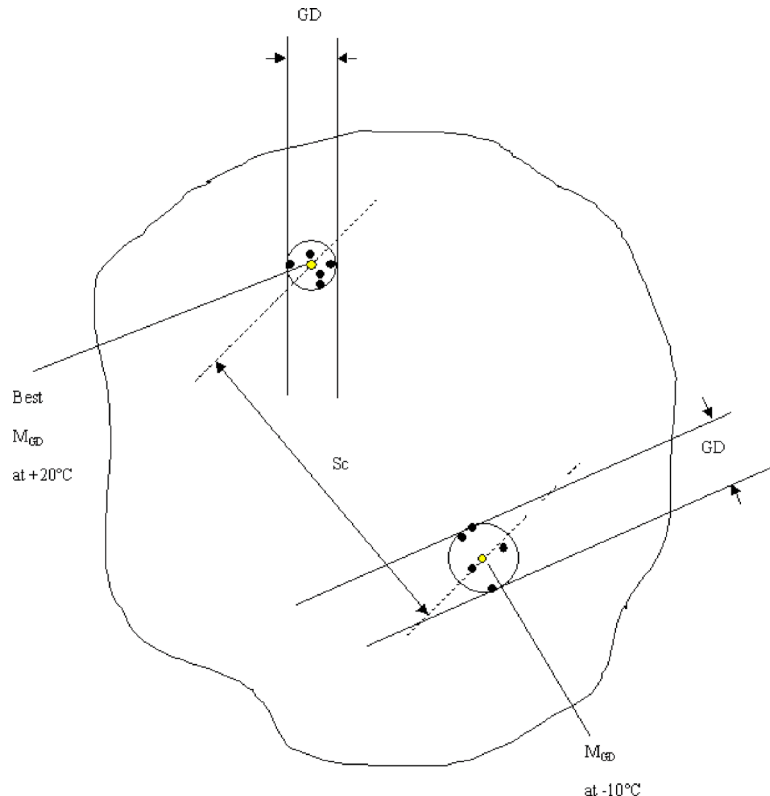


Fig. 4 Measurement (in mm) of the GD, and the Sc. M_{GD} =mean group diameter.

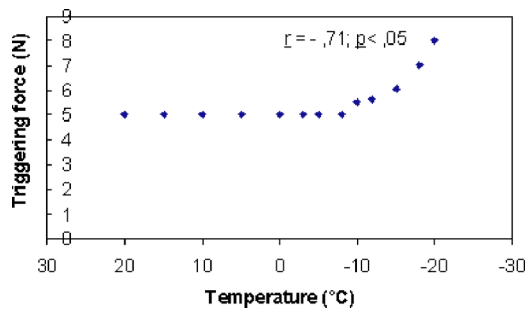
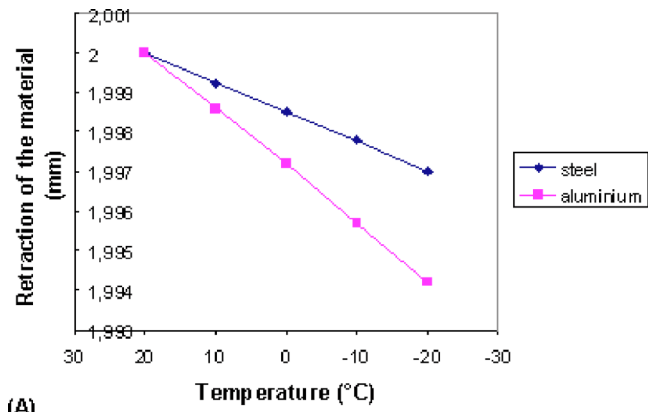


Fig. 5 Relationship between triggering force and temperatures with five cases of triggers

shooting group diameter was significantly lower ($P < 0.05$) than -3°C ($+26.6^{\circ}\text{F}$) (16.7 SD 1.2 mm) (0.65 SD 0.039 in.), -10°C ($+14^{\circ}\text{F}$) (17.1 SD 2.1 mm) (0.67 SD 0.082 in.) and -20°C (-4°F) (15.9 SD 6.2 mm) (0.62 SD 0.24 in.) measurements.

Concerning the shooting score, a significant decrease was found ($F(3)=19.72, P < 0.0001$). At $+20^{\circ}\text{C}$ ($+68^{\circ}\text{F}$) (0.9 SD 1.1 mm) (0.035 SD 0.043 in.) the shooting score was significantly better ($P < 0.05$) than at -3°C ($+26.6^{\circ}\text{F}$) (3.9 SD 1.2 mm) (0.15 SD 0.047 in.), at -10°C ($+14^{\circ}\text{F}$) (4.6 SD 1.5 mm) (0.18 SD 0.059 in.), and at -20°C (-4°F) (6.7 SD 2.4 mm) (0.26 SD 0.09 in.). Furthermore at -3°C ($+26.6^{\circ}\text{F}$), the shooting score was also significantly better ($P < 0.05$) than in the -20°C (-4°F) con-

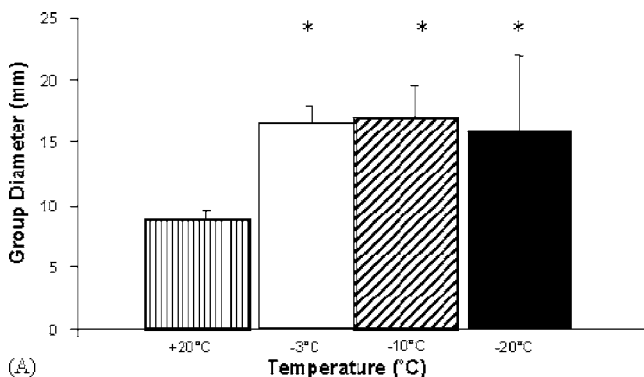


(A)

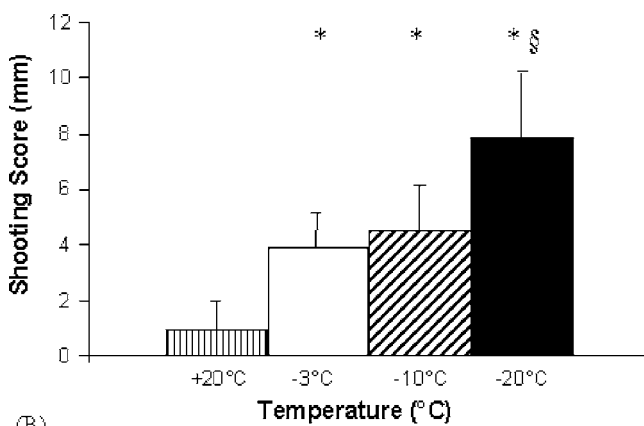
Temperature	Steel Nominal diameter of the pins: 2 mm +0.005 mm -0.005 mm	Aluminum Nominal diameter of the alestage of the case of the trigger: 2.01 mm +0.005 mm -0.005 mm
+20°C	2	2.01
	Difference of free motion : 0.01 mm	
-20°C	1.997	2.005
	Difference of free motion : 0.008 mm	

(B)

Fig. 7 Graph of retraction of aluminum and steel according to temperature (a) and values of retraction of the different materials constituting the pins and the case of the trigger at $+20^{\circ}\text{C}$ ($+68^{\circ}\text{F}$) and -20°C (-4°F) (b)



(A)



(B)

Fig. 6 Mean values and SD of the group diameter (a) and the shooting score (b) performed at different temperatures. * = $P < 0.05$ compared to $+20^{\circ}\text{C}$ ($+68^{\circ}\text{F}$); § = $P < 0.05$ compared to -3°C ($+26.6^{\circ}\text{F}$) condition.

dition. Note that the entire shots performed in negative temperatures were situated under the entire shot rounds performed at $+20^{\circ}\text{C}$ ($+68^{\circ}\text{F}$).

4 Discussion

4.1 Influence of Temperature on the Mechanism. The results of the present study revealed that the triggering force under -8°C ($+17.6^{\circ}\text{F}$) increased exponentially and attained 8 N (1.8 lb) at -20°C (-4°F). This increase of the triggering force could be caused by the retraction of the pins of the triggering mechanism. Thus, as the case of the trigger and the pins do not have the same material (i.e., the case of the trigger is made of aluminum while the pins are made of steel), these different materials have a different coefficient of expansion [23.2 ppm/°C (12.9 ppm/°F) for aluminum and 12.1 ppm/°C (6.7 ppm/°F) for steel] [9]. The retraction of materials shows that frictions increase with the decrease of the temperature (Figs. 7(a) and 7(b)). Knowing aluminum retracts more than steel, it can therefore be understood that the free motion decrease when temperatures decrease (Fig. 7(b)). Depending on the tolerances which are recommended for the standard functioning of a pin in its alestage, these tolerances could intervene with even less free motion in case the temperature decreases even more. When both parts are associated under -8°C ($+17.6^{\circ}\text{F}$), the increasing friction between the pins and the case of the trigger mechanism could increase the triggering force. Consequently, this mechanical response may impair the shooting accuracy of biathletes in low temperature conditions while increasing the latency time between aiming and triggering (Figs. 7(a) and 7(b)).

4.2 Influence of Temperature on the Ballistic. Concerning the influence of negative temperatures on the ballistic responses, the results showed a significant increase of the group diameter at -3°C ($+26.6^{\circ}\text{F}$), -10°C ($+14^{\circ}\text{F}$), and -20°C (-4°F) compared

to $+20^{\circ}\text{C}$ ($+68^{\circ}\text{F}$). Furthermore, the more the shots are performed in extreme negative temperatures the more the distance between the results at the referential condition increase. It can be supposed that the degradation of the group diameter and the shooting score could be caused by the formation of frost in the barrel. The materials' difference of coefficient of expansion between the bullet and the barrel can also explain this degradation. Yet, the lead of the bullet retracts more than the steel of the barrel (the coefficient of expansion of lead is $29.3\text{ ppm}/^{\circ}\text{C}$ ($16.3\text{ ppm}/^{\circ}\text{F}$), whereas it is of $12.1\text{ ppm}/^{\circ}\text{C}$ ($6.7\text{ ppm}/^{\circ}\text{F}$) concerning steel) [9]. Consequently, this difference of expansion coefficient could decrease the bullet's propulsion during firing. The frost and the difference between the coefficients of expansion could influence the shooting ballistic that could be more or less flat. However, this result must be interpreted with some caution because the present experiment does not take into account the possible effects of the atmosphere changes between the thermostat vessel and the atmosphere of the shooting range that may affect the external shooting ballistic.

To sum up, the results of the present study confirm the hypothesis that the triggering force and the ballistic is affected by negative temperatures. Under -8°C ($+17.6^{\circ}\text{F}$), the temperature could enhance the friction between the pins and the case of the trigger mechanism, which increases the triggering force while increasing the latency time between aiming and triggering. Furthermore, the negative temperatures could increase the frictions between the bullets and the barrel that influence the ballistic. Consequently, both mechanical responses could partly explain the shooting ac-

curacy impairment observed in negative temperature shooting conditions. After these thorough analysis of the material, some tests with biathletes could be made at different temperatures in order to couple the entity athletes material in a ulterior stage.

References

- [1] Simoneau, M., Bard, C., Fleury, N., Teasdale, N., and Boulay, M. R., 1997, "The Effects of Metabolic Activation on Postural Stability and Shooting Performance in Elite and Intermediate Biathletes," *Science et Motricité*, **29**, pp. 22–29.
- [2] Wick, J., 1990, *Untersuchungen zur Stellung, zum Ausprägungsgrad und zum Training der Reaktionsleistung Hinsichtlich der Erhöhung der Handlungsschnelligkeit und Genauigkeit beim Schiessen im Biathlon (Effects of a Reaction Time Training on the Shooting Time and Accuracy in Biathlon)*, Leipzig University Press.
- [3] Nitzsche, K., and Stolz, G., 1981, *Zur Objektivierung und Vervollkommung der Schiess Technik im Biathlon (Measurement and Enhancement of the Shooting Skills in Biathlon)*, Leipzig University Press.
- [4] Gros Lambert, A., Grappe, F., and Rouillon, J. D., 1998, "Adaptation Ventilatoire au Tir en Biathlon (Ventilatory Adaptations During Biathlon Shooting)," *Sciences et Sport*, **13**, 135–137.
- [5] Hoffman, M. D., Gilson, P. M., Westenburg, T. M., and Spencer, W. A., 1992, "Biathlon Shooting Performance After Exercise of Different Intensities," *Int. J. Sports Med.*, **13**, pp. 270–273.
- [6] Grebot, C., Gros Lambert, A., and Rouillon, J. D., 2002, "Relation Vitesse Précision au Tir en Biathlon (Velocity-Accuracy Relationship During Biathlon Shooting)," *Sciences et Sport*, **17**, pp. 35–37.
- [7] Gros Lambert, A., Candau, R., Hoffman, M. D., Bardy, B., and Rouillon, J. D., 1999, "Validation of Simple Tests of Biathlon Shooting Ability," *Int. J. Sports Med.*, **20**, 179–182.
- [8] Takahashi, A., 1996, "Ammo for Low Temperatures," www.biathlon.net.
- [9] Avril, J., 1984, *Encyclopédie d'Analyse des Contraintes (Encyclopaedia of stress analysis)*, Micromesures, Paris.

Fluid–Structure Interaction Effects in the Dynamic Response of Free-Standing Plates to Uniform Shock Loading

Nayden Kambouchev

Raul Radovitzky¹

e-mail: rapa@mit.edu

Department of Aeronautics and Astronautics,
Massachusetts Institute of Technology,
Cambridge, MA 02139

Ludovic Noels

Department of Aeronautics and Astronautics,
Massachusetts Institute of Technology,
Cambridge, MA 02139;
LTAS-Milieux Continus & Thermomécanique,
University of Liège,
Chemin des Chevreuils 1,
B-4000 Liège, Belgium

The problem of uniform shocks interacting with free-standing plates is studied analytically and numerically for arbitrary shock intensity and plate mass. The analysis is of interest in the design and interpretation of fluid–structure interaction (FSI) experiments in shock tubes. In contrast to previous work corresponding to the case of incident blast profiles of exponential distribution, all asymptotic limits obtained here are exact. The contributions include the extension of Taylor's FSI analysis for acoustic waves, the exact analysis of the asymptotic limits of very heavy and very light plates for arbitrary shock intensity, and a general formula for the transmitted impulse in the intermediate plate mass range. One of the implications is that the impulse transmitted to the plate can be expressed univocally in terms of a single nondimensional compressible FSI parameter. [DOI: 10.1115/1.2712230]

Keywords: uniform shock waves, fluid-structure interaction

1 Introduction

The reduction of impulse transmitted to structures subject to blast loading provided by the fluid–structure interaction (FSI) effect was recognized in the early work of Taylor [1] who studied the reflection of a blast wave with an exponential pressure profile in the case of negligible fluid compressibility. Taylor [1] showed that the impulse transmitted to the plate is reduced as the plate mass decreases because lighter plates acquire velocity quickly thus relieving the pressure acting on the surface of the plate. This peculiar property of the response of light structures to blast loads has been used in the design of sandwich panels with increased resistance to underwater explosions [2–10]. As part of the conclusions of the analysis, Taylor showed that the impulse transmission

depends on a single nondimensional parameter representing the relative time scales of the blast overpressure and of the fluid–structure interaction.

In previous work [11], we extended Taylor's results by incorporating the effect of compressibility, which is important in the case of blast waves propagating in air. The contributions included: the analysis of the asymptotic limits when the plate is very light and very heavy, the identification of an extended nondimensional FSI parameter which becomes relevant in the compressible range, and a practical formula for calculating the impulse transmitted to the plate for arbitrary plate weights and blast intensities. Whereas the light plate asymptotic result was exact, the heavy plate asymptotic result was approximate due to lack of an explicit solution for the flow field of an exponentially decaying blast pressure profile reflecting from a fixed boundary. An approximate asymptotic result was obtained by assuming that each level of the incident pressure reflected according to the Rankine–Hugoniot shock jump conditions.

In this technical brief, we discuss the simpler case of free-standing plates subject to uniform shock loading. The analysis should prove useful in the design and interpretation of FSI experiments using conventional shock tubes. It is shown that in the case of uniform shocks, the asymptotic case of heavy plates can be obtained exactly. As a first step, the analysis of the acoustic limit is discussed in Sec. 2 resulting in an explicit relationship between the transmitted impulse and the acoustic nondimensional parameter. Section 3 is devoted to the asymptotic analysis for heavy and light plates for arbitrary shock intensities. In Sec. 4 a practical formula interpolating the exact limits and encompassing the intermediate range is also proposed and verified.

The analysis reveals that the FSI is governed by a nondimensional parameter which is analogous to the acoustic parameter due to Taylor but incorporates the state of compressibility of the fluid and, thus, the intensity of the shock. Interestingly, it is found that the dependence of the transmitted impulse ratio with the mass of the plate collapses onto a single curve, independently of the shock intensity. It is concluded that significant reductions in impulse transmission are achievable by reducing the mass of the structure facing the shock.

2 Exact Solution for Acoustic Waves

The derivation presented in this section is a direct extension of Taylor's analysis for exponentially decaying pressure profiles [1] to the case of uniform waves. The problem setup is as follows. An infinite uniform pressure wave of overpressure p_s , propagating in the positive x direction in a fluid medium with density ρ_0 , sound speed a_0 , and pressure p_0 , impinges on an initially stationary plate of mass per unit area m_p located at $x=0$ m. The wave reaches the plate at time $t=0$ s. The location of the plate is denoted by $\xi = \xi(t)$. The pressure on the right side of the plate is assumed to stay constant and to be equal to the atmospheric pressure p_0 at all times.

Newton's second law gives the equation of motion of the plate

$$m_p \frac{d^2 \xi}{dt^2} = p(\xi, t) - p_0 \quad (1)$$

where $p(\xi, t) - p_0$ is the overpressure acting on the plate. The one-dimensional equation of motion for the fluid is

$$\rho \frac{du}{dt} = - \frac{\partial p}{\partial x} \quad (2)$$

where the convective derivative is given by $d/dt = \partial/\partial t + u(\partial/\partial x)$ and the density $\rho = \rho(x, t)$, the pressure $p = p(x, t)$, and the velocity $u = u(x, t)$ are all functions of both the particle location x and the time t . Acoustic waves will cause only small perturbations around the steady-state values, leading to: $\rho = \rho_0 + \tilde{\rho}$, $p = p_0 + \tilde{p}$, and $\xi = 0$

¹Corresponding author.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received July 4, 2006; final manuscript received September 16, 2006. Review conducted by Robert M. McMeeking.

+ $\tilde{\xi}$. The small pressure perturbation $\tilde{p}$ must then satisfy the wave equation

$$\frac{\partial^2 \tilde{p}}{\partial t^2} - a_0^2 \frac{\partial^2 \tilde{p}}{\partial x^2} = 0 \quad (3)$$

and therefore can be expressed as a sum of two waves propagating to the left and to the right at the sound speed a_0

$$\tilde{p}(x, t) = f(x - a_0 t) + g(x + a_0 t) \quad (4)$$

Using Eqs. (1) and (2), the function g is eliminated leading to

$$m_p \frac{d^3 \tilde{\xi}}{dt^3} + \rho_0 a_0 \frac{d^2 \tilde{\xi}}{dt^2} = -2a_0 f'(\tilde{\xi} - a_0 t) \quad (5)$$

If the shape f of the incoming wave is known then this equation can be solved for $\tilde{\xi}$. For uniform shock waves the shape is given by $f(\tilde{\xi} - a_0 t) = p_s = \text{const}$. The appropriate boundary conditions of the differential Eq. (5) are

$$\begin{aligned} \tilde{\xi}(t=0) &= 0 \\ \frac{d\tilde{\xi}}{dt}(t=0) &= 0 \end{aligned} \quad (6)$$

$$\frac{d^2 \tilde{\xi}}{dt^2}(t=0) = \frac{f(0) + g(0)}{m_p}$$

Since initially the plate is not moving and behaves as a rigid boundary that reflects the wave completely, $g(0) = f(0)$ and the solution is

$$\tilde{\xi} = \frac{2p_s m_p}{\rho_0 a_0^2} \left(e^{-\frac{\rho_0 a_0}{m_p} t} - 1 \right) + \frac{2p_s}{\rho_0 a_0} t \quad (7)$$

The solution clearly shows that the time scale of the fluid structure interaction is given by the time constant $t^* = m_p / \rho_0 a_0$. As the incident pressure wave lacks an intrinsic time scale, we arbitrarily choose a time scale t_i and interpret it as the time elapsed from the time of shock impact $t=0$. In what follows we consider the motion of the system comprising the fluid and the plate up to the fixed moment of time $t=t_i$, which can be chosen arbitrarily. Following Taylor [1], one can define a nondimensional parameter $\beta_0 = t_i / t^*$ which compares the relative time constants of the fluid–structure interaction t^* and the incident wave t_i .

The quantity $I_i = p_s t_i$ represents the impulse carried by the incident pressure wave through the point $x=0$ up to the moment of interest $t=t_i$. By noting that the acceleration of the plate $d^2 \tilde{\xi} / dt^2 (t > 0) = 2p_s / m_p e^{-(\rho_0 a_0 / m_p) t}$ remains positive at all times, one concludes that the maximum velocity and, therefore, the maximum impulse of the plate occur at time $t=t_i$. This maximum impulse when expressed in terms of the nondimensional parameter β_0 is

$$\frac{I_p}{I_i} = 2 \frac{1 - e^{-\beta_0}}{\beta_0} \quad (8)$$

Equation (8) is the uniform-wave analog to Taylor's result for exponentially decaying pressure waves [1]

$$\frac{I_p}{I_i} = 2\beta_0^{\beta_0 / (1 - \beta_0)} \quad (9)$$

where β_0 is defined as above and t_i is the decaying time of the incident exponential pressure wave. In both cases $\lim_{\beta_0 \rightarrow 0} (I_p / I_i) = 2$, which represents the fact that for fixed rigid plates the incident wave is reflected completely.

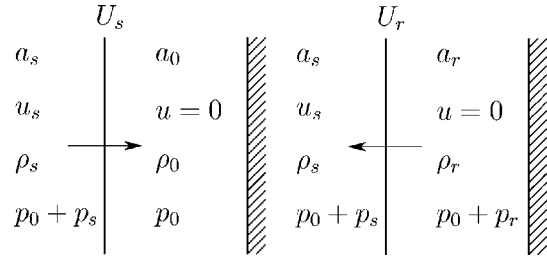


Fig. 1 Incident and reflected shock waves; (a) incident shock wave, (b) reflected shock wave

3 Extension to the Compressible Range

The problem considered in the previous section can be extended to compressible flows by eliminating the assumption that the overpressure of the incident wave is small. By contrast to the acoustic limit, the resulting coupled problem of a compressible nonlinear one-dimensional flow interacting with a plate is not amenable to analytical treatment. Instead we follow the approach in Ref. [11] and find the two asymptotic limits for heavy and light plates. In the intermediate range, we extrapolate a curve from the asymptotic limits and verify these results against numerical computation.

3.1 Heavy Plate Asymptotic Limit. In contrast to the case of exponential incident pressure waves considered in Ref. [11], for uniform incident pressure waves the impulse transmitted to heavy plates may be found exactly. In the heavy plate limit the plate may be considered as a fixed rigid boundary.

Figure 1 shows a schematic of the problem where a is the local speed of sound, u is the local particle velocity, ρ is the local particle density, p is the local pressure or overpressure, and U is the shock wave speed. Subscripts s , r , and 0 are used to denote states behind the incident shock, behind the reflected shock, and atmospheric, respectively. Assuming that the fluid is an ideal calorically perfect gas with constant specific heats ratio equal to γ , which has been shown to be a realistic assumption even for very strong explosions in air [12], and using the Rankine–Hugoniot relations it is found that [13]

$$\rho_s = \rho_0 \frac{(\gamma + 1)p_s + 2\gamma p_0}{(\gamma - 1)p_s + 2\gamma p_0} \quad (10)$$

$$U_s = a_0 \sqrt{\frac{(\gamma + 1)p_s + 2\gamma p_0}{2\gamma p_0}} \quad (11)$$

$$u_s = a_0 \sqrt{\frac{2p_s}{\gamma p_0}} \sqrt{\frac{p_s}{(\gamma + 1)p_s + 2\gamma p_0}} \quad (12)$$

Similarly for the reflected wave, the relations are

$$\rho_r = \rho_s \frac{\gamma p_s + \gamma p_0}{(\gamma - 1)p_s + \gamma p_0} \quad (13)$$

$$U_r = u_s \frac{(\gamma - 1)p_s + \gamma p_0}{p_s} \quad (14)$$

$$p_r = p_s \frac{(3\gamma - 1)p_s + 4\gamma p_0}{(\gamma - 1)p_s + 2\gamma p_0} \quad (15)$$

A direct consequence of these expressions is that the gas density neither in the incident wave ρ_s nor in the reflected wave ρ_r can grow unboundedly as the incident overpressure ratio p_s / p_0 becomes large. In the limit $p_s / p_0 \rightarrow \infty$, the incident density ratio tends to $(\gamma + 1) / (\gamma - 1)$, while the reflected density ratio tends to $\gamma / (\gamma - 1)$. For air ($\gamma = 1.4$) these two ratios are 6 and 3.5, respec-

tively. Another known consequence of the shock reflection is that the pressure reflection coefficient C_R defined as [13]

$$C_R = \frac{p_r}{p_s} = \frac{(3\gamma - 1)p_s + 4\gamma p_0}{(\gamma - 1)p_s + 2\gamma p_0} \quad (16)$$

is also bounded, $2 \leq C_R \leq (3\gamma - 1)/(\gamma - 1)$ (or $2 \leq C_R \leq 8$ for air). The lower limit corresponds to weak acoustic waves and the upper limit to very strong shocks. It should be emphasized that for air the reflection coefficient C_R departs from 2 for rather small overpressures. For example, $C_R = 2.75$ when $p_s = p_0 = 1$ atm, and therefore the use of results from the acoustic limit theory in the evaluation of the effects of blast loads on structures cannot be justified, as it has been pointed out recently [14,11].

The impulse per unit area transmitted to the plate is the time integral of the overpressure it experiences

$$I_p = \int_0^{t_i} p_r dt = p_r t_i \quad (17)$$

Consequently the impulse transmission coefficient is equal to the pressure reflection coefficient

$$\frac{I_p}{I_i} = \frac{p_r t_i}{p_s t_i} = C_R \quad (18)$$

in direct contrast with the result for exponential pressure profiles [11] in which it is always smaller. Another important difference is that this result is exact, while the result for exponential pressure profiles is approximate [11].

3.2 Light Plate Limit. Following Ref. [11], we first consider the acoustic limit. In this case $\beta_0 \rightarrow \infty$ and the plate instantaneously reaches its final velocity

$$\lim_{\beta_0 \rightarrow \infty} \frac{d\xi}{dt} = \lim_{\beta_0 \rightarrow \infty} \frac{2p_s}{\rho_0 a_0} (1 - e^{-\beta_0 t/t_i}) = \frac{2p_s}{\rho_0 a_0} \quad (19)$$

It is interesting to note that in the case of a uniform incident wave, the velocity of a very light plate remains constant in time, whereas in the case of an exponential profile the plate velocity decays exponentially. From Eq. (19) the transmitted impulse is

$$\lim_{\beta_0 \rightarrow \infty} \frac{I_p}{I_i} = \frac{2m_p}{\rho_0 a_0 t_i} = \frac{2}{\beta_0} \quad (20)$$

and independent of time $0 \leq t \leq t_i$.

Following Ref. [11], we assume that the maximum transmitted impulse in the nonlinear compressible range can be derived from the plate velocity u_p at time 0^+ . Toward this end, we consider the expansion wave produced by a fluid initially compressed at overpressure $p_r = C_R p_s$ on a free surface which is initially at rest. Instantaneously upon reflection, the fluid state is characterized by the normal shock reflection on a fixed boundary ($u_r = u_p = 0$) independently of the plate mass m_p . The reflected state can be characterized as

$$p_r = p_s \frac{(3\gamma - 1)p_s + 4\gamma p_0}{(\gamma - 1)p_s + 2\gamma p_0} \quad (21)$$

$$\rho_r = \rho_0 \frac{(\gamma + 1)p_s + 2\gamma p_0}{(\gamma - 1)p_s + 2\gamma p_0} \frac{\gamma p_s + \gamma p_0}{(\gamma - 1)p_s + \gamma p_0} \quad (22)$$

where Eqs. (10), (13), and (16) have been used.

In the limit $m_p \rightarrow 0$, the motion of the plate is equivalent to that of a free surface acted upon by the reflected fluid overpressure p_r on one side and zero overpressure on the other side. An expansion wave propagating at speed U_e is instantaneously formed with the overpressure $p_e = 0$ and the velocity of the fluid particles $u_e = u_p$ on the right and overpressure p_r and velocity $u_r = 0$ on the left. Applying mass and momentum conservation across the expansion wave

$$\rho_r(-U_e) = \rho_e(u_p - U_e) \quad (23)$$

$$-p_r = \rho_r(-U_e)^2 - \rho_e(u_p - U_e)^2 \quad (24)$$

the plate velocity is found to be

$$u_p^2 = \frac{p_r}{\rho_r} \left(\frac{p_r}{\rho_e} - 1 \right) = \frac{p_r}{\rho_r} \frac{2p_r}{(\gamma - 1)p_r + 2\gamma p_0} \quad (25)$$

where the ratio ρ_r/ρ_e has been expressed in terms of the pressure ratio by using the jump conditions Eq. (10). After some algebraic manipulation, the expression for the plate velocity can be written as

$$u_p = \frac{p_s}{\rho_s U_s} C_R \sqrt{\frac{[(\gamma + 1)p_s + 2\gamma p_0]^2 [(\gamma - 1)p_s + \gamma p_0]}{[(\gamma - 1)C_R p_s + 2\gamma p_0][(\gamma - 1)p_s + 2\gamma p_0](\gamma p_s + \gamma p_0)}} \quad (26)$$

where f_R is a nondimensional factor that depends exclusively on the incident overpressure ratio p_s/p_0 . From Eq. (26), the transmitted impulse ratio is finally obtained

$$\lim_{\beta_0 \rightarrow \infty} \frac{I_p}{I_i} = \frac{m_p u_p}{p_s t_i} = \frac{m_p C_R f_R}{\rho_s U_s t_i} \quad (27)$$

This equation clearly reveals that $t_s^* = m_p/\rho_s U_s$ is the time scale of the interaction of the shock wave with the plate and that the interaction process is characterized by the nondimensional parameter $\beta_s = t_i/t_s^*$. In terms of this parameter, Eq. (27) becomes

$$\lim_{\beta_0 \rightarrow \infty} \frac{I_p}{I_i} = \frac{C_R f_R}{\beta_s} \quad (28)$$

It is evident that this expression degenerates to Eq. (20) for weak sonic disturbances as $\lim_{p_s \rightarrow 0} C_R = 2$ and $\lim_{p_s \rightarrow 0} f_R = 1$. In the case of air, f_R remains close to one, reaches a maximum of 1.26 for $p_s/p_0 \approx 3.5$, and tends to $\sqrt{9/7}$ as $p_s/p_0 \rightarrow \infty$.

4 Intermediate Plate Weights and Numerical Verification

It is useful for the purpose of practical application to devise an expression for the maximum momentum transmission coefficient for arbitrary plate weights and shock intensities. As discussed in Ref. [11], the resulting expression should reduce to:

- The acoustic result Eq. (8) for very small overpressures;
- The heavy plate response Eq. (18) for small β_s and arbitrary shock intensities; and
- The light plate limit Eq. (28) for large β_s and arbitrary shock intensities.

A possible expression satisfying these requirements is

$$\frac{I_p}{C_R I_i} = \frac{1 - e^{-\beta_s f_R}}{\beta_s f_R} \quad (29)$$

This formula represents the ratio of momentum acquired by the plate for an arbitrary plate weight and shock intensity and the impulse that would otherwise be transmitted to the plate should fluid-structure interaction effects be ignored. It is interesting that in the case of a uniform incident shock considered in this paper, the resulting expression Eq. (29) collapses into a single curve as a function of the parameter β_s/f_R . The main difference between Eq. (29) and the result presented in Ref. [11] is that the expression proposed here is exact in the heavy plate limit.

A numerical method has been used for the purposes of verifying the various results of the analysis presented in the foregoing as well as the accuracy of the empirical formula Eq. (29) in the intermediate range of plate masses where exact solutions are not available. The numerical method employed as well as the simulations were reported elsewhere [15]. These consisted of generating

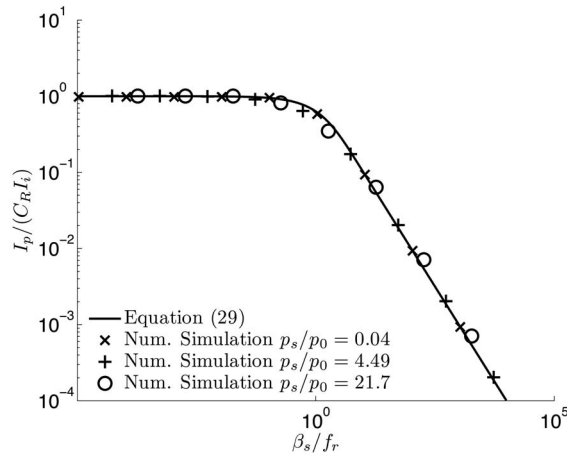


Fig. 2 Impulse transmission as function of the compressible parameter β_s for different values of the incident overpressure p_s/p_0

uniform shocks of varying intensity by applying a piston velocity at one end of the computational grid, followed by a computation of the propagation of the shock and its reflection on plates of varying mass modeled as a concentrated mass at the opposite extreme of the domain. The transmitted impulse I_p was extracted from the simulations and compared with the predictions of Eq. (29).

The numerical results as well as the comparisons with the theory are shown in Fig. 2 where a plot is given of the normalized transmitted impulse $I_p/C_R I_i$ versus the combination of parameters β_s/f_r . As it can be seen in this figure, an excellent agreement is found between the numerical results and the theory. Specifically, for $\beta_s \rightarrow 0$ the curve becomes horizontal supporting the correctness of the assumption that heavy plates behave as fixed walls and therefore absorb the same impulse independently of the plate mass. For $\beta_s \rightarrow \infty$ the curve has slope -1 which is consistent with the assumption that all plates acquire the same maximum velocity (specifically $I_p/I_i \propto m_p$ while $\beta_s \propto 1/m_p$, so that $I_p/I_i \propto 1/\beta_s$). In addition and most importantly, the numerical results support the predictions of the proposed formula Eq. (29) in the intermediate range.

5 Conclusions

Earlier work on the influence of compressibility on fluid–structure interaction effects in the case of exponential blast-wave profiles impinging on free-standing plates of varying mass has been extended to the case of uniform shocks. In this simplified problem, the asymptotic limits of very heavy and very light can be derived exactly for arbitrary shock intensities. The linearized problem for uniform shock profiles is also solved exactly.

It is found that the relative impulse transmitted to the plate can be described by a single nondimensional parameter which is an extension of Taylor's acoustic FSI parameter to the compressible range and incorporates the shock intensity. An explicit approximate expression for this dependence is proposed. In the intermediate range of plate masses the proposed general formula is verified numerically.

Similarly to what has been found before, the use of lighter plates has the benefit of reducing the transmitted impulse, which potentially can be exploited in structural designs with improved blast resistance, e.g., sandwich plates with light front sheets.

Acknowledgment

This research was supported by the U.S. Army through the Institute for Soldier Nanotechnologies, under Contract No. DAAD-19-02-D-0002 with the U.S. Army Research Office. The content does not necessarily reflect the position of the Government, and no official endorsement should be inferred.

References

- [1] Taylor, G. I., 1963, "The Pressure and Impulse of Submarine Explosion Waves on Plates," *The Scientific Papers of Sir Geoffrey Ingram Taylor, Vol. III: Aerodynamics and the Mechanics of Projectiles and Explosions*, G. Batchelor, ed., Cambridge University Press, Cambridge, UK, pp. 287–303.
- [2] Xue, Z., and Hutchinson, J. W., 2003, "Preliminary Assessment of Sandwich Plates Subject to Blast Loads," *Int. J. Mech. Sci.*, **45**(4), pp. 687–705.
- [3] Fleck, N., and Deshpande, V., 2004, "The Resistance of Clamped Sandwich Beams to Shock Loading," *Trans. ASME, J. Appl. Mech.*, **71**(3), pp. 386–401.
- [4] Xue, Z., and Hutchinson, J. W., 2004, "A Comparative Study of Impulse-Resistant Metal Sandwich Plates," *Int. J. Impact Eng.*, **30**(10), pp. 1283–1305.
- [5] Qiu, X., Deshpande, V., and Fleck, N., 2004, "Dynamic Response of a Clamped Circular Sandwich Plate Subject to Shock Loading," *J. Appl. Mech.*, **71**(5), pp. 637–645.
- [6] Qiu, X., Deshpande, V., and Fleck, N., 2005, "Impulsive Loading of Clamped Monolithic and Sandwich Beams Over a Central Patch," *J. Mech. Phys. Solids*, **53**(5), pp. 1015–1046.
- [7] Hutchinson, J. W., and Xue, Z., 2005, "Metal Sandwich Plates Optimized for Pressure Impulses," *Int. J. Mech. Sci.*, **47**(4–5), pp. 545–569.
- [8] Deshpande, V., and Fleck, N., 2005, "One-Dimensional Response of Sandwich Plates to Underwater Shock Loading," *J. Mech. Phys. Solids*, **53**(11), pp. 2347–2383.
- [9] Rabczuk, T., Kim, J. Y., Samaniego, E., and Belytschko, B. T., 2004, "Homogenization of Sandwich Structures," *Int. J. Numer. Methods Eng.*, **61**(7), pp. 1009–1027.
- [10] Liang, Y., Spuskanyuk, A. V., Flores, S. E., Hayhurst, D. R., Hutchinson, J. W., McMeeking, R. M., and Evans, A. G., 2007, "The Response of Metallic Sandwich Panels to Water Blast," *J. Appl. Mech.*, **74**(1), pp. 81–99.
- [11] Kambouchev, N., Noels, L., and Radovitzky, R., 2006, "Compressibility Effects in Fluid–Structure Interaction and their Implications on the Air-Blast Loading of Structures," *J. Appl. Phys.*, **100**, p. 063519.
- [12] Taylor, G. I., 1950, "The Formation of a Blast Wave by a Very Intense Explosion, II The Atomic Explosion of 1945," *Proc. R. Soc. London, Ser. A*, **201**(1065), pp. 175–186.
- [13] Anderson, J., 2001, *Fundamentals of Aerodynamics*, McGraw-Hill, New York.
- [14] Tan, P. J., Reid, S. R., and Harrigan, J. J., 2005, "Discussion: The Resistance of Clamped Sandwich Beams to Shock Loading (Fleck, N. A., and Deshpande, V. S., 2004, *ASME, J. Appl. Mech.*, **71**, pp. 386–401)," *Trans. ASME, J. Appl. Mech.*, **72**, pp. 978–979.
- [15] Kambouchev, N., Noels, L., and Radovitzky, R., 2007, "Numerical Simulation of the Fluid–Structure Interaction Between Air Blast Waves and Free-Standing Plates," *Comput. Struct.*, in press.

Integral Representation of Energy Release Rate in Graded Materials

Z.-H. Jin¹

C. T. Sun

e-mail: sun@purdue.edu

School of Aeronautics and Astronautics,
Purdue University,
West Lafayette, IN 47907

It is well known that, for homogeneous materials, the path-independent J contour integral is the (potential) energy release rate. For general nonhomogeneous, or graded materials, such a contour integral as the energy release rate does not exist. This work presents a rigorous derivation of the extended J integral for general graded materials from the potential energy variation with crack extension. Effects of crack shielding and amplification due to a graded interlayer in an elastic-plastic material system are discussed in terms of this integral. [DOI: 10.1115/1.2712236]

1 Introduction

For a homogeneous elastic material, Rice [1] introduced the path-independent J contour integral and derived it as the (potential) energy release rate G defined by

$$G = - \frac{d\Pi}{da} \quad (1)$$

where Π is the potential energy of the system and a is the crack length. The J integral remains path independent even for nonhomogeneous materials with properties graded in the direction perpendicular to crack extension. For general graded materials, however, the standard J integral loses its path independence and generally it does not represent the energy release rate. Eischen [2] introduced a modification to the J integral by adding an area integral term involving the explicit derivative of the strain energy density with respect to x , the coordinate along the crack direction. The modified integral (called J_1^* in Ref. [2]) reduces to the standard J integral when the integration contour shrinks to the crack tip and hence represents the energy release rate from a local field argument. Eischen [2] claimed that the J_1^* integral also represents the global strain energy release rate but did not offer the proof. For a less general nonhomogeneous material that possesses a constant Poisson's ratio and an exponentially graded shear modulus along the crack direction, Honein and Herrmann [3] introduced a path-independent J_e integral. In applying J_e integral to a split beam problem, the authors mistakenly claimed that this integral was not the energy release rate due.

The present study presents a mathematically rigorous derivation of the J integral as the energy release rate for general graded materials with continuous and piecewise differentiable properties. The implication of crack-tip stress singularity on the derivation is considered following the treatment for homogeneous materials by Jin and Sun [4]. It is shown that both the J_1^* integral [2] for general nonhomogeneous materials and the J_e integral [3] for a particular nonhomogeneous material can be derived as the energy release

rate from a global field argument. Effects of crack-tip shielding and amplification due to a graded interlayer in an elastic-plastic material system are discussed using the modified J integral.

2 Potential Energy Variation and the J Integral in Graded Materials

Consider a two-dimensional nonhomogeneous body with a crack of length a as shown in Fig. 1. The area of the body is denoted by A_0 and the boundary is Γ_0 which consists of the outer boundary Γ and the crack faces Γ_c , i.e.,

$$\Gamma_0 = \Gamma \cup \Gamma_c, \quad \Gamma_c = \Gamma_c^+ \cup \Gamma_a \cup \Gamma_c^- \quad (2)$$

The material is subjected to the prescribed tractions T_i on the boundary segment Γ_t and the prescribed displacements on the boundary segment Γ_u .

In the absence of body forces, the potential energy, Π , of the cracked body per unit thickness is a function of crack length a and can be expressed as

$$\Pi = \Pi(a) = \int \int_{A_0} W dX dY - \int_{\Gamma_t} T_i u_i d\Gamma \quad (3)$$

where (X, Y) is a stationary Cartesian coordinate system, u_i are the displacements corresponding to T_i on Γ_t , and W is the strain energy density given by

$$W = W(\epsilon_{ij}, X, Y) = \int_0^{\epsilon_{ij}} \sigma_{ij} d\epsilon_{ij} \quad (4)$$

Note that W explicitly depends on (X, Y) for graded materials. For an isotropic nonhomogeneous and linearly elastic material, for example, W is

$$W = \mu(X, Y) \left[\epsilon_{ij} \epsilon_{ij} + \frac{\nu(X, Y)}{1 - 2\nu(X, Y)} (\epsilon_{kk})^2 \right] \quad (5)$$

where $\mu(X, Y)$ and $\nu(X, Y)$ are the shear modulus and Poisson's ratio, respectively. The energy release rate associated with a quasi-static crack extension is defined by

$$G = - \frac{d\Pi}{da} = - \frac{d}{da} \int \int_{A_0} W dX dY + \frac{d}{da} \int_{\Gamma_t} T_i u_i d\Gamma \quad (6)$$

Introduce a local coordinate system (x, y) attached to the crack tip, i.e.,

$$x = X - a, \quad y = Y \quad (7)$$

Thus,

$$\frac{d}{da} = \frac{\partial}{\partial a} - \frac{\partial}{\partial x} \quad (8)$$

when the field variables are described in the local coordinate system (x, y) .

Because $\partial W / \partial x$ has a $1/r^2$ singularity at the crack tip (r is the distance from the tip) in linear elastic fracture mechanics since $W \sim 1/r$, as $r \rightarrow 0$ (this is true for graded materials, see Refs. [2,5]), the differentiation with respect to crack length may not be directly performed within the area integral sign in Eq. (6). To properly treat the stress singularity in performing the differentiation, Jin and Sun [4] considered a small square A_h with the center at the crack tip and the boundary denoted by Γ_h , as shown in Fig. 1. Equation (6) now can be written as

¹Currently at Department of Mechanical Engineering, University of Maine, Orono, ME 04469.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received November 7, 2006; final manuscript received November 21, 2006. Review conducted by Robert M. McMeeking.

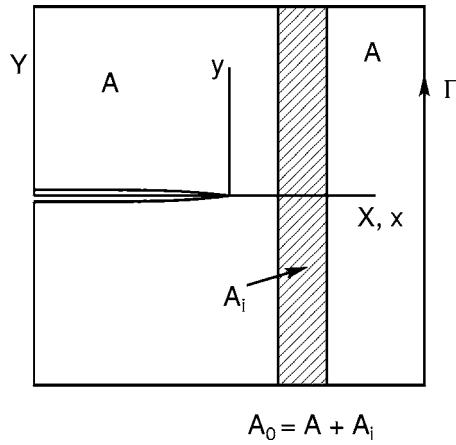


Fig. 2 A graded layer between two dissimilar homogeneous substrates

Assume that the material system is elastically homogeneous and the plastic flow properties are graded only in the interlayer. The deformation plasticity, or nonlinear elasticity, can be applied to stationary cracks. Under small scale yielding conditions, a K -field thus exists far away from the crack tip, but plastic deformations may develop near the crack tip and in the graded layer.

Consider an integration contour Γ in the K -dominance zone. According to Eq. (20), the energy release rate can be evaluated as

$$G = J_{app} - \int \int_{A_p} \frac{\partial W}{\partial x} \bigg|_{\exp l} dx dy \quad (21)$$

where A_p is part of the interlayer undergoing plastic deformations within the contour Γ as shown in Fig. 3 and J_{app} is the usual J integral given by

$$J_{app} = \int_{\Gamma} \left[W dy - T_i \frac{\partial u_i}{\partial x} d\Gamma \right] = \frac{1 - \nu^2}{E} K_I^2 \quad (22)$$

Note that G is not equal to J_{app} due to the graded plastic flow properties in the nonhomogeneous interlayer.

Now consider another integration contour Γ_{tip} within the cracked substrate as shown in Fig. 3. Because the material is homogeneous within Γ_{tip} , the energy release rate is equal to the usual J integral along Γ_{tip}

$$G = J_{tip} \quad (23)$$

It follows from Eqs. (21) and (23) that

$$J_{tip} = J_{app} - \int \int_{A_p} \frac{\partial W}{\partial x} \bigg|_{\exp l} dx dy \quad (24)$$

Equation (24) can be rewritten in the form

$$\frac{J_{tip}}{J_{app}} = 1 - \frac{E\Lambda}{(1 - \nu^2)K_I^2} \quad (25)$$

where the shielding/amplification factor Λ is

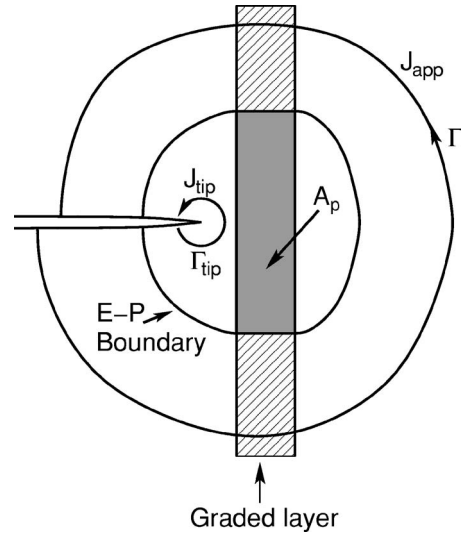


Fig. 3 Contours around the crack tip

$$\Lambda = \int \int_{A_p} \frac{\partial W}{\partial x} \bigg|_{\exp l} dx dy \quad (26)$$

Equation (25) indicates that crack shielding/amplification is caused by the gradation of plastic flow properties in the graded interlayer. The crack tip is shielded ($J_{tip} < J_{app}$) when $\Lambda > 0$, while the opposite is true if $\Lambda < 0$. For a power-law-graded material described by

$$\varepsilon = \frac{\sigma}{E}, \quad \sigma < \sigma_Y(x, y)$$

$$\varepsilon = \frac{\sigma_Y(x, y)}{E} \left(\frac{\sigma}{\sigma_Y(x, y)} \right)^{n(x, y)}, \quad \sigma \geq \sigma_Y(x, y) \quad (27)$$

the factor Λ is

$$\Lambda = \int \int_{A_p} \left(\frac{\partial W}{\partial \sigma_Y} \frac{\partial \sigma_Y}{\partial x} + \frac{\partial W}{\partial n} \frac{\partial n}{\partial x} \right) dx dy \quad (28)$$

References

- [1] Rice, J. R., 1968, "Mathematical Analysis in the Mechanics of Fracture," *Fracture*, Vol. 2, Liebowitz, H., ed., Academic Press, New York, pp. 191–311.
- [2] Eischen, J. W., 1987, "Fracture of Nonhomogeneous Materials," *Int. J. Fract.*, **34**, pp. 3–22.
- [3] Honein, T., and Herrmann, G., 1997, "Conservation Laws in Nonhomogeneous Plane Elastostatics," *J. Mech. Phys. Solids*, **45**, pp. 789–808.
- [4] Jin, Z.-H., and Sun, C. T., 2004, "On Potential Energy Variation and J -Integral," *Int. J. Fract.*, **126**, pp. L19–L24.
- [5] Jin, Z.-H., and Noda, N., 1994, "Crack Tip Singular Fields in Nonhomogeneous Materials," *ASME J. Appl. Mech.*, **61**, pp. 738–740.
- [6] Kim, A. S., Suresh, S., and Shih, C. F., 1997, "Plasticity Effects on Fracture Normal to Interfaces With Homogeneous and Graded Compositions," *Int. J. Solids Struct.*, **34**, pp. 3415–3432.

Elastic Fields Around the Cohesive Zone of a Mode III Crack Perpendicular to a Bimaterial Interface

W. Zhang

X. Deng

e-mail: deng@enr.sc.edu

Department of Mechanical Engineering,
University of South Carolina,
Columbia, SC 29208

Asymptotic stress and displacement fields near the cohesive zone ahead of a semi-infinite Mode III crack normal to a bimaterial interface are derived using elliptic coordinates.
[DOI: 10.1115/1.2723814]

1 Introduction

Fracture of composites often takes place near or at interfaces between dissimilar material layers. Pre-existing microcracks at an interface strongly affect the initiation and propagation of macrocracks. Accordingly, much research has been carried out on the interface crack problem, e.g., see review in Ref. [1].

A particular cracking configuration in composites involves cracks meeting bimaterial interfaces. Zak and Williams [2] first analyzed this type of problems with an eigenfunction expansion method, which was followed by further studies (e.g., Refs. [3–7]). These investigations revealed that stress singularity still exists when a crack touches a bimaterial interface of elastic materials, except in a special case of bending [4]. The singularity factor was found to depend on some dimensionless constants [3,4] or the Dundurs bimaterial parameters [5–7]. The well-known square-root singularity will be recovered if the interface is formed by identical materials.

In real materials infinite stresses do not actually occur since mechanisms such as finite plasticity will release the stresses. To address this issue, Dugdale [8] and Barenblatt [9] proposed that a cohesive zone (fracture process zone) exists ahead of a crack tip and it limits the stresses to physically meaningful magnitudes. In this zone, crack surfaces are partially developed and some traction is sustained between the partly developed crack surfaces, which are commonly referred to as the cohesive surfaces.

Extensions of the cohesive zone concept to situations in composites involving cracks meeting bimaterial interfaces have been very limited in the literature. A notable study by Romeo and Balarini [10] considered cracks perpendicular to and terminating at a bimaterial interface. By incorporating a cohesive zone, the authors were able to circumvent an unfavorable prediction in linear elastic fracture mechanics (LEFM) that a crack at either zero or infinite applied load might penetrate an interface from a relatively stiffer material to a more compliant one, or vice versa.

The purpose of the current technical brief is to provide an understanding of near-tip fields around a crack meeting a bimaterial interface when a cohesive zone is present ahead of the crack tip. To simplify the mathematics, a Mode III problem is considered in which a crack is normal to a bimaterial interface between two isotropic linear elastic materials, as described in detail in Sec. 2. It

is worth noting that the problem of an interfacial crack with a cohesive zone ahead of the crack tip has been treated recently by the authors [11].

2 Linear Elastic Fields

We consider a semi-infinite crack that is perpendicular to a bimaterial interface and with a cohesive zone ahead of the crack tip (see Fig. 1). To make the problem mathematically tractable and to obtain close-form solutions, we suppose that half of the cohesive zone connects the crack tip to the bond line. As illustrated in Fig. 1, the entire problem plane is divided into three regions: Regions 1 and 3 are occupied by Material 1, and Region 2 is occupied by Material 2.

2.1 General Solutions. To effectively account for the cohesive surface boundaries, we take advantage of elliptic cylindrical coordinates u, v, w , which are related to the Cartesian coordinates x, y, z in the plane of oxy by

$$\zeta = c \cosh(\omega) \quad (1)$$

where c is a parameter and is chosen to be one half of the cohesive zone size, and

$$\zeta = x + iy, \quad \omega = u + iv \quad (2)$$

are complex variables in the Cartesian and elliptic coordinate systems, respectively.

Figure 2 shows the locations of various quantities of the crack and the cohesive zone in the elliptic coordinate system. The two lines, $v = \pi$ and $v = -\pi$, represent the crack surfaces. The origin of the Cartesian coordinates is located at $u = 0$ and $v = \pm \pi/2$, the crack tip is placed at $x = -c$ and $y = 0$ (where $u = 0, v = \pm \pi$), and the two line segments with length $2c$ ($u = 0, -\pi \leq v \leq \pi$) stand for the cohesive surfaces. The bimaterial interface is represented by $x = 0$ ($v = \pm \pi/2$).

For a Mode III crack, the displacement and stresses can be expressed as the real and imaginary parts of analytic functions [12]

$$u_z^j = \text{Im}(X_j(\zeta))/\mu_j, \quad \sigma_{zx}^j = \text{Im}(X_j'(\zeta)), \quad \sigma_{zy}^j = \text{Re}(X_j'(\zeta)) \quad (3)$$

where j refers to the region number ($j = 1, 2, 3$); μ_j is the shear modulus; and $X_j'(\zeta)$ is the differentiation of $X_j(\zeta)$ and is chosen as

$$X_j'(\zeta(\omega)) = (\lambda^2 - 1)(b_j + ia_j)\mu_j e^{\lambda\omega} \quad (4)$$

with λ, a_j , and b_j being real parameters (as seen later, λ is the eigenvalue of a characteristic equation). Integration of Eq. (4) with respect to ζ yields

$$X_j(\zeta(\omega)) = \frac{c}{2}(b_j + ia_j)\mu_j[(\lambda - 1)e^{(\lambda+1)\omega} - (\lambda + 1)e^{(\lambda-1)\omega}] \quad (5)$$

Substituting Eqs. (4) and (5) into Eq. (3), we obtain the shear stresses and the out-of-plane displacement associated with Mode III in terms of elliptic coordinates

$$\sigma_{zx}^j = (\lambda^2 - 1)\mu_j e^{\lambda u} [b_j \sin(\lambda v) + a_j \cos(\lambda v)] \quad (6)$$

$$\sigma_{zy}^j = (\lambda^2 - 1)\mu_j e^{\lambda u} [b_j \cos(\lambda v) - a_j \sin(\lambda v)] \quad (7)$$

$$u_z^j = \frac{c}{2} e^{(\lambda+1)u} [(\lambda - 1)(b_j \sin((\lambda + 1)v) + a_j \cos((\lambda + 1)v)) - (\lambda + 1) \times (b_j \sin((\lambda - 1)v) + a_j \cos((\lambda - 1)v)) e^{-2u}] \quad (8)$$

To focus on the near-tip asymptotic behavior, and without specifying the details of a cohesive law, a solution possessing the following features is sought: (a) there are no tractions on the crack surfaces (however there are tractions on the cohesive surfaces); (b) stresses are finite everywhere; (c) tractions are continuous across the cohesive zone; (d) there is a finite displacement jump

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received March 3, 2006; final manuscript received February 2, 2007. Review conducted by Zhigang Suo.

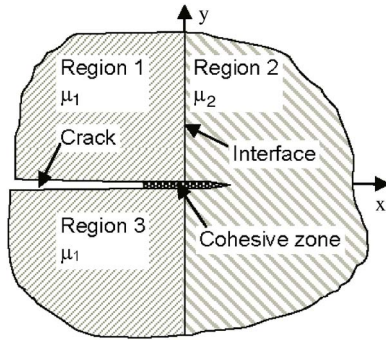


Fig. 1 A cracked bimaterial plate with a crack perpendicular to the bimaterial interface and with a crack-tip cohesive zone passing through the interface

across the cohesive zone; and (e) stresses tend to zero far away from the near-tip region. Some of these considerations have been made when Eq. (4) is chosen.

2.2 Boundary Value Problem. We assume that the same λ in Eqs. (6)–(8) applies to all three regions of the problem domain. Then the continuity at the interface is independent of the coordinate u . The boundary conditions on the crack surfaces ($v = \pm\pi$) and the continuity conditions across the interface ($v = \pm\pi/2$) are

$$\sigma_{zy}^1(u, \pi) = \sigma_{zy}^3(u, -\pi) = 0 \quad (9)$$

$$\sigma_{zx}^1(u, \pi/2) = \sigma_{zx}^2(u, \pi/2), \quad \sigma_{zx}^3(u, -\pi/2) = \sigma_{zx}^2(u, -\pi/2) \quad (10)$$

$$u_z^1(u, \pi/2) = u_z^2(u, \pi/2), \quad u_z^3(u, -\pi/2) = u_z^2(u, -\pi/2) \quad (11)$$

For brevity, and noting $\mu_3 = \mu_1$, let

$$\begin{aligned} c_1 &= a_1\mu_1, & c_2 &= b_1\mu_1, & c_3 &= a_2\mu_2, \\ c_4 &= b_2\mu_2, & c_5 &= a_3\mu_1, & c_6 &= b_3\mu_1 \end{aligned} \quad (12)$$

Substituting Eqs. (6)–(8) and (12) into Eqs. (9)–(11), we obtain the characteristic equation

$$\sin(\lambda\pi)[(\mu_1 + \mu_2)\cos(\lambda\pi) + \mu_1 - \mu_2] = 0 \quad (13)$$

which is a special form of the one obtained by Fenner [4]. The roots of Eq. (13) give the complete set of eigenvalues

$$\lambda = 2n \pm \Lambda, \quad \lambda = n \quad (14)$$

where n is an integer and Λ is determined by a bimaterial parameter k

$$\Lambda = \frac{1}{\pi} \arccos\left(\frac{k-1}{k+1}\right), \quad k = \mu_2/\mu_1 \quad (15)$$

If stress concentration is taken to occur only in the crack tip region (which is a natural and reasonable assumption because the crack tip region is where the only geometric discontinuity occurs

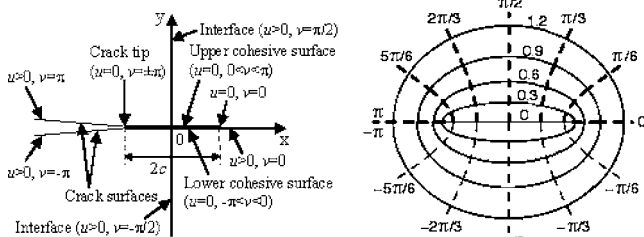


Fig. 2 Locations of the crack and cohesive zone quantities in terms of elliptic coordinate u (solid line) and v (dashed line)

in the current asymptotic analysis), then only negative values of λ are admissible. Noting that $0 < \Lambda < 1$ for $0 < k < \infty$ in Eq. (15), this means that the integer n can only have negative values for the second term in Eq. (14) and only zero (which must be accompanied by the choice of the “–” sign in front of Λ) and negative values for the first term in Eq. (14). As such, the leading eigenfunction term is associated with the leading eigenvalue $\lambda = -\Lambda$.

For illustration purpose, the leading solution term corresponding to $\lambda = -\Lambda$ will be obtained explicitly and will be used to view the near-tip field variation. For this eigenvalue, we solve Eqs. (9)–(11) for c_1, c_3, c_4, c_5 , and c_6 in terms of c_2 (which, for brevity, is replaced by $-p$)

$$c_1 = p/\tan(\Lambda\pi), \quad c_2 = -p, \quad c_3 = 0 \quad (16)$$

$$c_4 = -p(k+1)/2, \quad c_5 = -p/\tan(\Lambda\pi), \quad c_6 = -p \quad (17)$$

2.3 Leading-Term Field Solutions. Using the parameters and the leading eigenfunction term associated with $\lambda = -\Lambda$, stress and displacement field solutions in the three regions can be derived from Eqs. (6)–(8)

$$\sigma_{zx}^1 = -\tau_m e^{-\Lambda u} \cos(\Lambda(\pi - v)) \quad (18)$$

$$\sigma_{zy}^1 = \tau_m e^{-\Lambda u} \sin(\Lambda(\pi - v)) \quad (19)$$

$$u_z^1 = \frac{ku_m e^{(1-\Lambda)u}}{2(k+1)} [e^{-2u}(\Lambda-1)\cos(\Lambda\pi - (\Lambda+1)v) - (\Lambda+1)\cos(\Lambda\pi - (\Lambda-1)v)] \quad (20)$$

in Region 1 (containing Material 1)

$$\sigma_{zx}^2 = -\frac{k+1}{2} \tau_m e^{-\Lambda u} \sin(\Lambda\pi) \sin(\Lambda v) \quad (21)$$

$$\sigma_{zy}^2 = \frac{k+1}{2} \tau_m e^{-\Lambda u} \sin(\Lambda\pi) \cos(\Lambda v) \quad (22)$$

$$u_z^2 = \frac{u_m e^{(1-\Lambda)u}}{4} \sin(\Lambda\pi) [e^{-2u}(\Lambda-1)\sin((\Lambda+1)v) - (\Lambda+1)\sin((\Lambda-1)v)] \quad (23)$$

in Region 2 (containing Material 2), and

$$\sigma_{zx}^3 = \tau_m e^{-\Lambda u} \cos(\Lambda(\pi + v)) \quad (24)$$

$$\sigma_{zy}^3 = \tau_m e^{-\Lambda u} \sin(\Lambda(\pi + v)) \quad (25)$$

$$u_z^3 = \frac{ku_m e^{(1-\Lambda)u}}{2(k+1)} [(\Lambda+1)\cos(\Lambda\pi + (\Lambda-1)v) - e^{-2u}(\Lambda-1)\cos(\Lambda\pi + (\Lambda+1)v)] \quad (26)$$

in Region 3 (containing Material 1), where

$$\tau_m = \frac{(1-\Lambda^2)p}{\sin(\Lambda\pi)}, \quad u_m = \frac{(k+1)pc}{\mu_1 k \sin(\Lambda\pi)} \quad (27)$$

To illustrate the field solutions, the normalized quantities, σ_{zx}/τ_m , σ_{zy}/τ_m and u_z/u_m corresponding to the bimaterial parameter $k=2$, are plotted against the normalized Cartesian coordinates ($x/c = \cosh(u)\cos(v)$, $y/c = \sinh(u)\sin(v)$) in Fig. 3. Figures 3(a) and 3(b) clearly show that both shear stresses have finite magnitudes, and that σ_{zx} is continuous but σ_{zy} is discontinuous across the interface ($x=0$). Figure 3(c) reveals that the out-of-plane displacement varies smoothly from the cracked region ($x < -c$) to the uncracked region ($x > c$) through the cohesive zone ($-c \leq x \leq c$). This differs from the LEFM that predicts abrupt change of displacement field at the crack tip. Thus finite strain is sustained everywhere in the current solution but infinite strain arises at the crack tip in LEFM.

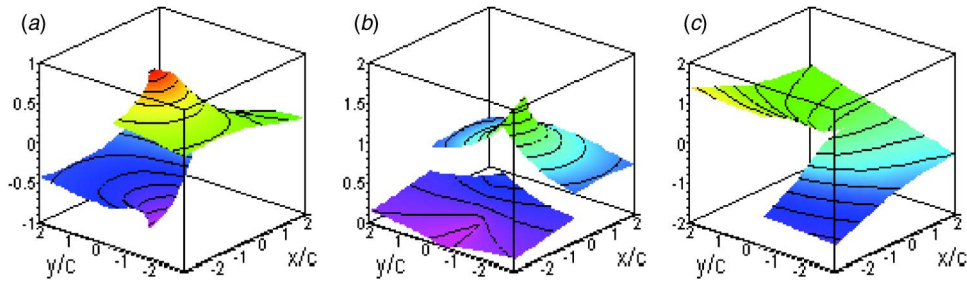


Fig. 3 Contour plots of normalized shear stresses and out-of-plane displacement near the crack tip for $k=2$: (a) σ_{zx}/τ_m ; (b) σ_{zy}/τ_m ; and (c) u_z/u_m

3 Discussion

In the example of $k=2$ shown in Fig. 3, Material 2 is stiffer, thus σ_{zy} is larger near the bond line in Material 2 ($x>0$) than in Material 1 ($x<0$). Figure 4 shows the profiles of normalized σ_{zy} along the crack path ($y=0^\pm$) and the interface ($x=0^\pm$) for $k=1/2$, where the maximum σ_{zy} happens inside the cohesive zone at the interface in Material 1 instead of at the tip of the cohesive zone in Material 2, since Material 1 is stiffer in this case.

The above comparison is based on the leading eigenfunction term (with $\lambda=-\Lambda$). A general solution would be a combination of Eqs. (18)–(26) and some higher-order terms with other eigenvalues (Eq. (14)). As an analogy, we can view the eigenfunctions (Eqs. (6)–(8)) as expansion terms of a “Fourier” series and interpret the eigenvalues in Eq. (14) as “frequencies” in the spectrum space. The selected eigenvalue $\lambda=-\Lambda$ has the lowest admissible “frequency” and hence should be the leading term so that stress concentration occurs in the crack tip region. In a homogeneous material, the leading term corresponding to $\lambda=-1/2$ has been shown to be a good asymptote of the classical LEFM solution [13]. The leading term for $\lambda=-\Lambda$ can also be compared with the first-order LEFM solution in a similar way. It is emphasized that other eigenvalues may also be needed for a more accurate analysis.

A few notes are necessary for the current approach. First, we have assumed that the cohesive zone is divided equally into two halves (Fig. 2). This is an idealized simplification when a Mode III crack is penetrating the interface. Second, the traction continuity conditions across the cohesive zone interface are not used explicitly in the derivations but they have been taken into consideration when choosing the basic function (Eq. (4)) which leads to eigenfunctions that satisfy the boundary conditions. Third, boundary conditions along the cohesive surfaces are not prescribed, which implies that a cohesive zone model is not predefined. Fourth, to accommodate a particular cohesive zone model, it is expected that stress and displacement field terms associated with

multiple eigenvalues may be needed in order to properly characterize the traction and separation variations along the cohesive surfaces.

To illustrate the last point conveniently, let Material 1 and Material 2 be identical (so that $\Lambda=1/2$ in Eq. (14)). The corresponding cohesive stress (along $u=0$), when all admissible eigenterms are considered, can be written as

$$\sigma_{zy}(\nu) = \sum_n \tau_n \cos((n-1/2)\nu), \quad n=0, -1, -2, \dots \quad (28)$$

where antisymmetric conditions have been accounted for (therefore only the first set of eigenvalues in Eq. (14) is activated).

Equation (28) is actually a special form of Fourier series [14], where all terms with sine functions are zero. Thus for a specific cohesive law, an asymptotic solution can be obtained by selecting a sufficient number of eigenterms. Since the cosine functions form an orthogonal system, the coefficients in Eq. (28) can be readily derived by

$$\tau_n = \frac{2}{\pi} \int_0^\pi \sigma_{zy}(\nu) \cos((n-1/2)\nu) d\nu \quad (29)$$

As an example, let us consider the Dugdale cohesive model $\sigma_{zy}=\sigma_0$ (i.e., traction is constant in the cohesive zone), and the maximum separation $\Delta u_z=\delta_0$ occurs at the crack tip. Here, σ_0 and δ_0 are material constants. Once all coefficients τ_n are determined using Eq. (29), the half cohesive zone size, c , can be computed by summing the eigenterms for the displacement and equating the crack opening displacement at the crack tip to δ_0 . Figure 5 shows

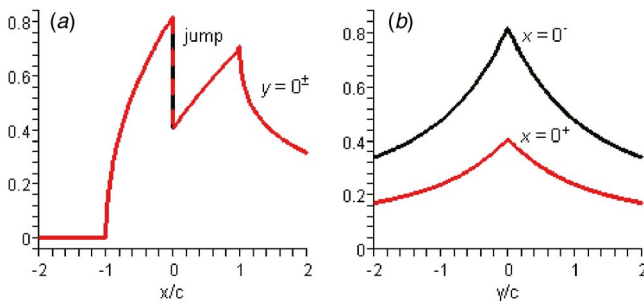


Fig. 4 Spatial distributions of normalized shear stresses for $k=1/2$: (a) σ_{zy}/τ_m along $y=0^\pm$; and (b) σ_{zy}/τ_m along $x=0^\pm$

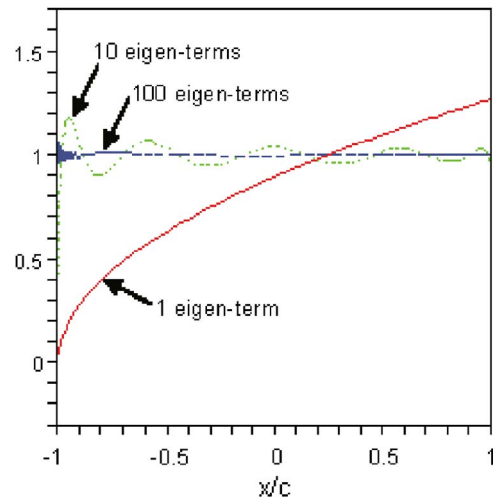


Fig. 5 Approximation of cohesive stress distribution of the Dugdale model $\sigma_{zy}/\sigma_0=1$ for identical materials ($k=1$)

the approximation of the Dugdale model when totally 1 ($n=0$), 10 ($n=-9, -8, \dots, 0$), and 100 ($n=-99, -98, \dots, 0$) eigenterms are used in Eq. (28). It clearly demonstrates that the given $\sigma_{zy}=\sigma_0$ in the cohesive zone ($-1 \leq x/c \leq 1$, $y=0$) can be approached to a certain cutoff error by using an adequate number of eigenfunction terms.

It is worth noting that in the above analysis the enforcement of a cohesive law ahead of the crack tip implies that the cohesive zone is fully developed, the crack is fully open, and crack growth is impending (that is, the stress field around the crack tip is in a steady state, as in quasi-static crack growth, and thus is independent of the external load). In particular, this implies that the coefficients of the stress and displacement eigenterms and the half cohesive zone size, c , depend only on the material constants σ_0 and δ_0 but do not depend on the external load. On the other hand, if the cohesive zone is considered not fully developed, then an interpretation of the above analysis is that the crack opening δ_0 at the crack tip is not critical, thus it (and hence the cohesive zone size) is controlled by and depends on the external load.

4 Summary

This Technical Brief provides an understanding of the near-tip elastic fields around a Mode III crack with a cohesive zone. The crack is perpendicular to a bimaterial interface and the crack-tip cohesive zone penetrates the interface, with the midpoint of the cohesive zone being situated right at the interface. Stress and displacement field variations based on the leading eigenfunction term have been presented. The stresses do not have any singularity and the displacement field has a finite jump across the cohesive zone. The results of this study can serve as a basis for further investigations relevant to the analyses of fracture of materials near interfaces when the effects of crack-tip cohesive zones are considered.

Acknowledgment

The authors gratefully acknowledge the support provided by the ONR DEPSCoR program through Grant No. N00014-03-1-0807.

References

- [1] Suo, Z., 1990, "Singularities, Interfaces and Cracks in Dissimilar Anisotropic Media," *Proc. R. Soc. London, Ser. A*, **A427**, pp. 331–358.
- [2] Zak, A. R., and Williams, M. L., 1963, "Crack Point Stress Singularities at a Bimaterial Interface," *ASME J. Appl. Mech.*, **30**, pp. 142–143.
- [3] Cook, T. S., and Erdogan, F., 1972, "Stresses in Bonded Materials With a Crack Perpendicular to the Interface," *Int. J. Eng. Sci.*, **10**, pp. 677–697.
- [4] Fenner, D. N., 1976, "Stress Singularities in Composite Materials With an Arbitrarily Oriented Crack Meeting an Interface," *Int. J. Fract.*, **12**, pp. 705–721.
- [5] Comninou, M., and Dundurs, J., 1979, "A Closed Crack Tip Terminating at an Interface," *ASME J. Appl. Mech.*, **46**, pp. 97–100.
- [6] Gharpuray, V. M., Dundurs, J., and Keer, L. M., 1991, "A Crack Terminating at a Slipping Interface Between Two Materials," *ASME J. Appl. Mech.*, **58**, pp. 960–963.
- [7] Chen, D. H., 1994, "A Crack Normal to and Terminating at a Bimaterial Interface," *Eng. Fract. Mech.*, **49**, pp. 517–532.
- [8] Dugdale, D. S., 1960, "Yielding of Steel Sheets Containing Slits," *J. Mech. Phys. Solids*, **8**, pp. 100–104.
- [9] Barenblatt, G. I., 1962, "The Mathematical Theory of Equilibrium Cracks in Brittle Fracture," *Adv. Appl. Mech.*, **7**, pp. 55–129.
- [10] Romeo, A., and Ballarini, R., 1997, "A Cohesive Zone Model for Cracks Terminating at a Bimaterial Interface," *Int. J. Solids Struct.*, **34**, pp. 1307–1326.
- [11] Zhang, W., and Deng, X., 2006, "Asymptotic Fields Around an Interfacial Crack With a Cohesive Zone ahead of the Crack Tip," *Int. J. Solids Struct.*, **43**, pp. 2989–3005.
- [12] Unger, D. J., 1995, *Analytical Fracture Mechanics*, Academic, San Diego.
- [13] Zhang, W., and Deng, X., 2005, "Formulation of a Cohesive Zone Model for a Mode III Crack," *Eng. Fract. Mech.*, **72**, pp. 1818–1829.
- [14] Weisstein, E. W. "Fourier Series," *MathWorld—A Wolfram Web Resource*, <http://mathworld.wolfram.com/FourierSeries.html>

Lateral Motion of an Axially Moving Tape on a Cylindrical Guide Surface

Bart Raeymaekers¹

e-mail: bart@talkelab.ucsd.edu

Frank E. Talke

University of California, San Diego,
Center for Magnetic Recording Research,
9500 Gilman Drive,
La Jolla, CA 92093-0401

The lateral motion of a tape moving axially over a cylindrical guide surface is investigated. The effects of lateral bending stiffness and friction force are studied and the attenuation of lateral tape motion as a function of the guide radius and friction coefficient is determined. Good agreement between numerical predictions and experimental results is observed.

[DOI: 10.1115/1.2723823]

Keywords: magnetic tape, lateral tape motion, bending stiffness, friction

1 Introduction

Lateral tape motion (LTM) is the time-dependent displacement of magnetic tape perpendicular to the tape transport direction. Lateral tape motion causes track misregistration between the read/write head and a previously written track, thereby limiting the track density that can be achieved [1]. In order to increase the track density further, the lateral displacement of tape must be decreased [2].

Several researchers have studied the dynamic behavior and vibrations of a moving web between rollers [3–11]. The effect of guides on the lateral tape motion in a tape path has been studied by only a few researchers. Ono [12] described the lateral displacement of an axially moving string on a cylindrical guide surface. Bending stiffness was not included in his model. He showed that the lateral motion was governed by a second order differential equation similar to that for one-dimensional heat flow. More recently, O'Reilly and Varadi [13] studied the dynamics of a closed loop of inextensible string which is undergoing an axial motion and of which one point is in contact with a singular supply of momentum. Taylor and Talke [14] investigated the interactions between rollers and flexible tape and showed that friction between the tape and the roller affects the lateral displacement of tape.

In this paper we have studied the lateral motion of magnetic tape as it moves over a stationary guide. We have included the effect of bending stiffness, since the area moment of the tape for the transverse direction is very large.

2 Theoretical Study

2.1 Lateral Tape Motion Including Bending Stiffness. In Figs. 1(a) and 1(b), a section of tape is shown as it moves over a

cylindrical guide surface. Vector $\mathbf{M}$ denotes the position of a point on the centerline of the tape between boundary points s_1 and s_2 in a fixed cylindrical coordinate system $r\varphi z$ with unit vectors $\mathbf{e}_r, \mathbf{e}_\varphi, \mathbf{k}$.

To model the system, we make the following simplifying assumptions:

- 1) The magnitude of the friction force is proportional to the normal force, $F = \mu N$.
- 2) The lateral tape motion is sufficiently small so that only the first-order terms of the lateral displacement and the derivatives with respect to space or time must be retained.
- 3) Deformation of the tape is assumed small in the lateral direction (θ small). Bernoulli-Euler assumptions are employed.
- 4) The tape moves at a constant speed $\bar{v}_0$. Hence, $\partial v_0 / \partial s = 0$.

The forces acting on a small element ds of tape are the tension force $\mathbf{T}$, the normal force $\mathbf{N}$, the friction force $\mathbf{F}$, and the shear force $\mathbf{V}$. Expressing $d\mathbf{T}$, $\mathbf{N}$, and $\mathbf{F}$ in r, φ, z coordinates, we obtain

$$d\mathbf{T}(s) = \left[-Ta \left(\frac{\partial \varphi}{\partial s} \right)^2 ds \right] \mathbf{e}_r + \left[\left(\frac{\partial T}{\partial s} a \frac{\partial \varphi}{\partial s} + Ta \frac{\partial^2 \varphi}{\partial s^2} \right) ds \right] \mathbf{e}_\varphi + \left[\left(\frac{\partial T}{\partial s} \frac{\partial z}{\partial s} + T \frac{\partial^2 z}{\partial s^2} \right) ds \right] \mathbf{k} \quad (1)$$

$$\mathbf{N} = N \mathbf{e}_r \quad (2)$$

and

$$\mathbf{F} = -\frac{Na}{v_0} \mu_\varphi \left[\frac{\partial \varphi}{\partial t} + \frac{\partial \varphi}{\partial s} \frac{\partial \varphi}{\partial s} \right] \mathbf{e}_\varphi - \frac{N}{v_0} \mu_z \left[\frac{\partial z}{\partial t} + \frac{\partial z}{\partial s} \frac{\partial z}{\partial s} \right] \mathbf{k} \quad (3)$$

where s is the spatial coordinate along the tape centerline, a is the radius of the cylindrical guide, and μ_φ and μ_z are the friction coefficient in the circumferential direction and the z direction, respectively. T is the magnitude of the tape tension vector and t represents the time. In addition, $d\mathbf{V}$ can be shown to be (Appendix)

$$d\mathbf{V}(s) = \left[2EI \frac{\partial^5 z}{\partial s^5} ds \right] \mathbf{e}_\varphi - \left[EI \frac{\partial \varphi}{\partial s} \frac{\partial^4 z}{\partial s^4} ds \right] \mathbf{e}_r - \left[EI \frac{\partial \varphi}{\partial s} a \frac{\partial^3 z}{\partial s^3} \left(\frac{\partial z}{\partial s} + \frac{\partial \varphi}{\partial s} \right) ds \right] \mathbf{k} \quad (4)$$

Using equilibrium of forces, we have

$$\rho w ds \frac{d^2 \mathbf{M}}{dt^2} = d\mathbf{T} + \mathbf{F} ds + N ds + d\mathbf{V} \quad (5)$$

where ρ is the tape density and w is the tape width.

Introducing (1)–(4) in (5), one can obtain equations for the $\mathbf{e}_r$, $\mathbf{e}_\varphi$, and $\mathbf{k}$ directions. We assume that v_0 is constant. In addition, we assume that the local stress change in the tape can be neglected and that the LTM is small. Hence, we can write that

$$\frac{\partial \varphi}{\partial t} = \frac{\partial^2 \varphi}{\partial t^2} = 0$$

$$\frac{\partial^2(a\varphi)}{\partial s^2} = -\sin \theta \frac{\partial \theta}{\partial s} \ll 1$$

$$\frac{\partial z}{\partial s} = \sin \theta \ll 1$$

$$\frac{\partial(a\varphi)}{\partial s} = \cos \theta \cong 1$$

For a typical tape transport speed $\bar{v}_0 = 4$ m/s and tension $T = 1$ N, we can neglect the inertia terms $\rho w \bar{v}_0^2$ versus the tension T for a

¹Corresponding author.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received February 19, 2006; final manuscript received February 8, 2007. Review conducted by Oliver M. O'Reilly.

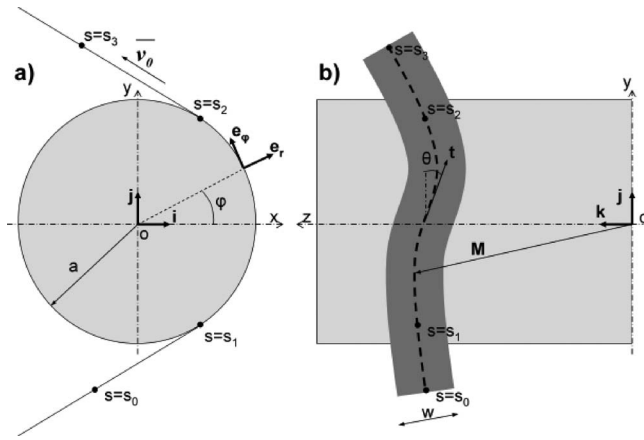


Fig. 1 Tape on a cylindrical guide

9 μm thick magnetic tape (mylar-PET) with $\rho=0.012 \text{ kg/m}^2$ and $w=12.7 \text{ mm}$. Hence, we obtain the equation of motion of an axially moving tape on a cylindrical guide surface as

$$-EI \frac{\partial^4 z}{\partial s^4} + T \frac{\partial^2 z}{\partial s^2} - \frac{\mu_z T}{a} (1 - \nu) \frac{\partial z}{\partial s} - \frac{\mu_z T}{v_0 a} \frac{\partial z}{\partial t} = 0 \quad (6)$$

$\nu = \mu_\phi / \mu_z$ is the ratio of the friction coefficients in the circumferential and vertical direction, respectively. E represents the Young's modulus and I is the area moment. In Eq. (6), the fourth order derivative represents the bending stiffness of the tape. If the bending stiffness is neglected, i.e., $I=0$, Eq. (6) becomes identical to the equation of motion for a string derived by Ono [12].

3 Numerical Solution

To solve Eq. (6), an implicit Euler finite difference scheme was implemented. As illustrated in Figs. 1(a) and 1(b), the tape makes contact with the cylindrical guide at point s_1 and comes off the cylindrical guide at point s_2 . We assume that the tape is wound on a reel with zero run-out at s_3 , i.e., $z(s_3, t)=0$. At s_0 , the lateral displacement $f_0(t)$ is assumed to be known from experimental measurements, i.e., $z(s_0, t)=f_0(t)$. In addition we postulate that the tape moves like a rigid body between s_0 and s_1 and s_2 and s_3 , where it is not supported by the guide. The distances $l_1=|s_0 s_1|$ and $l_3=|s_2 s_3|$ are taken much shorter than the distance $l_2=|s_1 s_2|$.

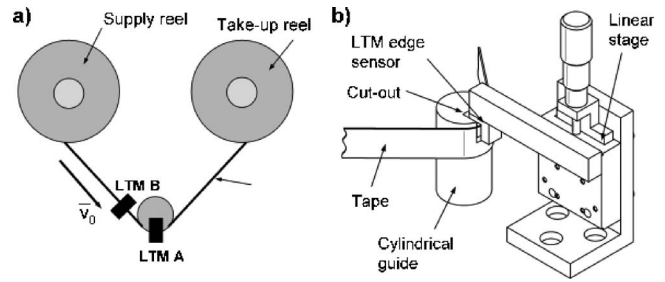


Fig. 2 Experimental apparatus to measure LTM on a cylindrical guide

The solution of Eq. (6) for the domain $s_1 s_2$ requires four boundary conditions. The slope of the tape at $s=s_1$ must be equal to the slope at $s=s_0$ and the slope of the tape at $s=s_2$ must be equal to the slope at $s=s_3$. Thus,

$$l_1 \frac{\partial z}{\partial s} - (z - f_0(t)) = 0 \quad \text{at } s = s_1 \quad (7)$$

and

$$l_3 \frac{\partial z}{\partial s} + z = 0 \quad \text{at } s = s_2 \quad (8)$$

In addition, the curvature at s_1 and s_2 has to be zero to insure a smooth tape path, i.e.,

$$\left. \frac{\partial^2 z}{\partial s^2} \right|_{s_1, s_2} = 0 \quad (9)$$

4 Experimental Validation and Discussion

4.1 Apparatus. To allow a comparison of numerical results with experimental data, the apparatus shown in Fig. 2 was used. The setup consists of a tape moving from a supply reel to a take-up reel at $v_0=4 \text{ m/s}$ over a cylindrical guide with a radius of $a=10 \text{ mm}$. The nominal tape tension $T=1 \text{ N}$. The run-out of the supply reel creates LTM in the tape path. A cutout was provided in the cylinder, for the placement of a lateral tape motion edge sensor [1] (LTM A in Fig. 2(a)), mounted on a linear microstage for vertical positioning. Additionally, we measured the LTM as close as possible to the point where the tape makes contact with the guide (LTM B in Fig. 2(b)) and used this as input (boundary condition) $z(s_0, t)=f_0(t)$ for our numerical model.

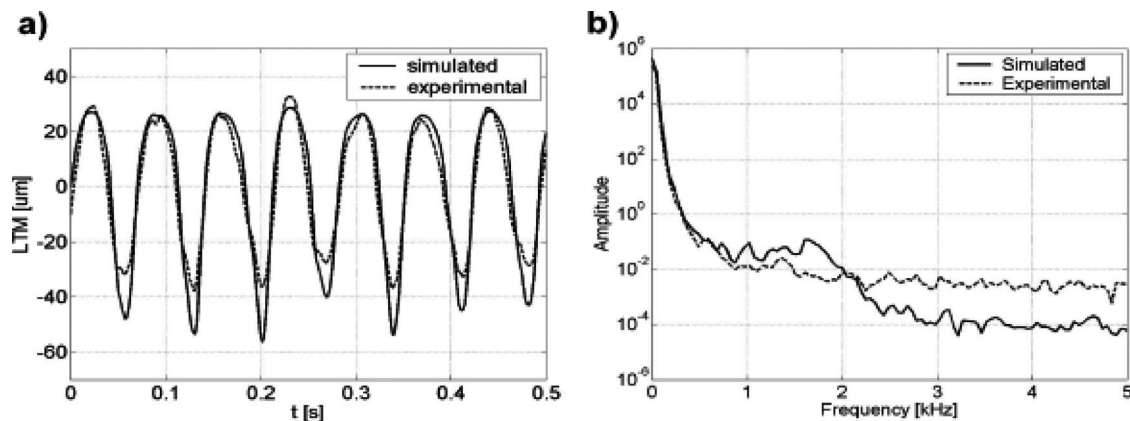


Fig. 3 Comparison of experimental measurements and numerical predictions in the middle of the cylindrical guide (a) in time domain and (b) in frequency domain

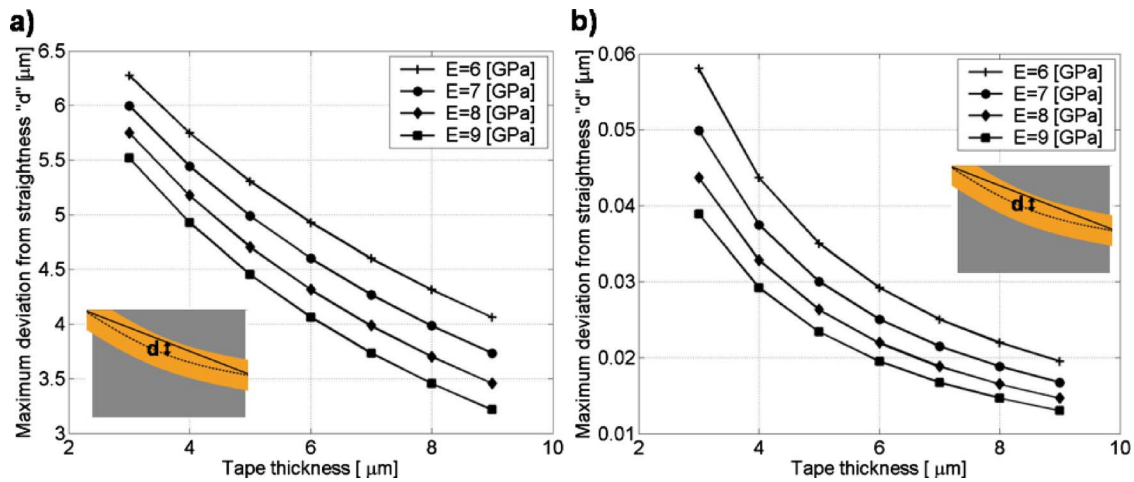


Fig. 4 Maximum deviation from straightness for (a) a guide with radius 100 mm and (b) a guide with radius 10 mm ($s_0=20 \mu\text{m}$ and $s_3=0 \mu\text{m}$, $\mu_\varphi=\mu_z=0.3$)

4.2 Experimental Results. To verify our model, we have compared the values of numerically calculated lateral tape motion in the middle of the cylindrical guide with experimentally measured LTM values at the same position. Figure 3(a) shows simulated and experimentally measured values of lateral tape displacement in the middle of the cylindrical guide, while Fig. 3(b) shows the simulated and experimentally determined frequency spectrum at the same position. In our simulation we have used the following parameters: $E=7 \text{ GPa}$, $w=12.7 \text{ mm}$, $\rho=0.012 \text{ kg/m}^2$, $v_0=4 \text{ m/s}$, $\rho=0.012 \text{ kg/m}^2$, $a=10 \text{ mm}$, $T=1 \text{ N}$, $\nu=1$, and tape thickness $b=9 \mu\text{m}$. These values represent typical values for state-of-the-art magnetic tapes [15,16]. From Fig. 3 we observe good agreement between experimental measurements and numerical results, especially in the low frequency region. The increased deviation between experimental measurements and the numerical predictions for increasing frequencies is most likely related to the presence of the cutout in the stationary guide for positioning of the lateral tape displacement sensor.

4.3 Bending Stiffness. The path of a tape with bending stiffness on a cylindrical guide is different from the path followed by a string. Figure 4 shows the maximum displacement d that the tape trajectory deviates from a straight line, versus tape thickness, as a function of Young's modulus.

The results in Fig. 4(a) are for a cylindrical guide with a radius of 100 mm, while the results in Fig. 4(b) are for a cylindrical

guide with a radius of 10 mm. We note that the vertical plotting scale for both pictures is different. We observe that the maximum displacement d decreases with increasing Young's modulus and increasing tape thickness, i.e., d decreases with increasing bending stiffness. In addition, the maximum deviation d decreases with decreasing guide radius.

4.4 Effect of a Guide in the Tape Path. In order to investigate the effect of a guide on the lateral tape displacement, we have calculated the amplitude ratio of the lateral tape displacement at positions s_1 and s_2 . A sine wave with constant frequency was applied as an input at $s=s_1$ and the output at $s=s_2$ was simulated. For each frequency, the ratio of output to input amplitude was calculated. Figure 5(a) shows the amplitude ratio for a constant friction coefficient $\mu_\varphi=\mu_z=0.3$ as a function of frequency and guide radius a , while Fig. 5(b) shows the amplitude ratio for a guide radius of 200 mm for different friction coefficients.

We observe that the amplitude ratio decreases with increasing guide diameter and with increasing friction coefficient. By positioning guides in the tape path at specific positions, high frequency LTM could be filtered out before the tape moves over the magnetic read/write head. Lateral tape motion above 500 Hz cannot be followed by the servo system of the magnetic read/write head. When the amplitude of high frequency lateral tape motion is larger than 10% of the track width, read/write errors are likely to

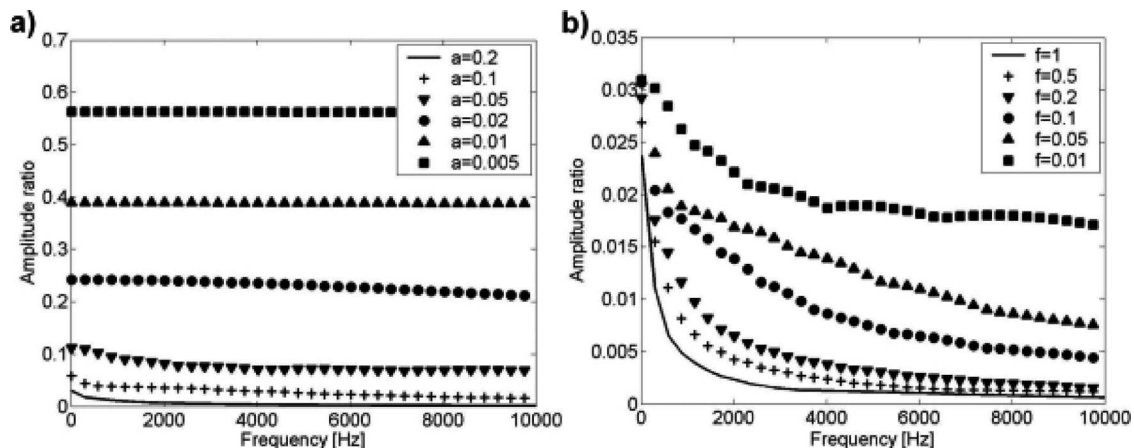


Fig. 5 Amplitude ratio $z(s_2)/z(s_1)$ (a) for different guide radii and (b) for different friction coefficients

occur. Thus, reducing high frequency lateral tape motion before it reaches the head is desirable since it would allow narrower tracks, resulting in higher track density.

The tape model with bending stiffness also has an important application in the tape slitting process where lateral tape displacement has to be minimized in order to manufacture tapes with sufficiently “straight” edges for servo track writing. If the radius of the guides that transport the tape is increased, the amplitude of the LTM is attenuated more strongly and the edge quality of the tape produced during the slitting process should be improved. Thus, optimization of the tape path by increasing the diameter of the guides should be considered in tape slitting machines to improve the quality of future tapes.

5 Conclusion

The results obtained in this paper show that:

1. Bending stiffness is an important parameter in describing the lateral displacement of a tape on a cylindrical guide surface. When modeling a tape, shear forces have to be included.
2. The effect of a cylindrical guide in the tape path can be characterized by the amplitude ratio of output and input lateral tape displacement. The amplitude ratio depends on the friction coefficient and on the contact length between tape and guide surface, i.e., the guide radius and the wrap angle. An increase in the guide radius of the tape guide or an increase in the friction coefficient between the guide surface and the tape improves the damping of both low and high LTM frequencies.
3. The bending stiffness of the tape affects its trajectory over the cylindrical guide. As the bending stiffness increases, the tape trajectory deviates less from a straight line.

Appendix

The shear force vector V with respect to the r, φ, z coordinate system can be written as

$$V = V\mathbf{v} = V \frac{\partial z}{\partial s} \mathbf{e}_\varphi - V \frac{a}{\partial s} \frac{\partial \varphi}{\partial s} \mathbf{k} \quad (\text{A1})$$

The unit shear vector $\mathbf{v}$ can be written as

$$\mathbf{v} = \frac{\partial z}{\partial s} \mathbf{e}_\varphi - \frac{a}{\partial s} \frac{\partial \varphi}{\partial s} \mathbf{k} \quad (\text{A2})$$

thus,

$$\frac{\partial \mathbf{v}}{\partial s} = \frac{\partial^2 z}{\partial s^2} \mathbf{e}_\varphi - \frac{\partial z}{\partial s} \frac{\partial \varphi}{\partial s} \mathbf{e}_r - \frac{a}{\partial s^2} \frac{\partial^2 \varphi}{\partial s^2} \mathbf{k} \quad (\text{A3})$$

Combining (A1)–(A3) with $dV(s) = \mathbf{v} \frac{\partial V}{\partial s} ds + V \frac{\partial \mathbf{v}}{\partial s} ds$, $V = \frac{\partial C}{\partial s}$, and $C = EI \frac{\partial^2 z}{\partial s^2}$, yields Eq. (4).

References

- [1] Taylor, R. J., Strahle, P., Stahl, J., and Talke, F. E., 2000, “Measurement of Cross-Track Motion of Magnetic Tapes,” *J. Inf. Storage Process. Syst.*, **2**, pp. 255–261.
- [2] Richards, D. B., and Sharrock, M. P., 1998, “Key Issues in the Design of Magnetic Tape for Linear Systems of High Track Density,” *IEEE Trans. Magn.*, **34**(4), pp. 1878–1882.
- [3] Norwood, R. E., 1969, “Effects of Bending Stiffness in Magnetic Tape,” *IBM J. Res. Dev.*, **13**(2), pp. 205–208.
- [4] Shelton, J. J., and Reid, K. N., 1971, “Lateral Dynamics of an Idealized Moving Web,” *ASME J. Dyn. Syst., Meas., Control*, **93**, pp. 187–192.
- [5] Shelton, J. J., and Reid, K. N., 1971, “Lateral Dynamics of a Real Moving Web,” *ASME J. Dyn. Syst., Meas., Control*, **93**, pp. 180–186.
- [6] Wickert, J. A., and Mote, Jr., C. D., 1990, “Classical Vibration Analysis of Axially Moving Continua,” *ASME J. Appl. Mech.*, **57**, pp. 738–744.
- [7] Young, G. E., and Reid, K. N., 1993, “Lateral and Longitudinal Dynamic Behavior and Control of Moving Webs,” *ASME J. Dyn. Syst., Meas., Control*, **115**, pp. 309–317.
- [8] Lee, S., and Mote, Jr., C. D., 1999, “Wave Characteristics and Vibration Control of Translating Beams by Optimal Boundary Damping,” *ASME J. Appl. Mech.*, **121**, pp. 18–25.
- [9] Garziera, R., and Amabili, M., 2000, “Damping Effect of Winding on the Lateral Vibration of Axially Moving Tapes,” *ASME J. Vib. Acoust.*, **122**, pp. 49–53.
- [10] Benson, R. C., 2002, “Lateral Dynamics of a Moving Web With Geometrical Imperfection,” *ASME J. Dyn. Syst., Meas., Control*, **124**, pp. 25–34.
- [11] Yerashunas, J. B., De Abreu-Garcia, J. A., and Hartley, T. T., 2003, “Control of Lateral Motion in Moving Webs,” *IEEE Trans. Control Syst. Technol.*, **11**, pp. 684–693.
- [12] Ono, K., 1979, “Lateral Motion of an Axially Moving String on a Cylindrical Guide Surface,” *ASME J. Appl. Mech.*, **46**, pp. 905–912.
- [13] O'Reilly, O. M., and Varadi, P. C., 2004, “On Some Peculiar Aspects of Axial Motions of Closed Loops of String in the Presence of a Singular Supply of Momentum,” *ASME J. Appl. Mech.*, **71**, pp. 541–545.
- [14] Taylor, R. J., and Talke, F. E., 2005, “Investigation of Roller Interactions With Flexible Tape Medium,” *Tribol. Int.*, **38**, pp. 599–605.
- [15] Berry, B. S., and Pritchett, W. C., 1988, “Elastic and Viscoelastic Behaviour of a Magnetic Recording Tape,” *IBM J. Res. Dev.*, **32**(5), pp. 682–694.
- [16] Bhushan, B., and Tao, Z., 2004, “Mechanical, Hygroscopic, and Thermal Properties of Metal Particle and Metal Evaporated Tapes and Their Individual Layers,” *J. Appl. Polym. Sci.*, **92**, pp. 1319–1345.

Discussion: “A Mathematical Model for Frictional Elastic-Plastic Sphere-on-Flat Contacts at Sliding Incipient” (Chang, L., and Zhang, H., 2007, ASME J. Appl. Mech., 74, pp. 100–106)

Izhak Etsion

Fellow ASME

Dept. of Mechanical Engineering,

Technion,

Haifa 32000, Israel

e-mail: etsion@tx.technion.ac.il

The authors presented a model for sliding inception that is essentially based on the 1949 Mindlin approach (see Ref. [1]) for frictional contact with extension to elastic-plastic normal loading. According to this concept the interfacial shear stress is assumed to be proportional to the contact pressure until it reaches a limiting value that is related to the shear strength of the sphere bulk (see Eq. (1)). In other words, the Mindlin approach assumes a “local Coulomb friction law” that requires an input of a certain proportionality constant that relates the shear stress distribution to the normal pressure distribution in the contact interface. It may lead to unrealistic situations in which the local equivalent von Mises stress can exceed the yield strength of the sphere material. Another shortcoming of this concept is that sliding inception always occurs in the contact interface regardless of the level of normal loading. This would contradict the well known phenomenon of

material transfer, which is associated with high normal loading of adhesive frictional contacts and requires slip below the contact interface.

A completely different approach to frictional contact was recently presented in a series of papers by Brizmer et al. assuming full stick contact condition (see Refs. [2–4]). These three papers deal with all the aspects of the present paper, namely, the critical interference [2], the elastic-plastic loading regime [3], and the sliding inception [4], respectively. The full stick contact condition captures very well the concept of an adhesive joint formed in the contact interface [5]. It does not require any assumption of a proportionality constant that relates local shear stress and local pressure, and therefore it never violates the von Mises yield criterion. Furthermore, it allows analyzing the sliding inception as a failure mechanism and it utilizes first principles to predict the sliding inception and to obtain the resulting corresponding static friction coefficient and junction growth. The full stick contact condition does not impose slip at the contact interface and hence, allows for the possibility of material transfer under severe normal loads. Additionally the results of Ref. [4] correlate well with some preliminary experimental results obtained by Ovcharenko et al. [5].

References

- [1] Mindlin, R. D., 1949, “Compliance of Elastic Bodies in Contact,” *ASME J. Appl. Mech.*, **16**, pp. 259–268.
- [2] Brizmer, V., Kligerman, Y., and Etsion, L., 2006, “The Effect of Contact Conditions and Material Properties on the Elasticity Terminus of a Spherical Contact,” *Int. J. Solids Struct.*, **43**, pp. 5736–5749.
- [3] Brizmer, V., Zayit, Y., Kligerman, Y., and Etsion, L., 2006, “The Effect of Contact Conditions and Material Properties on Elastic-Plastic Spherical Contact,” *Journal of Mechanics of Solids and Structures*, **1**, pp. 865–879.
- [4] Brizmer, V., Kligerman, Y., and Etsion, L., 2007, “Elastic-Plastic Spherical Contact Under Combined Normal and Tangential Loading,” *Tribol. Lett.*, **25**, pp. 61–70.
- [5] Ovcharenko, A., Halperin, G., Etsion, L., and Varenberg, M., 2006, “A Novel Test Rig for *In Situ* and Real Time Optical Measurement of the Contact Area Evolution During Pre-Sliding of a Spherical Contact,” *Tribol. Lett.*, **23**, pp. 55–63.