# Transactions of the ASME®

# Journal of Applied Mechanics

Published Quarterly by The American Society of Mechanical Engineers

VOLUME 67 • NUMBER 4 • DECEMBER 2000



Technical Editor, **LEWIS T. WHEELER**

**APPLIED MECHANICS DIVISION**

Executive Committee
(Chair) **T. J. R. HUGHES**
**D. KRAJCINOVIC**
**S. KYRIAKIDES**
**P. D. SPANOS**
**M. C. BOYCE**
Associate Technical Editors
**J. R. BARBER (2003)**
**R. C. BENSON (2003)**
**A. A. FERRI (2003)**
**H. GAO (2003)**
**J. W. JU (2001)**
**V. K. KINRA (2002)**
**D. A. KOURIS (2002)**
**A. K. MAL (2001)**
**B. M. MORAN (2002)**
**A. NEEDLEMAN (2001)**
**M. ORTIZ (2001)**
**N. C. PERKINS (2002)**
**M.-J. PINDERA (2003)**
**K. T. RAMESH (2003)**
**K. RAVI-CHANDAR (2003)**
**W. S. SARIC (2003)**
**D. A. SIGINER (2003)**
**T. E. TEZDUYAR (2003)**
**N. TRIANTAFYLLIDIS (2003)**

**BOARD ON COMMUNICATIONS**

Chairman and Vice-President
**R. K. SHAH**

**OFFICERS OF THE ASME**

President, **J. R. PARKER**
Executive Director, **D. L. BELDEN**
Treasurer, **J. A. MASON**

**PUBLISHING STAFF**

Managing Director, Engineering
**CHARLES W. BEARDSLEY**
Director, Technical Publishing
**PHILIP DI VIETRO**
Managing Editor, Technical Publishing
**CYNTHIA B. CLARK**
Managing Editor, Transactions
**CORNELIA MONAHAN**
Production Coordinator
**JUDITH SIERANT**
Production Assistant
**MARISOL ANDINO**



**Transactions of the ASME, Journal of Applied Mechanics** (ISSN 0021-8936) is published quarterly (Mar., June, Sept., Dec.) The American Society of Mechanical Engineers, Three Park Avenue, New York, NY 10016. Periodicals postage paid at New York, NY and additional mailing office. 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 © 2000** 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.


**INDEXED** by Applied Mechanics Reviews and Engineering Information, Inc. Canadian Goods & Services Tax Registration #126148048.


## TECHNICAL PAPERS


645 **On the Propagation of Lüders Bands in Steel Strips**
S. Kyriakides and J. E. Miller

655 **Common Errors on Mapping of Nonelliptic Curves in Anisotropic Elasticity**
T. C. T. Ting

658 **On the Interior Stress Problem for Elastic Bodies**
J. Helsing

663 **A Stress and Relative Density-Dependent Dynamic Compliance Spectra Model of the Creep Response of Microcellular Polycarbonate**
W. D. Armstrong and V. Kumar

667 **Effective Elastic Properties of Porous Materials With Randomly Dispersed Pores: Finite Deformation**
V. A. Levin, V. V. Lokhin, and K. M. Zingerman

671 **Development of One-Dimensional Models for Elastic Waves in Heterogeneous Beams**
H. Murakami and J. Yamakawa

685 **Asymptotic Stability Analysis for Sheet Metal Forming—Part I: Theory**
W. Scherzinger and N. Triantafyllidis

691 **Asymptotic Stability Analysis for Sheet Metal Forming—Part II: Application**
W. Scherzinger, E. Chu, and N. Triantafyllidis

697 **Modified Lagrange Method to Analyze Problems of Sliding and Rolling**
A. Rosen

705 **On the General Solution for Piezothermoelasticity for Transverse Isotropy With Application**
W. Q. Chen

712 **Dynamic Wetting: Displacement of the Contact Line Induced by Air Pressure Oscillations in Its Close Vicinity**
P. Woehl and P. Bourgin

717 **Stability Analysis of an Initially Inclined Ship in Following Sea**
S. K. Lee

720 **Stick-Slip, Imposition-Removal of Constraints and the Spinning Ball Problem**
S. Djerassi

727 **The Fiber Composite With Nonlinear Interface—Part I: Axial Tension**
A. J. Levy

733 **The Fiber Composite With Nonlinear Interface—Part II: Antiplane Shear**
A. J. Levy

740 **Bond-Induced Longitudinal Fracture in Reinforced Concrete**
M. Ghandehari, S. Krishnaswamy, and S. Shah

749 **Reactive Flow in a Porous Medium: Formulation for Spatially Periodic Hexagonally Packed Cylinders**
G. Rajesh, R. B. Bhagat, and K. A. Fichthorn

758 **Dynamic Plastic Response of a Submarine Oil Pipeline to an Underwater Explosion Bubble**
Z. Zong and K. Y. Lam



(**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.

(Contents continued)

Journal of Applied Mechanics Volume 67, Number 4 DECEMBER 2000

763 Energy-Based Stochastic Estimation for Nonlinear Oscillators With Random Excitation
J. B. Roberts and M. Vasta

772 Vibrations of a Taut Cable With an External Damper
S. Krenk

777 Initial Post-Buckling and Growth of a Circular Delamination Bridged by Nonlinear Fibers
S. Li, J. Nie, J. Qian, Y. Huang, and Y. Hu

785 Structural Dynamic Effects on Interface Response: Formulation and Simulation Under Partial Slipping Conditions
E. J. Berger, M. R. Begley, and M. Mahajani

793 Deformation, Stress State, and Thermodynamic Force for a Transforming Spherical Inclusion in an Elastic-Plastic Material
F. D. Fischer and E. R. Oberaigner

797 On Symmetrizable Systems of Second Kind
S. Adhikari

803 Localization-Induced Band and Cohesive Model
S. Hao, W. K. Liu, and D. Qian

813 Bending and Extension of Thin-Walled Composite Beams of Open Cross-Sectional Geometry
O. Rand

819 Use of Graded Finite Elements to Model the Behavior of Nonhomogeneous Materials
M. H. Santare and J. Lambros

823 Solution of the Moving Mass Problem Using Complex Eigenfunction Expansions
K.-Y. Lee and A. A. Renshaw

## BRIEF NOTES

828 On the Relation Between the *L*-Integral and the Bueckner Work-Conjugate Integral
J. P. Shi, X. H. Liu, and J. M. Li

829 A Note on the Driving Traction Acting on a Propagating Interface: Adiabatic and Non-Adiabatic Processes of a Continuum
R. Abeyaratne and J. K. Knowles

831 Characterizing Damping and Restitution in Compliant Impacts via Modified K-V and Higher-Order Linear Viscoelastic Models
E. A. Butcher and D. J. Segalman

834 Finite-Amplitude Elastic Instability of Plane-Poiseuille Flow of Viscoelastic Fluids
R. E. Khayat and N. Ashrafi

837 In-Plane Gravity Loading of a Circular Membrane
R. O. Tejeda, E. G. Lovell, and R. L. Engelstad

839 Free Vibration of a Spinning Stepped Timoshenko Beam
S. D. Yu and W. L. Cleghorn

841 Penetration Experiments With Limestone Targets and Ogive-Nose Steel Projectiles
D. J. Frew, M. J. Forrestal, and S. J. Hanchak

846 Eccentric Crack in a Piezoelectric Strip Under Electromechanical Loading
Jeong Woo Shin, Soon Man Kwon, and Kang Yong Lee

## ANNOUNCEMENTS AND SPECIAL NOTES

848 Author Index

854 Information for Authors

855 New Reference Format

856 Solicitation for Nominations for the J.P. Den Hartog aand N.O. Myklestad Awards

# On the Propagation of Lüders Bands in Steel Strips

**S. Kyriakides**
Fellow ASME

**J. E. Miller**

Research Center for Mechanics of Solids,
Structures & Mechanics,
WRW 110,
The University of Texas at Austin,
Austin, TX 78712

*The initiation and propagation of Lüders-type localized deformation in thin, fine grained steel strips in tension is studied through combined experimental and analytical efforts. Purely elastic deformation is terminated (upper yield stress) by localized deformation which tends to initiate along preferred directions. The strain level associated with this material instability is limited to two to five percent. When this strain level is achieved locally, the instability propagates via inclined fronts which separate coexisting regions of essentially elastic and plastically deformed materials. Under displacement controlled stretching, one or two fronts propagate in a steady-state manner (lower yield stress). The propagation of one and two fronts are simulated numerically using finite element models in which the material is modeled as a finitely deforming elastoplastic solid with an up-down-up nominal stress-strain response. The simulations capture the major events observed in the experiments such as the initiation process, the propagation of inclined fronts, kinking of the strip and the build up of moments, and the periodic straightening and moment reduction through transient events. This confirms that structural effects play a major role in the evolution of observed events.* [S0021-8936(00)01604-4]

## 1 Introduction

For several metals, the transition from elastic to plastic deformation is characterized by a material instability known as *Lüders strain*. The macroscopic effect of the instability is inhomogeneous deformation. Figure 1 shows a typical response recorded in a uniaxial test on a low carbon steel, which is the best known example of this class of behavior. Localized plastic deformation starts with a sudden drop in stress. Under displacement control, the deformation spreads along the length of the specimen while the stress remains essentially constant. When the whole test section has been deformed to the Lüders strain ($\Delta\epsilon_L$), the material hardens and the specimen returns to homogeneous deformation. The initial stress peak (*upper yield stress* $\sigma_U$), as well as the level and extent of the stress plateau (*lower yield stress* $\sigma_L$ and $\Delta\epsilon_L$, respectively) depend on the grain size (the finer the grain the higher the values of $\sigma_U$, $\sigma_L$, and $\Delta\epsilon_L$) and on alloy content. They are also sensitive to temperature and to the rate of loading ([1]).

Two main theories of the microscopic events behind the instability have survived scientific scrutiny. Cottrell and Bilby [2] attributed the upper yield stress to the pinning of dislocations by carbon and nitrogen atoms which naturally tend to form ''atmospheres'' around them. They postulated that initial yielding requires a higher stress in order to pull the dislocations out of their atmospheres. Once released, the dislocations can be moved by a lower stress. The pinning effect of interstitial impurities is widely accepted for several reasons including the reappearance of Lüders strain following mild heat treatment which allows these atoms to migrate and repin the dislocations (*strain aging*). The second theory, due to Johnston and Gilman [3], attributes the load drop to multiplication of dislocations. As their number increases, the stress required to move them decreases. Indeed, simple considerations of dislocation motion expanded upon in Hahn [4] showed that an up-down-up stress-strain response like ones measured in LiF crystals could be derived from such a premise.

Shaw and Kyriakides [5,6], in their studies of the evolution of stress-induced transformations in nearly equiatomic NiTi wires ([7]) and strips ([5,6]), showed that for some temperature regimes, austenite $\rightleftharpoons$ martensite transformations result in inhomogeneous deformation. The microscopic reason for this behavior is that at some stress level, the cubic lattice of austenite snaps to the monoclinic lattice of martensite and the reverse. Despite this fundamentally different underlying mechanism from the dislocation governed Lüders strain, many macroscopic events observed in NiTi strips undergoing localized deformation by uniaxial tension matched those described in the literature for low carbon steel strips (e.g., [1,8,9]). These include the speed of the bands, the angle of their fronts, the in-plane kinking of the strips, the formation of stress valleys on the load plateau when two converging fronts coalesce, etc. A proviso for these similarities is that both materials have fine enough grains (~a few microns).

Shaw and Kyriakides [6] concluded that macroscopic geometric effects and the underlying mechanism of continuum strain localization must play an important role in these common phenomena. They modeled the loading part of the response of NiTi as a finitely deforming elastic-plastic solid with a trilinear up-down-up stress-strain response (see also [10]) and were able to reproduce the evolution of events observed in isothermal tests on strips.

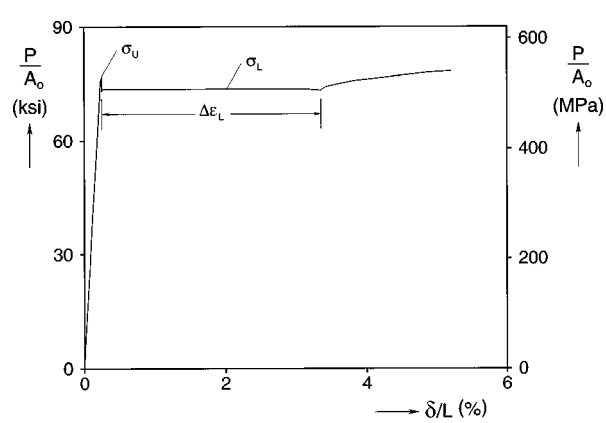



**Fig. 1 Typical force-elongation response of mild steel exhibiting Lüders-type deformation**

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received and accepted by the ASME Applied Mechanics Division, Apr. 18, 2000. Associate Technical Editor: L. T. Wheeler. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

 **Copyright © 2000 by ASME** 

In this paper, the subject is revisited but from the perspective of the Lüders phenomena in steel. The evolution of localized deformation in strips is first examined in new experiments similar to those of Butler [9]. The phenomena are then simulated numerically using a model like that in Shaw and Kyriakides [6] calibrated to the properties of the steel tested.

## 2 Experiments

The experiments were conducted on thin strips machined from plates of fine grained mild steel typical in pipeline applications (grain size ~100-120 gr/mm). The strips were initially 6 in. (152 mm) long, 0.25 in. (6.4 mm) wide, and were ground finished to a thickness of 0.030 in. (0.76 mm). A three-inch (76.2-mm) long ($L$) test section was machined to a width of 0.178 in. (4.52 mm) as shown in Fig. 2($b$). The test section terminates into fillets at the ends ($r=0.067$ in.—1.6 mm). A thin layer (~0.003 in. —75 $\mu$m) of *Stresscoat* was applied to one side of the strip using an airbrush. Stresscoat is a brittle, ceramic coating which shatters when strained sufficiently. Shattering causes a change in the reflectivity of the coating, and by illuminating the surface appropriately, areas of localized deformation can be identified. The shattering strain is not known exactly, but its value is within the strain level of the Lüders phenomenon of this steel. On the back surface, a strain gage was mounted axially at midlength.

The specimen was mounted in a standard uniaxial testing machine using two solid grips, approximately 3.5 in. apart. The experiment was run at a constant crosshead displacement rate ($\dot{\delta}$) of 0.018 in./min which, for homogeneous deformation, corresponds to a strain rate of $\dot{\epsilon}=10^{-4}\,\mathrm{s}^{-1}$. The strain gage strain, the crosshead displacement, and the load were monitored and recorded using a data acquisition system. The full length of the coated side of the test section was monitored with a video camera so that the evolution of the Lüders bands could be later ascertained. The data acquisition system and the video system were run on a common time base so that the measured variables and the optical recordings could be matched during data processing.

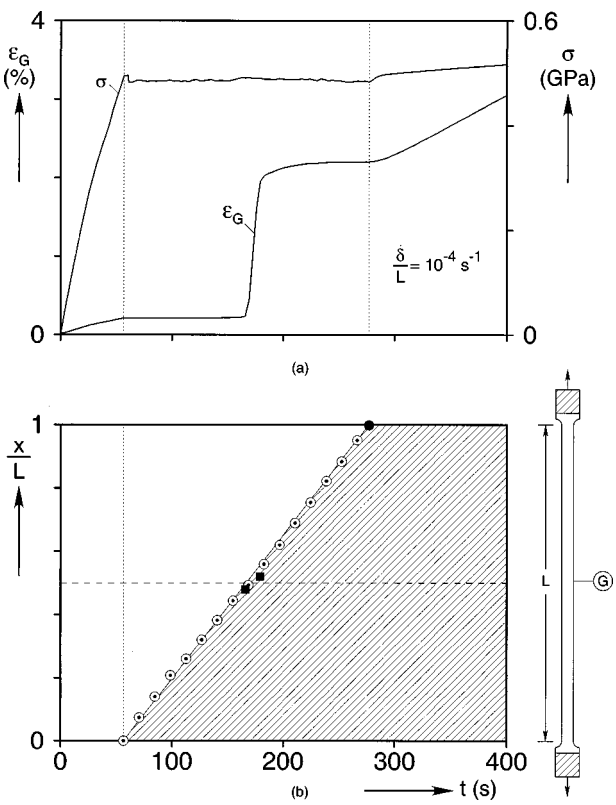



**Fig. 2 Results from uniaxial test on steel strip; (*a*) stress and strain histories recorded, (*b*) *x-t* diagram of evolution of deformation front**

Figure 2($a$) shows the engineering stress ($\sigma$) history recorded. The stress rises to a local maximum at which localized deformation in the form of Lüders strain is initiated. It then maintains a relatively constant level around 70.6 ksi (487 MPa) during which the Lüders bands propagate along the length of the specimen. Approximately 277 s into the test the stress starts to monotonically increase once more. The strain ($\epsilon_G$) history seen by the 0.125 in. (3.2 mm) gage is also shown in Fig. 2($a$). During the initial loading before the stress peak, it records the induced elastic strain. (Note that the small nonlinearities in both the stress and the strain records before the stress peak are due to deformation of the gripped ends; when the stress is plotted against the local strain the response is linear with modulus of approximately $30\times10^3$ ksi—207 GPa). At the onset of localized deformation, the strain at the site of the gage remains unchanged until the deformation front reaches it. At $t=166$ s, the local strain experiences a jump of approximately 1.75 percent over a period of just over 13 s duration followed by a second strain plateau. The recorded strain resumes its increasing trend at the termination of the stress plateau.

A sequence of 17 images from the video record, taken at time intervals of 14 s are shown in Fig. 3. For improved clarity, the images were enhanced using Adobe Photoshop by increasing the contrast between the shattered and intact parts of the stresscoat. Light gray represents material that has experienced Lüders strain while the dark color represents relatively undeformed material. Lüders strain first initiated on the left end of the test section where, despite the presence of fillets, some stress concentration developed. This stress concentration also influences to some extent the initial stress peak recorded. Thus, in such experiments the initiation stress is always a lower bound of the true value.

The Lüders strain propagates from left to right. An inclined front separates the Lüders strained and unstrained regions. The shattering of the Stresscoat did not result in very sharp fronts. In other similar experiments, the deformation process was interrupted with the front partially across the length of the strip which resulted in freezing the front. Measurements performed under a microscope confirmed the inclination to be in the neighborhood of the expected value of 55 deg to the axis of loading ($\alpha$ in Fig. 4). At the interface of the Lüders deformed and undeformed material, the strip kinked by approximately 1.2 deg in the fashion shown in Fig. 4 (angle $\gamma$).

The position of the middle of each front was determined and is plotted in the $x-t$ diagram in Fig. 2($b$). (The last image is not marked, instead the time of the stress valley at the end of the stress plateau is marked with a solid bullet—●). The linear distribution of the points indicates that the front propagated at a nearly constant velocity. Included with a solid line is the front velocity based on the expression postulated by Butler [9] and Shaw and Kyriakides [7]

$$c=\frac{\dot{\delta}}{n\Delta\epsilon_L} \tag{1}$$

where $\dot{\delta}$ is the displacement rate applied to the end of the strip and $\Delta\epsilon_L$ is the Lüders strain (extent of stress plateau taken to be 2.20 percent) and $n$ is the number of fronts ($n=1$ here). The data is in very good agreement with this construction.

Included in the $x-t$ diagram are the positions of the start and termination of the passage of the Lüders band through the strain gage. The two data points are again in good agreement with the construction line and the photographic data. The strain gage signal also indicates that additional deformation of approximately 0.25 percent takes place after the main Lüders strain front passes by. This was found to be repeatable in other experiments performed. If the width of the active front is $w$, then the average strain rate

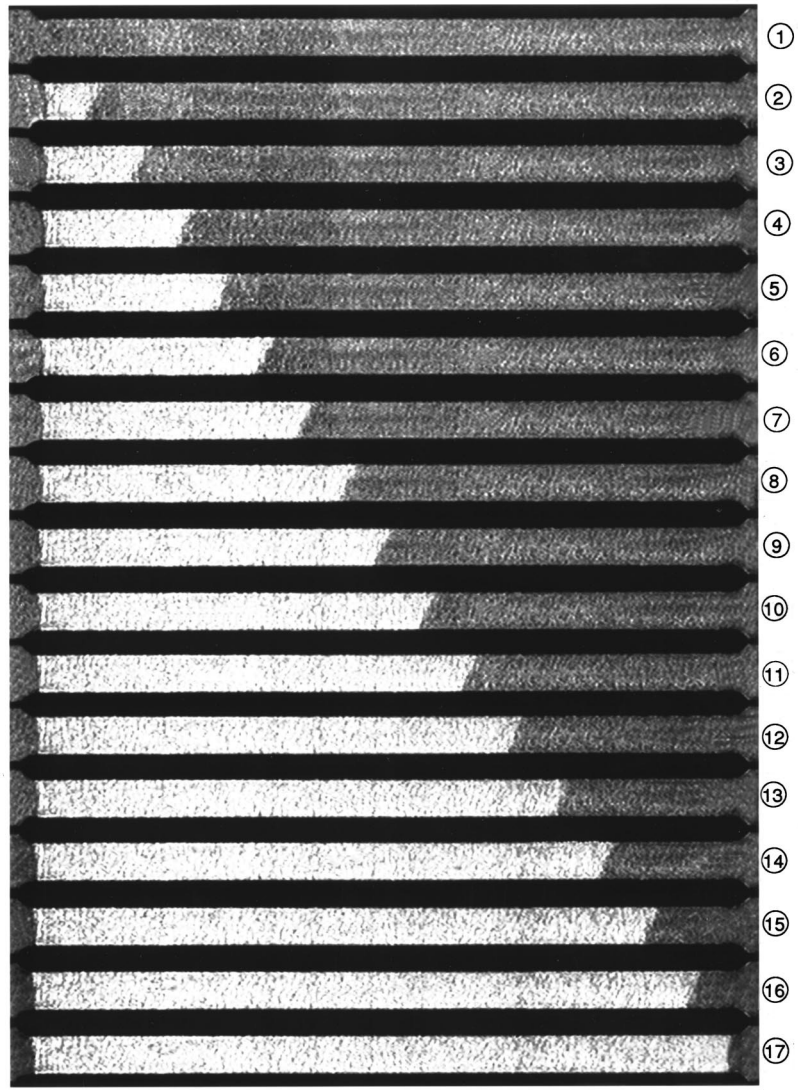



**Fig. 3 Photographic sequence of the evolution of Lüders strain (shown at 14 s intervals)**

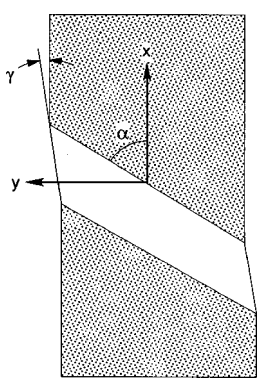



**Fig. 4 Definition of geometric variables of inclined Lüders band**

inside it is $\dot{\epsilon} = \dot{\delta}/w$. Guided by observations of others (e.g., [11]), we suggest that the front width is only a few grains wide and, as a result, the strain rate inside it is rather high. It seems that some dislocations unable to move at the high strain rate continue to move at a slower velocity after the front has passed them.

We point out that in such experiments, two propagating fronts, usually emanating from the ends, are more common than the single propagating front of this experiment. For example, two such fronts are shown in the photograph in Fig. 5 taken from another experiment (see also Figs. 1 and 2 in [9]). Both fronts usually emanate from the ends of the specimen and propagate towards each other at half the present speed of the single front. At the same time, the moment that builds up in the specimen due to

**Fig. 5 Photograph of two coexisting Lüders fronts propagating towards each other**

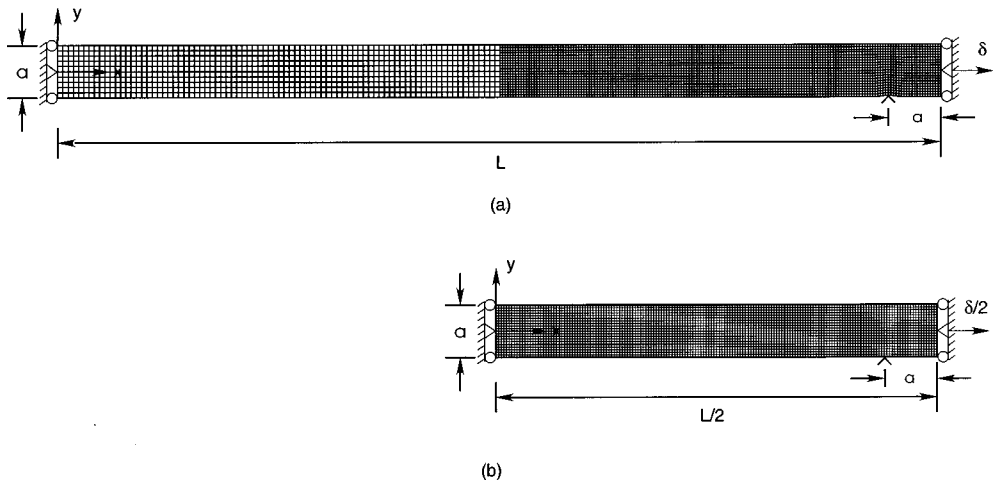



**Fig. 6 Geometry and mesh of model strips for (*a*) single Lüders front propagation and (*b*) two-band propagation**

kinking is reduced. The strip in the test in Fig. 2 had a small taper in its width which in conjunction with small geometric differences in the fillets machined at the two ends, increased the stress concentration at the left end.

## 3 Analysis

The evolution of localized Lüders strain in steel strips was simulated using finite element models similar to those of Shaw and Kyriakides [6]. The length ($L$), the width ($a$), and thickness ($t$) of the strip were the same as in the experiments ($L=100t$, $a=6t$; see Fig. 6). The steel was modeled as a finitely deforming $J_2$-type, elastic-plastic solid with a nonmonotonic engineering stress-strain response drawn with a dashed line in Fig. 7($a$) (the true stress-logarithmic strain version of this is used in the analysis). Although the phenomenon is known to be influenced by rate effects ([1]), in this first attempt at the problem a rate independent model is adopted. The initial elastic modulus and the slope of the positive modulus section following the linear sector with negative slope match the corresponding experimental parts. The level of the upper yield stress was found in the past to affect the stability of the localized deformation patterns that develop. In all our experiments, the value of $\sigma_U$ was masked to some degree by stress concentrations at the fillets which could not easily be avoided. We chose $\sigma_U=79$ ksi (545 MPa) as a reasonable estimate of the true value. The Maxwell stress of this construction was selected to correspond to the plateau stress in Fig. 2. The extent of the Maxwell stress was selected to match the experimental value. In addition, in order to make the events somewhat sharper, the level of the stress knee in the second rising part of the response was increased by 8 ksi (55 MPa) from the experimental value.

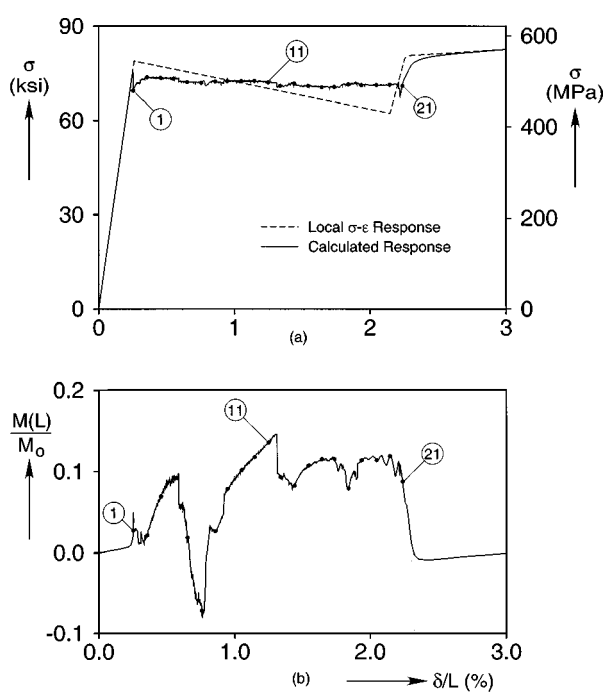



**Fig. 7 Simulation of single Lüders front propagation; (*a*) stress-displacement response and (*b*) moment-displacement response**

The strips were discretized with three-dimensional 20-node (quadratic), isoparametric, reduced integration, brick elements using ABAQUS. The elements had a one-to-one aspect ratio in the $x-y$ plane. This ''isotropic'' mesh was adopted to avoid introducing any preferred directions into the results. The midplane ($z=0$) was assumed to be a plane of symmetry and the half thickness was represented by a single element. The density of the mesh in the $x-y$ plane was patterned after Shaw [5], such that all relevant features of the strain fields were represented without becoming computationally prohibitive.

Two cases will be simulated: The first involves the propagation of a single front and the second the simultaneous propagation of two fronts. Unlike the experiments, the model width was uniform and free of clamping forces. The instability was initiated close to the ends by introducing geometric imperfections in the form of small indentations along one side of the strips (denoted by ''<'' in Fig. 6). A single indentation was introduced in the first case and two in the second (via the antisymmetry assumed about $x=0$). The indentations were symmetric about their middle, with an exponentially decaying amplitude along their length. This type of imperfection allowed the localization to initiate without any artificial directional bias across the width.

**(a) Propagation of a Single Lüders Front.** For this case, two mesh densities were adopted for computational efficiency. On the right half ($L/2\leqslant x\leqslant L$) the density was 20 elements across the width and in the left half ($0\leqslant x\leqslant L/2$) it was 10. The imperfection was $3.33t$ wide and $0.33t$ deep. The following boundary conditions were applied to the ends of the strip:

$$u(0,y,z)=0, \quad v(0,0,z)=0 \quad \text{and}$$
$$u(L,y,z)=\delta, \quad v(L,0,z)=0. \tag{2}$$

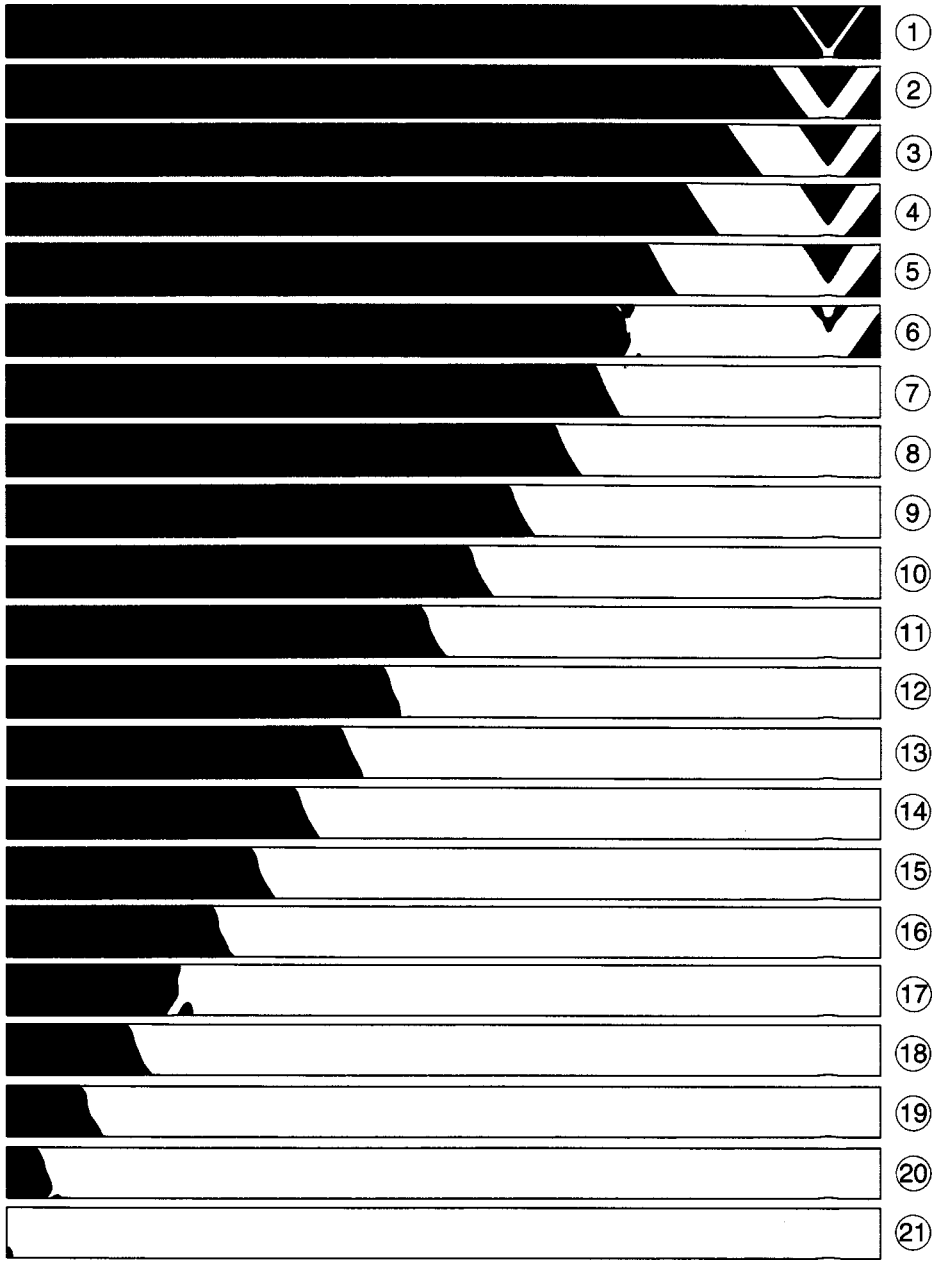



**Fig. 8 Sequence of black-white deformed configurations corresponding to responses in Fig. 7. White depicts zones with axial strain higher than 1.2 percent and black depicts zones lower than 1.2 percent.**

For most of the loading history, the strip was loaded incrementally by prescribing $\delta$. In places where the stress undergoes sudden changes in direction, Riks' path following scheme was adopted instead for easier convergence. The results are shown in Figs. 7–9. Figure 7(*a*) shows the calculated $\sigma$–$\delta$ response. Its general features are similar to what is seen experimentally. It consists of an initial linear part which terminates into a sharp limit load at 75.5 ksi (512 MPa) at which the instability in the form of localized deformation is initiated. The value of the initiation stress is governed by the size of the imperfection used. The load experiences a sharp drop down to a local stress valley of 69.4 ksi (479 MPa). (A small load valley can also be seen in the experimental response in Fig. 1.) Tracing these sharp changes in the response required use of Riks' method. Beyond the first stress valley, the end deflection $\delta$ was prescribed. The response follows a relatively flat plateau with an average stress of 72.2 ksi (498 MPa). Some detailed features of the stress plateau will be discussed below. The response terminates into a stress valley of 67.8 ksi (468 MPa) and subsequently the stress increases monotonically.

A sequence of 21 deformed configurations of the strip is shown in Fig. 8. They correspond to the points marked on the $\sigma$-$\delta$ response in Fig. 7(*a*) with solid dots. The three numbered dots correspond to the configurations with the same numbers in Fig. 8. For clarity, we have chosen to depict regions with surface strain higher than 1.2 percent in white and regions at lower strain levels in black. This cutoff level of strain is a bit less than 50 percent of the Lüders strain. Other cutoff strains closer to the Lüders strain level were also considered but the main features of interest did not change. (Note that the shattering strain of Stresscoat in the experiments is not known but is somewhere within the Lüders strain range.) The instability initiates via two narrow fingers which emanate from the imperfection and cross the width of the specimen as

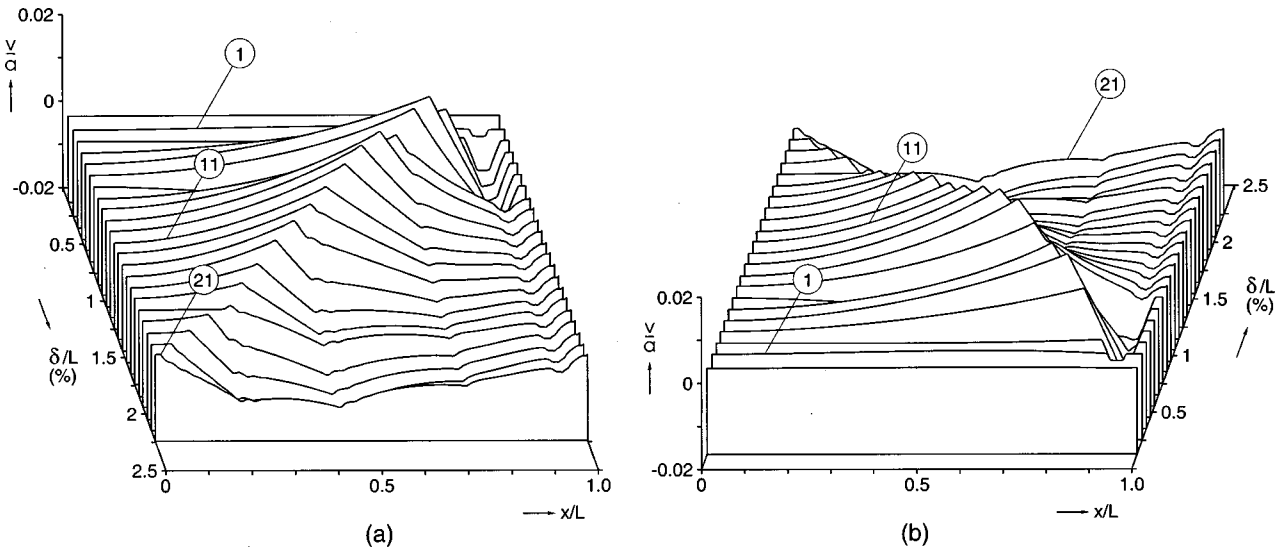



**Fig. 9 Lateral displacement of axis of strip as a function of $\delta$ for simulation in Figs. 7 and 8**

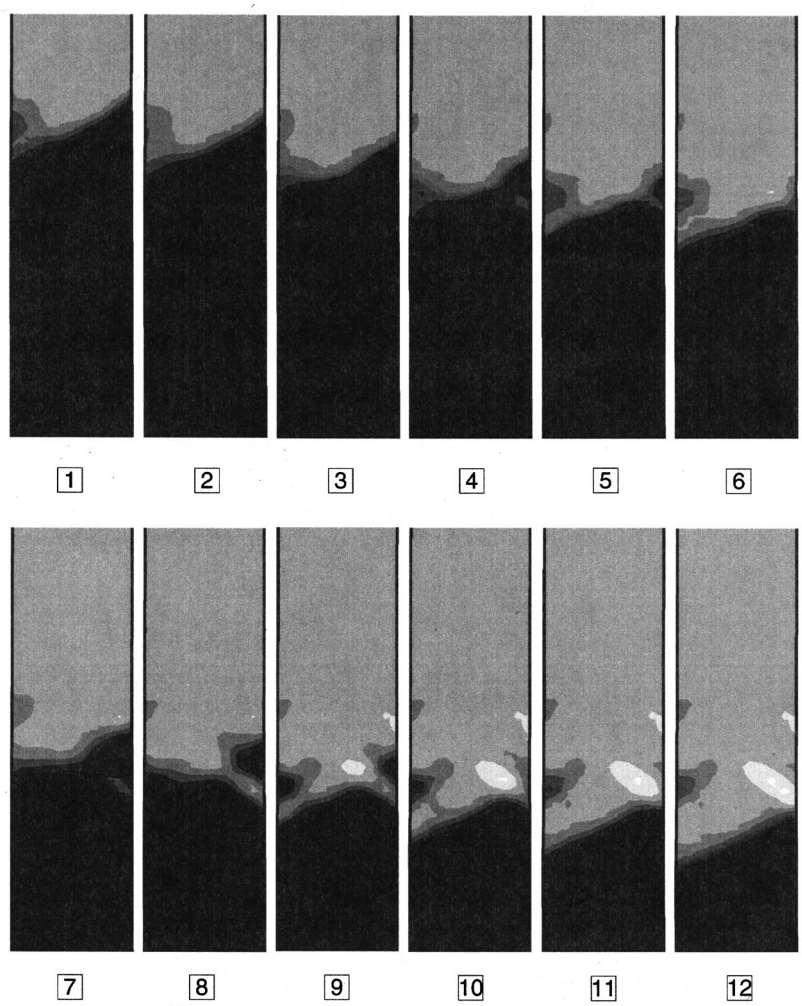



**Fig. 10 Expanded deformed configurations in gray scale showing one of the transient events**

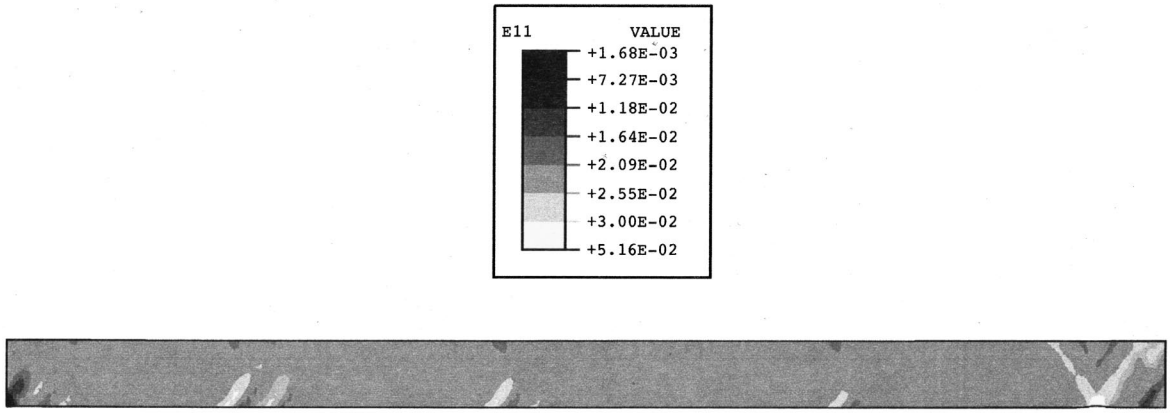



**Fig. 11 Gray scale plot of configuration (21) showing islands of uneven deformation left by five transient events**

shown in configuration (1) in Fig. 8. The two fingers are along the two characteristics of Hill [12] (see also [13]) oriented at approximately $\pm 55$ deg to the axis of loading. The local deformation quickly reaches the Lüders strain. When this is achieved, further deformation in these zones requires additional stress. Thus, spreading of the instability along the length of the strip becomes energetically preferred.

The localized deformation propagates through a front aligned approximately to one of the characteristics. The width of the front is of the order of one to two elements. As the material inside the front deforms to the Lüders strain, it provides the stress riser necessary to initiate deformation in the intact material adjacent to it. The stress riser required is the difference between the first stress peak of the local $\sigma$–$\varepsilon$ response and the current stress. The elements in the path of the front deform one, or a few, at a time to the Lüders strain. This produces small stress fluctuations just discernable in Fig. 7(*a*). The stress fluctuations double in size when the Lüders front propagates in the left half of the strip with the coarser mesh.

The steel used in the experiments had grains of the order of 10 $\mu$m in diameter. Thus, there were approximately 450 grains across the width and 75 through the thickness of the strip. The actual Lüders front is a few grains wide ([11]). In the model, the element size represents the main characteristic microstructural length. Since we have a maximum of 20 elements across the strip width, the calculated front width is larger than the actual one. Despite this difference, the main features of the evolution of deformation are very similar to those seen in the experiments. This is because many of the features are due primarily to structural effects and less dependent on the microstructure.

The two bands are seen to initially broaden simultaneously and, in the process, two islands of relatively undeformed material are formed on the right of the strip (see (2)). The induced deformation has a significant shear component which kinks the strip in the $x-y$ plane. Kinking was also observed in the experiments and has also been reported in the past (Lomer, [8]; Hall, [1]; Shaw & Kyriakides, [5]). As in Shaw and Kyriakides [6], we quantify the extent of kinking by plotting in Fig. 9 the transverse displacement of the axis of the strip ($v(x)$, $0 \leq x \leq L$) for each of the 21 configurations in Fig. 8 (due to the relief of the plot, views from two directions are provided). Because of the kink, the islands of relatively undeformed material, and the boundary conditions at the ends, a moment builds up in the strip. The moment at the end of the strip, $M(L)$, normalized by the yield moment ($M_o$) is plotted as a function of the applied end-displacement in Fig. 7(*b*). In configurations (1) to (4), the left front propagates down the strip while the islands of undeformed material on the right end remain relatively unchanged. In configuration (4), the kink angle ($\gamma$ in Fig. 4) has grown to 1 deg. A positive moment builds up which is relieved when the islands start to deform just after configuration (4). As this is taking place, the front on the left reverts to the more symmetric v-shape reported in the experiments of Shaw and Kyriakides [5] which temporarily removes the kink. Some fluctuations in stress are recorded while these transient events take place. In addition, some regions of slightly more deformed material are left behind which influence the moment in the strip. Between configurations (7) and (11), a single inclined front propagates towards the left. The kink has reappeared and positive moment builds up once more. Just after configuration (11), the moment can no longer be sustained and one more transient event takes place.

This event is shown in some detail in Fig. 10 where 12 expanded views of the front are plotted in gray scale. They correspond to events which occurred between configurations (11) and (13) in Figs. 7–9. In configurations [1] to [3], the front is inclined and the moment is rising. Configurations [4] to [9] show the transient switch to the v-shaped front which occurred with a drop in stress of 1.34 ksi (9.26 MPa) and a drop in moment of 0.048 $M_o$.

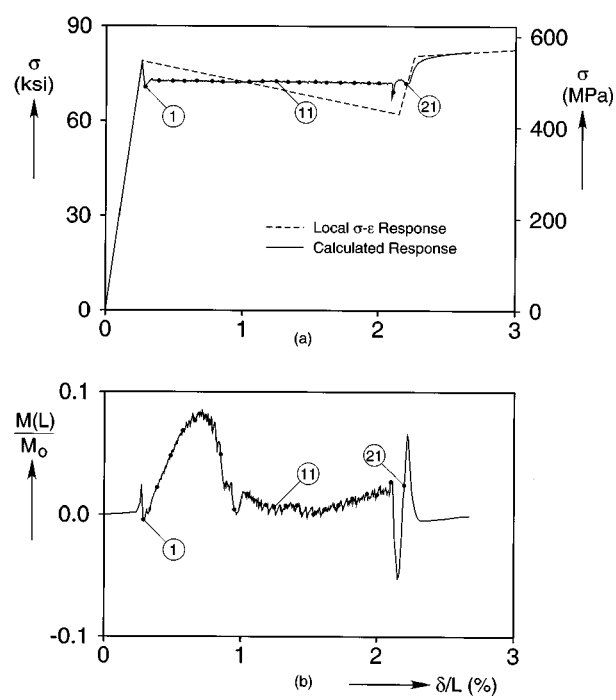



**Fig. 12 Simulation of propagation of two Lüders fronts; (*a*) stress-displacement response and (*b*) moment-displacement response**

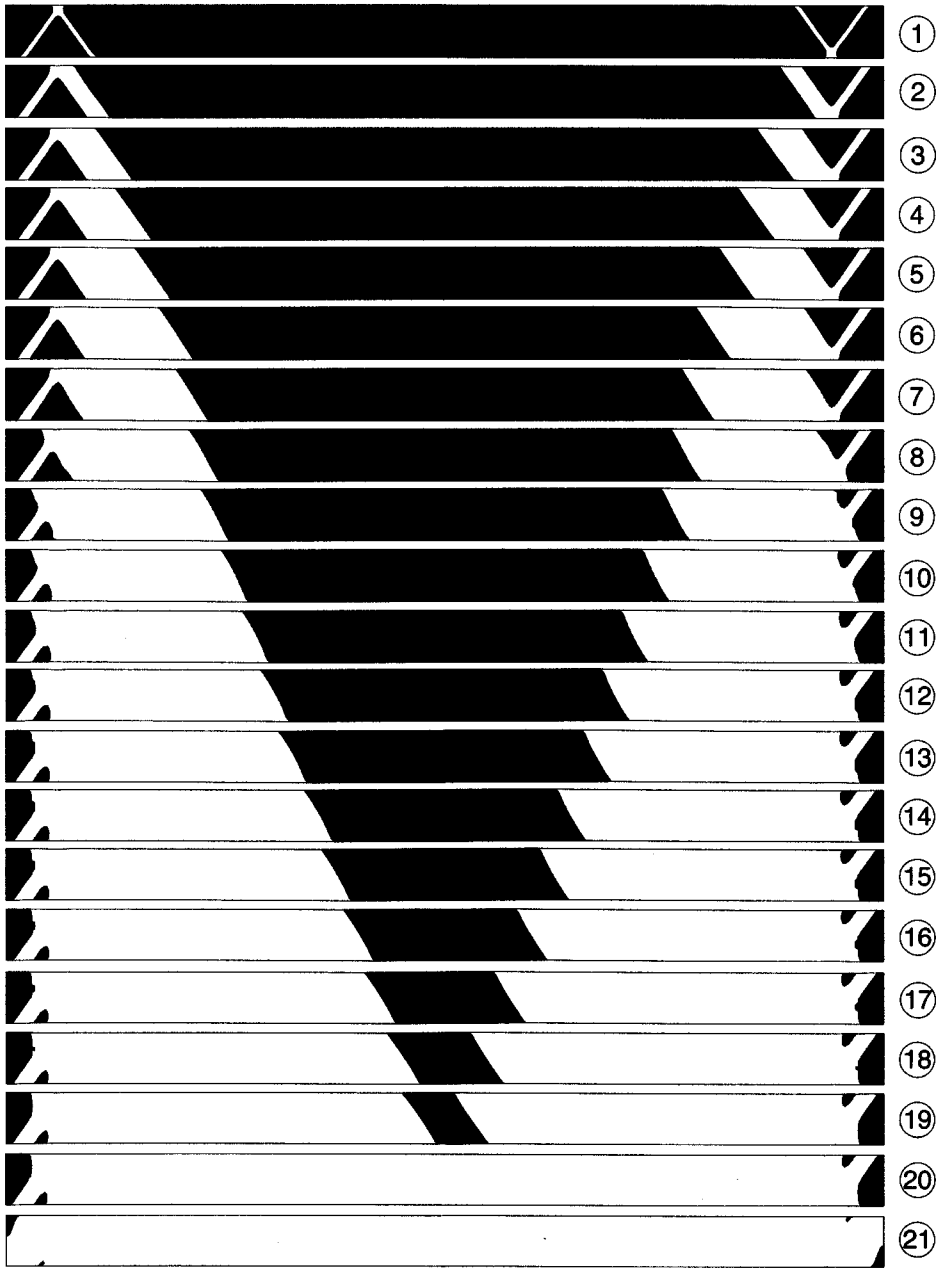



**Fig. 13 Sequence of black-white deformed configurations corresponding to responses in Fig. 12. White depicts zones with axial strain higher than 1.2 percent and black depicts zones lower than 1.2 percent.**

The finger-like deformation features associated with the v-front leave behind small islands of relatively undeformed material. By configuration 12, the transient is over and the front returns to the single inclination once more. Left behind are one small island of not totally deformed material and one island of somewhat more deformed material. The two islands are sources of stress nonuniformity which tends to reduce the moment. The transient has a permanent straightening effect on the strip seen in the plots of the axis transverse deflections in Fig. 9 for all configurations following 11.

The band propagates essentially in a steady-state manner until configuration 17 when another transient event takes place (see Fig. 8). A final transient event takes place when the left end of the strip starts to deform. The stress valley at the end of the stress plateau in Fig. 7(*a*) (67.8 ksi—468 MPa) corresponds to the deformation of the last bit of material. Riks' path following scheme had to be applied again in order to capture this stress valley. Such stress valleys at the termination of the unstable regime are usually present in experiments (see Figs. 1 and 2(*a*)). Under continued stretching, the strip deforms uniformly and the stress grows monotonically once more.

It is instructive to see on gray scale the deformation of the strip at the end of the stress plateau (configuration 21). In Fig. 11, we see remnants of five transient events that took place during the propagation of the Lüders deformation. Each produced islands of slightly more deformed material as well as islands of slightly less deformed material. These islands tend to have a straightening effect on the strip and are sources of stress nonuniformity. Each of the events can also be traced on the transverse deflection maps in Fig. 9.

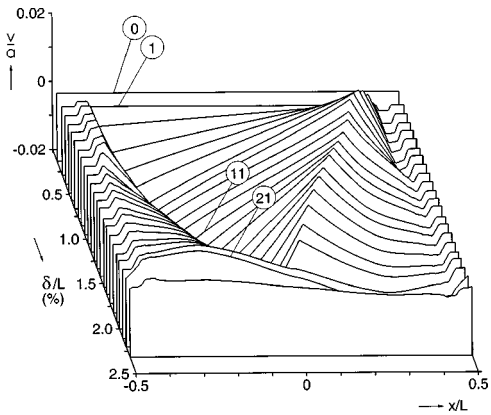



**Fig. 14 Lateral displacement of axis of strip as a function of $\delta$ for simulation in Figs. 12 and 13**

Some transient events like the ones described above have been observed in the experiments. They were difficult to capture with clarity due to the poor resolution provided by the visualization scheme used. Load fluctuations, which probably correspond to such transients, can also be seen in Fig. 2(*a*). However, as stated earlier, the 79 ksi (545 MPa) initiation stress adopted was chosen somewhat arbitrarily and is probably too low. We expect an increase of this variable to reduce the number of transients events seen in the simulation.

**(b) Propagation of Two Lüders Fronts.** Next we consider the case of two coexisting Lüders fronts emanating from the ends and propagating towards the middle of the specimen. In this case antisymmetry of deformation about the mid-length applies. Half of the length is then modeled as shown in Fig. 6(*b*) with the following boundary conditions at the plane of antisymmetry:

$$u(0,0,z)=v(0,0,z)=0, \quad u(0,y,z)=-u(0,-y,z), \quad v(0,y,z)=-v(0,-y,z) \quad w(0,y,z)=w(0,-y,z). \tag{3a}$$

The following boundary conditions are applied to the end of the strip:

$$u\left(\frac{L}{2},y,z\right)=\frac{\delta}{2}, \quad v\left(\frac{L}{2},0,z\right)=0. \tag{3b}$$

The domain was modeled with square elements with a mesh consisting of 20 elements across the width. For most of the loading history, $\delta$ is prescribed incrementally. The results are shown in Figs. 12–14. The main features of the stress-displacement response shown in Fig. 12(*a*) are similar to those in Fig. 7(*a*) but the deformation patterns are different. The instability is initiated with an imperfection with smaller depth $(0.033t)$ at a somewhat higher stress (78.5 ksi—541MPa). The initial stress drop is milder and can be traced by displacement control. Again two fingers of deformed material oriented along the two characteristics emanate from the imperfection (Fig. 13) and a stress valley develops down to 70.6 ksi (487 MPa, see Fig. 12(*a*). Initially, the two bands broaden together but by configuration ② deformation is limited to the inner fronts which propagate towards the middle of the strip. The strip kinks as shown in Fig. 14 and a moment starts to build up. In the neighborhood of configuration ⑥, the stress nonuniformity forces the islands of undeformed material to deform into a more symmetric pattern seen in ⑧ and in the process the moment is reduced. From configuration ⑨ onwards, the two middle fronts propagate towards each other eventually coalescing just before configuration ⑳. The coalescence causes the stress to drop down to 67.5 ksi (466 MPa) as the two fronts interact. Tracing of this stress valley required temporary use of Riks' method. The islands of undeformed material at the ends deform last and subsequently the strip returns to homogeneous deformation.

Overall, the two fronts keep the moment in the strip at a lower level and, as a result, events such as those seen in the previous case do not take place. Thus, a much flatter stress plateau is traced between the initiation of localized deformation and the coalescence of the two fronts in the middle of the strip.

## 4 Summary and Conclusions

The phenomenon of propagation of Lüders bands in mild steel is perhaps the oldest example of a propagating instability due to micromechanical causes (see Kyriakides [14] for related phenomena at the structural level). In thin strips, the Lüders deformation tends to initiate along preferred directions (characteristics) from small stress concentrations often present at the ends of the specimens. Lüders deformation spreads by the propagation of inclined fronts which separate coexisting regions of relatively undeformed material and material deformed to the Lüders strain. For fine grained steels, the fronts tend to be well defined and sharp and, under relatively slow displacement controlled stretching, they propagate in a steady-state manner. In the present experimental study, the surface of the strips was coated with a brittle coating. The coating shatters at a level within the Lüders strain and this enables monitoring of the evolution of the fronts along the strip. In most experiments, from the present study and from the literature, either one or two propagating Lüders fronts developed. The velocity of propagation of the fronts is captured well by the kinematical condition given in Eq. (1).

The Lüders phenomena in thin steel strips were modeled with suitably nonlinear finite elements with a relatively fine and isotropic mesh. The material was modeled as a finitely deforming $J_2$-type elastic-plastic solid. A key aspect of the constitutive model is the adoption of a trilinear up-down-up stress-strain response termed the local stress-strain response. It was chosen to have a Maxwell stress which corresponds to the plateau stress measured experimentally and extending to the experimentally measured strain level. The upper yield stress of the steel is rather difficult to accurately determine experimentally as it is masked by ever present stress concentrations. Furthermore, it is rather sensitive to the rate of loading. Thus, the choice of the difference between the initial stress peak and the stress valley of the trilinear stress-strain response was somewhat arbitrary.

The model strips had the same dimensions as the strips tested but they had a uniform cross section. In addition, the boundary conditions used at the ends did not produce stress concentrations. Lüders deformation was initiated from small initial geometric imperfections at the ends of the strips. In the first case, only one initiation site was provided and in the second two at opposite ends of the strip. Two finger-like deformation features emanated from the imperfections inclined at $\pm 55$ deg to the axis of loading. The net stress dropped sharply and a local stress valley formed as in the experiments. The propagation of the single front was found to be sustainable. The shear deformation associated with the inclined front tends to kink the strip and a moment builds up. For a single propagating front, the asymmetry was more pronounced and the moments developed were of larger amplitude. Because of the relatively low difference between the assumed upper yield stress and the plateau stress, the moment caused periodic transient events during which the front tended to propagate with a more symmetric v-shape. This, and islands of nonuniformly deformed material left behind, tended to straighten the strip and relieve the moment. Several such transient events took place while, between them, the front propagated at a relatively constant inclination along one of the characteristics.

The propagation of two fronts antisymmetrically oriented was found to be less eventful. The initiation is similar but steady-steady propagation was more sustainable because the antisymmetry kept the kinked strip straighter and the moments at a much lower level. The stress plateau was flatter until the end when the two fronts coalesced in the middle of the strip producing a stress valley like ones seen in experiments.

The width of the propagating front depends on the mesh size and as a result we suspect it to be wider than the actual ones in our fine grained steel strips. This difference did not seem to affect most other events which were more affected by structural effects.

## Acknowledgments

The financial support of the National Science Foundation through grant CMS-95-03396 is acknowledged with thanks. Computer resources were provided through NSF cooperative agreement ACI-9619020 through the National Partnership for Advanced Computational Infrastructure at the San Diego Supercomputer Center. The numerical model used and particularly the reduction of the numerical results were influenced by the framework developed by J. A. Shaw in the related study of SMAs.

## References

[1] Hall, E. O., 1970, *Yield Point Phenomena in Metals and Alloys*, Plenum Press, New York.

[2] Cottrell, A. H., and Bilby, B. A., 1948, ''Dislocation Theory of Yielding and Strain Ageing of Iron,'' Proc. Physical Society, **62/I-A**, pp. 49–62.

[3] Johnston, W. G., and Gilman, J. J., 1959, ''Dislocation Velocities, Dislocation Densities, and Plastic Flow in Lithium Fluoride Crystals,'' J. Appl. Phys., **30**, pp. 129–144.

[4] Hahn, G. T., 1962, ''A Model for Yielding With Special Reference to the Yield-Point Phenomena of Iron and Related BCC Metals,'' Acta Metall., **10**, pp. 727–738.

[5] Shaw, J. A., and Kyriakides, S., 1997, ''On the Nucleation and Propagation of Phase Transformation Fronts in a NiTi Alloy,'' Acta Mater., **45**, pp. 683–700.

[6] Shaw, J. A., and Kyriakides, S., 1998, ''Initiation and Propagation of Localized Deformation in Elasto-Plastic Strips Under Uniaxial Tension,'' Int. J. Plast., **13**, pp. 837–871.

[7] Shaw, J. A., and Kyriakides, S., 1995, ''Thermomechanical Aspects of NiTi,'' J. Mech. Phys. Solids, **43**, pp. 1243–1281.

[8] Lomer, W. M., 1952, ''The Yield Phenomenon in Polycrystalline Mild Steel,'' J. Mech. Phys. Solids, **1**, pp. 64–73.

[9] Butler, J. F., 1962, ''Lüders Front Propagation in Low Carbon Steels,'' J. Mech. Phys. Solids, **10**, pp. 313–334.

[10] Abeyaratne, R., and Knowles, J. K., 1993, ''A Continuum Model of a Thermoelastic Solid Capable of Undergoing Phase Transitions,'' J. Mech. Phys. Solids, **41**, pp. 541–571.

[11] Morrison, W. B., and Glenn, R. C., 1968, ''Examination of the Lüders Front in a Low-Carbon Steel by Transmission Electron Microscopy,'' J. Iron Steel Inst., **206**, pp. 611–612.

[12] Hill, R., 1952, ''On Discontinuous Plastic States, With Special Reference to Localized Necking in Thin Sheets,'' J. Mech. Phys. Solids, **1**, pp. 19–30.

[13] Bijlaard, P. P., 1940, ''Theory of Local Plastic Deformations,'' Publication of the International Association of Bridge and Structural Engineers, **6**, pp. 27–44.

[14] Kyriakides, S., 1993, ''Propagating Instabilities in Structures,'' *Advances in Applied Mechanics*, Vol. 30, J. W. Hutchinson and T. Y. Wu, eds., Academic Press, San Diego, pp. 67–189.

# Common Errors on Mapping of Nonelliptic Curves in Anisotropic Elasticity

**T. C. T. Ting**
Fellow ASME
Professor of Applied Mechanics,
Department of Civil and Materials Engineering,
University of Illinois at Chicago,
842 W. Taylor Street, M/C 246,
Chicago, IL 60607-7023

*For an isotropic elastic material under a two-dimensional plane-stress or plane-strain deformation that involves a closed boundary of a given geometry Γ, it is a common practice to use a conformal mapping to map Γ to a circle of unit radius. For an anisotropic elastic material, the mapping produces the correct solution to the physical problem when Γ is an ellipse. There are published (and unpublished) papers that report solutions to anisotropic elasticity problems with a nonelliptic curve Γ. They all fail to solve the physical problem correctly. The purpose of this paper is to show the common errors made by many researchers, and to alert newcomers on the subject so that they do not fall into the same trap.* [S0021-8936(00)02403-X]

## 1 Introduction

Consider an infinite isotropic elastic material under a two-dimensional plane-stress or plane-strain deformation. The material has a hole of given geometry Γ. The boundary conditions are prescribed on Γ and at infinity. The solution for the displacement and stress can be expressed in terms of the complex variable $z$ and its complex conjugate $\bar{z}$ where

$$z = x_1 + i x_2. \tag{1}$$

It is a common practice to transform the curve Γ to a circle of unit radius by a conformal mapping

$$z = w(\zeta). \tag{2}$$

The function $w(\zeta)$ is chosen such that, when $z$ is on Γ, $\zeta$ is on the unit circle. It is important that the mapping be reversible and single-valued. Every point outside Γ in the $z$-plane is mapped to only one point in the region outside the unit circle in the $\zeta$-plane, and vice versa. In other words, the mapping must be ''one-to-one.'' This condition is satisfied if ([1])

$$\frac{d}{d\zeta} w(\zeta) \neq 0 \quad \text{for } |\zeta| > 1. \tag{3}$$

For an anisotropic elastic material under a two-dimensional deformation the solution can be expressed in terms of three complex variables ([2,3])

$$z_\alpha = x_1 + p_\alpha x_2 \quad (\alpha = 1,2,3) \tag{4}$$

and their complex conjugates $\bar{z}_\alpha$. The complex constants $p_\alpha$ ($\alpha = 1,2,3$) depend on elastic constants only. The imaginary part of $p_\alpha$ is positive and nonzero. $p_\alpha = i$ ($\alpha = 1,2,3$) for isotropic materials. To map the curve Γ to a unit circle we need three mapping functions

$$z_\alpha = w_\alpha(\zeta_\alpha) \quad (\alpha = 1,2,3). \tag{5}$$

Two important conditions must be met before one starts to solve the boundary value problem.

(i) Each point on Γ in the $(x_1, x_2)$-plane is mapped to the *same* point on the unit circle in the $\zeta_\alpha$-plane for all three mapping functions $w_\alpha(\zeta_\alpha)$.

(ii) The mapping is one-to-one between the region outside Γ and the region outside the unit circle for each $w_\alpha(\zeta_\alpha)$. This is assured if

$$\frac{d}{d\zeta_\alpha} w_\alpha(\zeta_\alpha) \neq 0 \quad \text{for } |\zeta_\alpha| > 1 \quad (\alpha = 1,2,3). \tag{6}$$

All published (and unpublished) papers on nonelliptic Γ that this author has read fail to meet the condition (i) or (ii). When the condition (i) is violated, the boundary condition on Γ cannot be satisfied for all points on Γ because each point is mapped to three different points on the unit circle in the $\zeta_\alpha$-plane for $\alpha = 1,2,3$. If the condition (ii) is violated, a Riemann surface with a branch cut must be introduced in the $(x_1, x_2)$-plane so that the mapping is one to one. The solution so obtained does not represent the original physical problem because the region outside Γ has a branch cut, and is no longer a continuous medium.

It is unfortunate that a few incorrect papers on nonelliptic Γ that have appeared in print have induced countless number of unpublished papers to fall into the same trap. What is particularly sad for the authors of these unpublished papers on nonelliptic Γ is that they have spent so much of their time in solving the boundary value problems, not knowing that they have made a fundamental error on mapping at the very beginning of their papers. It is hoped that this paper would prevent newcomers on the subject from falling into the same trap.

We will elaborate on conditions (i) and (ii) separately below.

## 2 The Mapping for Anisotropic Elasticity

Equation (2) for isotropic elastic material can be approximated by a polynomial of the form ([1])

$$z = w(z) = c_0 \zeta + \frac{c_1}{\zeta} + \frac{c_2}{\zeta^2} + \ldots + \frac{c_k}{\zeta^k} \tag{7}$$

where $c_0, c_1, c_2, \ldots$ are constants and $k$ is a positive integer. We have deleted a constant term in (7) that does not affect the general results. The following are some of the special cases ([1,4]).

$$\text{Ellipse:} \quad w(\zeta) = \frac{a+b}{2}\zeta + \frac{a-b}{2\zeta}, \tag{8}$$

$$\text{Triangle:} \quad w(\zeta) = a\left\{\zeta + \frac{1}{3\zeta^2} + \frac{1}{45\zeta^5} - \ldots\right\}, \tag{9}$$

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Sept. 10, 1999; final revision, Dec. 15, 1999. Associate Technical Editor: W. J. Drugan. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

 **Copyright © 2000 by ASME** 

$$\text{Square:}\quad w(\zeta)=a\left\{\zeta-\frac{1}{6\zeta^3}+\frac{1}{56\zeta^7}-\ \dots\right\},\tag{10}$$

$$\text{Pentagon:}\quad w(\zeta)=a\left\{\zeta+\frac{1}{10\zeta^4}+\frac{1}{75\zeta^9}-\ \dots\right\},\tag{11}$$

where $a\geqslant b$ are constants.

For an anisotropic elastic material we may first map the $(x_1,x_2)$-plane to the $(x_1^\alpha,x_2^\alpha)$-plane by

$$z^\alpha=x_1^\alpha+ix_2^\alpha=x_1+p_\alpha x_2=x_1+p'_\alpha x_2+ip''_\alpha x_2\quad(\alpha=1,2,3),\tag{12}$$

where $p'_\alpha$ and $p''_\alpha>0$ are the real and imaginary parts of $p_\alpha$. This transformation can be seen as stretching the $(x_1,x_2)$-plane uniformly in the $x_2$-direction with the stretch ratio $p''_\alpha$ and, if $p'_\alpha\neq 0$, shearing the $(x_1,x_2)$-plane with the $x_1$-axis being held fixed ([5], p. 73). The curve $\Gamma$ maps to a curve $\Gamma^\alpha$ in the $z^\alpha$-plane. If $\Gamma$ is a square $\Gamma^\alpha$ is a parallelogram. If $\Gamma$ is a circle $\Gamma^\alpha$ is an ellipse. If $\Gamma$ is an ellipse $\Gamma^\alpha$ is an ellipse of different shape. The mapping in this step is one to one, but *not* conformal (see Section 4). The second step is to map the curves $\Gamma^\alpha$ to a unit circle in the $\zeta_\alpha$-plane by a conformal mapping. It should be noted that the curves $\Gamma^\alpha$ for $\alpha=1,2,3$ are different curves unless $p_1=p_2=p_3$. Therefore this two-step mapping maps each point on $\Gamma$ to *three* different points on the unit circle in the $\zeta_\alpha$-plane for $\alpha=1,2,3$. This violates condition (i).

To map each point on $\Gamma$ to the *same* point on the unit circle in the $\zeta_\alpha$-plane for $\alpha=1,2,3$, many papers make use of the mapping (2) that is valid for isotropic elastic materials. By writing (2) and its complex conjugate as

$$x_1+ix_2=w(\zeta),\quad x_1-ix_2=\overline{w}(\overline{\zeta}),\tag{13}$$

they can be solved for $x_1$ and $x_2$, and $z_\alpha$ can be written as

$$z_\alpha=x_1+p_\alpha x_2=\frac{1-ip_\alpha}{2}w(\zeta)+\frac{1+ip_\alpha}{2}\overline{w}(\overline{\zeta}).\tag{14}$$

Since $\overline{\zeta}=\zeta^{-1}$ on the unit circle, (14) is replaced by

$$z_\alpha=x_1+p_\alpha x_2=\frac{1-ip_\alpha}{2}w(\zeta)+\frac{1+ip_\alpha}{2}\overline{w}(\zeta^{-1}).\tag{15}$$

Equation (15) maps each point on $\Gamma$ to the same point on the unit circle for $\alpha=1,2,3$. Therefore it satisfies condition (i).

Unfortunately, as we will show in the next section, the mapping (15) is no longer one to one for the region outside $\Gamma$ unless $\Gamma$ is an ellipse. The error was made when $\overline{\zeta}$ in (14) was replaced by $\zeta^{-1}$ to obtain (15). This is the crucial error made by most papers. It is also an error easily missed by most reviewers.

## 3 The Mapping (15) is *Not* One to One for Nonelliptic $\Gamma$

If we take the first two terms in (8)–(11), substitute them into (15) and write $\zeta$ as $\zeta_\alpha$, we have

$$\text{Ellipse:}\quad z_\alpha=w_\alpha(\zeta_\alpha)=\frac{a-ip_\alpha b}{2}\zeta_\alpha+\frac{a+ip_\alpha b}{2\zeta_\alpha},\tag{16}$$

$$\text{Triangle:}\quad z_\alpha=w_\alpha(\zeta_\alpha)=\frac{a}{2}\left\{(1-ip_\alpha)\left(\zeta_\alpha+\frac{1}{3\zeta_\alpha^2}\right)+(1+ip_\alpha)\left(\frac{1}{\zeta_\alpha}+\frac{\zeta_\alpha^2}{3}\right)\right\},\tag{17}$$

$$\text{Square:}\quad z_\alpha=w_\alpha(\zeta_\alpha)=\frac{a}{2}\left\{(1-ip_\alpha)\left(\zeta_\alpha-\frac{1}{6\zeta_\alpha^3}\right)+(1+ip_\alpha)\left(\frac{1}{\zeta_\alpha}-\frac{\zeta_\alpha^3}{6}\right)\right\},\tag{18}$$

$$\text{Pentagon:}\quad z_\alpha=w_\alpha(\zeta_\alpha)=\frac{a}{2}\left\{(1-ip_\alpha)\left(\zeta_\alpha+\frac{1}{10\zeta_\alpha^4}\right)+(1+ip_\alpha)\left(\frac{1}{\zeta_\alpha}+\frac{\zeta_\alpha^4}{10}\right)\right\}.\tag{19}$$

Equation (16) is the mapping of an ellipse for anisotropic elastic materials first employed by Lekhnitskii [2,6]. It is the only one that is in the form of (7). It can be shown that (16) satisfies (6) ([5], p. 85). Hence the mapping (16) is one to one. Equation (17) has the $\zeta_\alpha^2$ term while (18) and (19) have the $\zeta_\alpha^3$ and $\zeta_\alpha^4$ terms. These terms do not appear in (7). Of course, these terms disappear when $p_\alpha=i\ (\alpha=1,2,3)$, which is the case when the material is isotropic. There are anisotropic elastic materials for which $p_\alpha=i\ (\alpha=1,2,3)$. We will discuss these materials in Section 5. It is easy to see that the mapping (17) is not one to one. When $|\zeta_\alpha|$ is very large (17) can be approximated by

$$z_\alpha=w_\alpha(\zeta_\alpha)=a(1+ip_\alpha)\frac{\zeta_\alpha^2}{3}.\tag{20}$$

The mapping is not one to one because, when we go around the origin of the $\zeta_\alpha$-plane one full circle, we will be going around the origin of the $(x_1,x_2)$-plane two full circles. For the mapping to be one to one we have to introduce a branch cut in the $(x_1,x_2)$-plane.

As an illustration let $p_\alpha=\frac{3}{5}i$ and $a=1$ so that (17) is

$$x_1+i\frac{3}{5}x_2=w_\alpha(\zeta_\alpha)=\frac{1}{5}\left\{4\left(\zeta_\alpha+\frac{1}{3\zeta_\alpha^2}\right)+\left(\frac{1}{\zeta_\alpha}+\frac{\zeta_\alpha^2}{3}\right)\right\}.\tag{21}$$

Figure 1 shows the mapping between the $(x_1,x_2)$-plane and the $\zeta_\alpha$-plane. A branch cut $\kappa$ in the $(x_1,x_2)$-plane is introduced so that the mapping is one to one. The branch cut $\kappa$ is mapped to the curve $\kappa_\zeta$ in the $\zeta_\alpha$-plane. The *branch points* for $\kappa_\zeta$ and $\kappa$ are computed from (6) (with the $\neq$ replaced by $=$) and (21) as

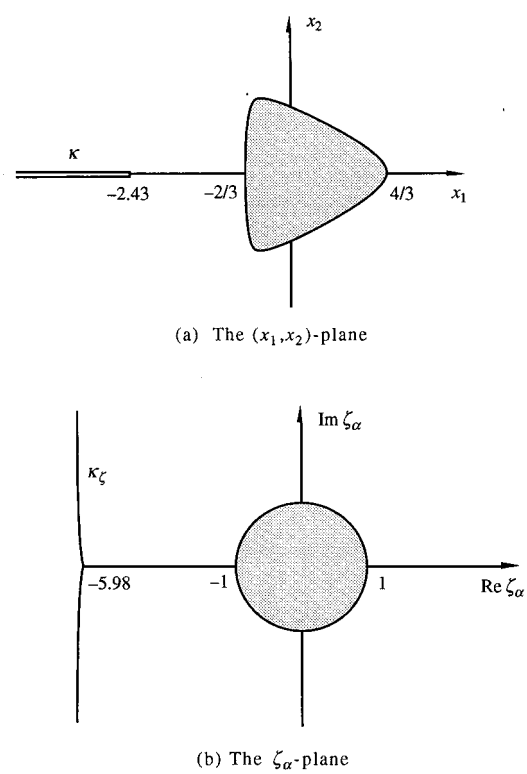



**Fig. 1 A triangle in the $(x_1,x_2)$-plane is mapped to a circle in the $\zeta_\alpha$-plane through Eq. (21)**

$$\zeta_\alpha = -5.98, \quad z_\alpha = -2.43. \tag{22}$$

The curve $\kappa_\zeta$ is practically a straight vertical line with $\zeta_\alpha = -6$ as its asymptotic line. The *infinite* region in the $(x_1, x_2)$-plane outside the triangle is mapped to the *semi-infinite* region in the $\zeta_\alpha$-plane outside the unit circle. The region to the left of $\kappa_\zeta$ is excluded. This mapping creates two problems. First, due to the presence of the branch cut $\kappa$, the region outside the triangle in the $(x_1, x_2)$-plane is no longer a continuous medium. Secondly, the boundary conditions prescribed at infinity in the $(x_1, x_2)$-plane do not correspond to the boundary conditions at the *whole* infinity in the $\zeta_\alpha$-plane because the region in the $\zeta_\alpha$-plane is a semi-infinite region. Any one of the two problems invalidates the analysis.

Likewise, one can show that the mappings (18) and (19) are not one to one. They create the same problems as the mapping (17) does.

## 4 The Mapping for Anisotropic Elastic Material is *Not Conformal*

For an isotropic elastic material the mappings (8)–(11) are *conformal* because the angle between any two infinitesimal lines that intersect at a point in the $(x_1, x_2)$-plane and the angle between the two corresponding lines in the $\zeta$-plane are identical. Moreover, the *sense* of the angle is preserved. A mapping that preserves the magnitudes of the angles but not necessarily the sense is called *isogonal* ([7]).

For an anisotropic elastic material for which $p_\alpha \neq i$ the angle between any two infinitesimal lines that intersect at a point in the $(x_1, x_2)$-plane and the angle between the two corresponding lines in the $\zeta_\alpha$-plane are in general different. Therefore the mapping is *not* conformal, *nor* isogonal. This is irrespective of whether the mapping function satisfies condition (i) or (ii). The mappings (16)–(19) are all nonconformal, even though (16) satisfies conditions (i) and (ii). Most papers that dealt with mapping for anisotropic elastic materials used the word conformal mapping indiscriminately. A nonconformal mapping poses no problems in anisotropic elasticity. It is completely irrelevant in finding a correct solution. Nevertheless, the term conformal mapping does not apply to anisotropic elasticity.

## 5 Concluding Remarks

We have shown that, for an anisotropic elastic material with a hole of given geometry $\Gamma$ under a two-dimensional deformation, a mapping of $\Gamma$ to a unit circle does not provide a correct solution to the physical problem unless $\Gamma$ is an ellipse. Lekhnitskii [2,6] was the first one to solve the anisotropic elasticity problems with an elliptic hole. He did consider nonelliptic holes. However, he assumed that the hole was a *small* perturbation of an ellipse, or the material was a *small* perturbation from isotropic elastic material. That was as far as he could go. Had he known how to solve the problem for a nonelliptic hole, he would have done it a long time ago.

This does *not* mean that the problem cannot be solved for *any* anisotropic elastic material when it has a nonelliptic hole. There are *degenerate* anisotropic elastic materials for which $p_\alpha = i$ for $\alpha = 1,2,3$ ([8]). There are also degenerate anisotropic elastic materials for which $p_\alpha \neq i$, but $p_1 = p_2 = p_3$ ([9–11]). These materials are degenerate in the mathematical sense, not necessarily in the physical sense. They need not be a small perturbation from an isotropic elastic material. They can be *triclinic*, i.e., they may not possess a plane of material symmetry. For these degenerate materials conditions (i) and (ii) can be easily satisfied. Therefore, a correct solution to the physical problem can be obtained for a nonelliptic hole for these materials. Recently, Ting [12] has shown that there are anisotropic elastic materials that uncouple antiplane and in-plane displacements but not antiplane and in-plane stresses, and vice versa. Monoclinic material with the symmetry plane at $x_3 = 0$ is a special case that uncouples both displacements and stresses. For these materials, if the solution depends on $p_3$ alone or if the solution depends on $p_1$ and $p_2$ with $p_1 = p_2$, a correct solution for a nonelliptic hole can be obtained.

There are papers on nonelliptic $\Gamma$ that provide an approximate solution with a small parameter $\varepsilon$. The exact solution to the problem is obtained when $\varepsilon \to 0$. This is certainly acceptable except that, for these papers, the geometry of $\Gamma$ reduces to an ellipse when $\varepsilon \to 0$. Thus the problem is equivalent to that of a small perturbation from an ellipse considered by Lekhnitskii [2,6]. An approximate solution for a nonelliptic $\Gamma$ is justified if, in the limit when the approximate solution becomes the exact solution, the geometry of $\Gamma$ remains nonelliptic. The question is open if an exact solution or an approximate solution in the sense mentioned above (other than a complete numerical solution) for a nonelliptic $\Gamma$ can be obtained for a general anisotropic elastic material. It is a challenging problem for the new millennium.

## References


[1] Muskhelishvili, N. I., 1953, *Some Basic Problems of the Mathematical Theory of Elasticity* (translation by J. R. M. Radok, Noordhoff, Groningen).

[2] Lekhnitskii, S. G., 1950, *Theory of Elasticity of an Anisotropic Body*, Gostekhizdat, Moscow (in Russian), *Theory of Elasticity of an Anisotropic Elastic Body*, Holden-Day, San Francisco (in English, 1963), and Mir Pub. Moscow (in English, 1981).

[3] Stroh, A. N., 1958, ''Dislocations and Cracks in Anisotropic Elasticity,'' Philos. Mag., **3**, pp. 625–646.

[4] Savin, G. N., 1961, *Stress Concentration Around Holes*, Pergamon Press, London.

[5] Ting, T. C. T., 1996, *Anisotropic Elasticity: Theory and Applications*, Oxford University Press, New York.

[6] Lekhnitskii, S. G., 1957, *Anisotropic Plates*, 2nd Ed., Gostekhizdat, Moscow (in Russian), (translation by S. W. Tsai and T. Cheron, Gordon and Breach, New York (1968, 1984, 1987).

[7] Churchill, R. V., 1948, *Introduction to Complex Variables and Applications*, McGraw-Hill, New York.

[8] Ting, T. C. T., 1994, ''On Anisotropic Elastic Materials That Possess Three Identical Stroh Eigenvalues as do Isotropic Materials,'' Q. Appl. Math., **52**, pp. 363–375.

[9] Ting, T. C. T., 1996, ''Existence of an Extraordinary Degenerate Matrix **N** for Anisotropic Elastic Materials,'' Q. J. Mech. Appl. Math., **49**, pp. 405–417.

[10] Yin, W.-L., 2000, ''Deconstructing Plane Anisotropic Elasticity, Part I: The Latent Structure of Lekhnitskii's Formalism,'' Int. J. Solids Struct., **37**, pp. 5257–5276.

[11] Yin, W.-L., 2000, ''Deconstructing Plane Anisotropic Elasticity, Part II: Stroh's Formalism Sans Frills,'' Int. J. Solids Struct., **37**, pp. 5277–5296.

[12] Ting, T. C. T., 2000, ''Anisotropic Elastic Materials That Uncouple Antiplane and Inplane Displacements but Not Antiplane and Inplane Stresses, and Vice Versa,'' Math. Mech. Solids, **5**, pp. 139–156.

# On the Interior Stress Problem for Elastic Bodies

**J. Helsing**
Department of Solid Mechanics and NADA,
Royal Institute of Technology,
SE-100 44 Stockholm, Sweden

*The classic Sherman-Lauricella integral equation and an integral equation due to Muskhelishvili for the interior stress problem are modified. The modified formulations differ from the classic ones in several respects: Both modifications are based on uniqueness conditions with clear physical interpretations and, more importantly, they do not require the arbitrary placement of a point inside the computational domain. Furthermore, in the modified Muskhelishvili equation the unknown quantity, which is solved for, is simply related to the stress. In Muskhelishvili's original formulation the unknown quantity is related to the displacement. Numerical examples demonstrate the greater stability of the modified schemes.* [S0021-8936(00)01304-0]

## 1 Introduction

The task of computing the elastic field inside an unconstrained body subjected to external stress is a basic one in applied mechanics. A variety of numerical methods exist, leading to the solution of systems of linear equations. Finite element methods and integral equation methods are two examples. A problem, for any method, is a certain undeterminacy in the solution—when stress is applied, displacement is not unique.

There are standard ways to get a well-posed problem. In a finite element program one can prescribe also the displacement at some points (to prevent rigid-body movements). In the context of integral equations, the integral operator can be completed with an extra operator, containing another arbitrary point, which makes the solution unique. The choice of particular representations of the unknown fields and placements of arbitrary points will, of course, affect the stability of a numerical code. With a direct solver and for simple problems this may not be an issue. The computational work only depends on the size of the system matrix. With the faster iterative solvers used by many engineers today and in difficult situations, the stability of the code and the condition number of the system matrix is suddenly important. The lower the condition number, the faster the solver will converge. A stable algorithm can give a solution with better quality.

This paper focuses on the Sherman-Lauricella integral equation and an integral equation due to Muskhelishvili for the interior stress problem in two-dimensional elastostatics. The classic way to get a unique solution for these equations is to complete them with an operator $B$ containing an arbitrary point $z^*$ ([1]). We show that the convergence properties of iterative algorithms based on these equations can be sensitive to the placement of $z^*$. To remedy this situation, we introduce a uniqueness condition with a clear physical interpretation. This leads us to a new operator $B$ which is free from the arbitrary point $z^*$. We then derive a modification of the Muskhelishvili equation which is better suited to compute stress fields. Numerical experiments indicate that our modified equations give more efficient algorithms.

## 2 Potential Representation

A finite, linearly elastic body occupies a domain $D$. Its two-dimensional elastic bulk and shear moduli are $\kappa$ and $\mu$. The boundary of the body is denoted $\Gamma$and is given positive (counter-clockwise) orientation. Traction is prescribed at $\Gamma$. We would like to compute the deformation and the stress field inside $D$.

Let $U$ denote the Airy stress function. Since $U$ satisfies the biharmonic equation inside $D$ it can be represented as

$$U=\Re\mathrm{e}\{\bar{z}\phi+\chi\}, \tag{1}$$

where the potentials $\phi$ and $\chi$ are single-valued analytic functions of the complex variable $z=x+iy$. For a thorough discussion of the complex variable approach to elasticity problems, see ([2–5]). For our purposes it is sufficient to observe a few relations that link the complex potentials to quantities of physical interest: The displacement $(u_x,u_y)$ in the material satisfies

$$u_x+iu_y=\left(\frac{1}{2\mu}+\frac{1}{\kappa}\right)\phi-\frac{1}{2\mu}(z\bar{\phi}'+\bar{\psi}), \tag{2}$$

where $\psi=\chi'$. The integral of traction $(t_x,t_y)$ along a curve $\gamma(s)$ can be obtained from the relation

$$\int_{s_0}^{s}(t_x+it_y)ds=-\Big|_{s_0}^{s} i(\phi+z\bar{\phi}'+\bar{\psi}), \tag{3}$$

where $s$ denotes arclength along $\gamma(s)$. Complex differentiation of the expression (2) along the tangent to $\Gamma(s)$ gives

$$\frac{d}{dz}(u_x+iu_y)=\left(\frac{1}{2\mu}+\frac{1}{\kappa}\right)\Phi-\frac{1}{2\mu}\left(\bar{\Phi}-\frac{\bar{n}}{n}z\overline{\Phi'}-\frac{\bar{n}}{n}\bar{\Psi}\right), \tag{4}$$

and differentiation with respect to arclength in (3) gives

$$t_x+it_y=\Phi n+\bar{\Phi}n-z\overline{\Phi'}\bar{n}-\bar{\Psi}\bar{n}, \tag{5}$$

where $\Phi=\phi'$, $\Psi=\chi''$, and $n=n_x+in_y$ is the outward unit normal vector on $\Gamma$. The components of the stress tensor can be computed via

$$\sigma_{xx}+\sigma_{yy}=4\,\Re\mathrm{e}\{\Phi\}, \tag{6}$$

$$\sigma_{yy}-\sigma_{xx}+2i\sigma_{xy}=2(\bar{z}\Phi'+\Psi). \tag{7}$$

A natural starting point for elastostatic problems is to represent the potentials $\phi$ and $\psi$, or $\Phi$ and $\Psi$, in the form of Cauchy-type integrals

$$\phi(z)=\frac{1}{2\pi i}\int_{\Gamma}\frac{\omega(\tau)d\tau}{(\tau-z)},\quad z\in D, \tag{8}$$

and

$$\psi(z)=\frac{1}{2\pi i}\int_{\Gamma}\frac{\rho(\tau)d\tau}{(\tau-z)},\quad z\in D, \tag{9}$$

or

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Dec. 2, 1999; final revision, May 10, 2000. Associate Technical Editor: J. R. Barber. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

 **Copyright © 2000 by ASME** 

$$\Phi(z)=\frac{1}{2\pi i}\int_{\Gamma}\frac{\Omega(\tau)d\tau}{(\tau-z)},\quad z\in D, \tag{10}$$

and

$$\Psi(z)=\frac{1}{2\pi i}\int_{\Gamma}\frac{\Xi(\tau)d\tau}{(\tau-z)},\quad z\in D, \tag{11}$$

where $\omega$ and $\rho$, or $\Omega$ and $\Xi$, are unknown layer densities on $\Gamma$. Values of the potentials $\phi$, $\psi$, $\Phi$, and $\Psi$ on $\Gamma$ are defined as limits of $\phi$, $\psi$, $\Phi$, and $\Psi$ in $D$ as $\Gamma$ is approached. Since the equations of elasticity now are satisfied everywhere, it remains only to solve the problem which consists of enforcing the boundary condition of prescribed traction $(t_x^{\mathrm{pr}},t_y^{\mathrm{pr}})$ along $\Gamma$. This can be done in various ways, leading to various integral equations.

## 3 The Classic Sherman-Lauricella Integral Equation

A classic choice for the interior stress problem is to take the unknown layer density $\rho$ of (9) in the following way ([6]):

$$\rho(z)=\overline{\omega(z)}-\overline{z}\omega'(z). \tag{12}$$

The choice (12) makes $\psi$ of (9) assume the form

$$\psi(z)=\frac{1}{2\pi i}\int_{\Gamma}\frac{\omega(\tau)d\overline{\tau}}{(\tau-z)}+\frac{1}{2\pi i}\int_{\Gamma}\frac{\overline{\omega(\tau)}d\tau}{(\tau-z)}-\frac{1}{2\pi i}\int_{\Gamma}\frac{\overline{\tau}\omega(\tau)d\tau}{(\tau-z)^2}. \tag{13}$$

The requirement of prescribed traction on $\Gamma$ leads, via (3), to the Lauricella integral equation for $\omega$

$$(I+M_{\mathrm{SL}})\omega(z)=g(z),\quad z\in\Gamma, \tag{14}$$

accompanied with the solvability conditions that $g(z)$ must be single valued and

$$Q_1 g=0. \tag{15}$$

In (14)–(15) the notation $g(z)$ has been introduced for the integral of traction along $\Gamma$ from a point $z(s_0)$ as

$$g(z)=i\int_{s_0}^{s(z)}t\,ds,\quad z\in\Gamma,$$

where $t=t_x^{\mathrm{pr}}+it_y^{\mathrm{pr}}$, and the operator $Q_1$ is a mapping from $\Gamma$ to $\mathbb{R}$, defined by

$$Q_1 g=\frac{1}{S}\Re\mathrm{e}\left\{\int_{\Gamma}g(z)d\overline{z}\right\}, \tag{16}$$

where $S$ is the perimeter of the body, and $M_{\mathrm{SL}}$ is a compact integral operator given by

$$M_{\mathrm{SL}}\omega(\tau)=\frac{1}{2\pi i}\left[\int_{\Gamma}\frac{\omega(\tau)d\tau}{(\tau-z)}-\int_{\Gamma}\frac{\omega(\tau)d\overline{\tau}}{(\overline{\tau}-\overline{z})}-\int_{\Gamma}\frac{\overline{\omega(\tau)}d\tau}{(\overline{\tau}-\overline{z})}+\int_{\Gamma}\frac{(\tau-z)\overline{\omega(\tau)}d\overline{\tau}}{(\overline{\tau}-\overline{z})^2}\right]. \tag{17}$$

Consider now the integral operator $B$ suggested by Sherman [1] and defined by

$$B\omega(z)=\left(\frac{1}{(z-z^*)}-\frac{1}{(\overline{z}-\overline{z}^*)}+\frac{(z-z^*)}{(\overline{z}-\overline{z}^*)^2}\right)\times\frac{1}{\pi i}\Re\mathrm{e}\left\{\int_{\Gamma}\frac{\omega(\tau)d\tau}{(\tau-z^*)^2}\right\},\quad z\in\Gamma, \tag{18}$$

where $z^*$ is an arbitrary point in $D$. Parton and Perlin [5] suggest a simpler operator $B$

$$B\omega(z)=\frac{1}{(\overline{z}-\overline{z}^*)}\frac{1}{\pi i}\Re\mathrm{e}\left\{\int_{\Gamma}\frac{\omega(\tau)d\tau}{(\tau-z^*)^2}\right\},\quad z\in\Gamma. \tag{19}$$

Addition of the operator $B$ to the left-hand side of (14) gives the Sherman-Lauricella integral equation

$$(I+M_{\mathrm{SL}}+B)\omega(z)=g(z),\quad z\in\Gamma. \tag{20}$$

Uniqueness of the solution to (20), with the choice (18) for $B$, is proven in paragraph 56 of Mikhlin [4]. Uniqueness, with the choice (19), is proven in paragraph 19 of Parton and Perlin [5].

Once Eq. (20) is solved for $\omega$, various quantities of physical interest can be computed. The displacement on $\Gamma$, for example, can be obtained from

$$u_x+iu_y=\frac{1}{2}\left(\frac{1}{\mu}+\frac{1}{\kappa}\right)(I+M_1)\omega(z)-\frac{g(z)}{2\mu},\quad z\in\Gamma, \tag{21}$$

where

$$M_1\omega(z)=\frac{1}{\pi i}\int_{\Gamma}\frac{\omega(\tau)d\tau}{(\tau-z)},\quad z\in\Gamma. \tag{22}$$

From the viewpoint of numerical efficiency, the choice between (18) and (19) for $B$ is perhaps not so important. Greenbaum, Greengard, and Mayo [7] use (20) with (18). Greengard, Kropinski, and Mayo [8] and Strandberg [9] use (20) with (19). No author comments on the relative merits of the two choices. The next section presents yet another choice $B$. As we shall see in the last section, this new ''twist'' can make a substantial difference.

## 4 A New Operator *B*

Equation (14) with the solvability condition (15) does not have a unique solution. The operator on the left-hand side of (14) is rank-one deficient. On the unit disk, for example, a null-vector is $\omega_n=iz$. We suggest the uniqueness condition

$$Q_1(I+M_1)\omega=0. \tag{23}$$

We see, from (16) and (21), that the condition (23) has a physical interpretation in terms of average tangential displacement on $\Gamma$. Two useful relations are

$$Q_1(I+M_{\mathrm{SL}})\omega=0, \tag{24}$$

$$Q_1 in=1. \tag{25}$$

We are now in the position to propose a new equivalent formulation for (14) and (23), assuming that (15) holds. The new formulation is based on the choice

$$B\omega(z)=\frac{in}{2}Q_1(I+M_1)\omega(z),\quad z\in\Gamma. \tag{26}$$

This choice for $B$ differs from the choices (18) and (19) in two respects: It has a clear physical interpretation and, more importantly, it does not involve the arbitrary point $z^*$. The Sherman-Lauricella equation now reads

$$\left(I+M_{\mathrm{SL}}+\frac{in}{2}Q_1(I+M_1)\right)\omega(z)=g(z),\quad z\in\Gamma. \tag{27}$$

Equation (27) trivially follows from (14) and (23). To prove the converse, we apply $Q_1$ from the left in (27) and use the relations (24)–(25) and (15). This gives (23). Subtraction of (23) from (27) gives back (14).

Uniqueness of the solution to (27) can be proven using the same technique as in paragraph 56 of Mikhlin [4] and observing the relations

$$\phi(z)=\frac{1}{2}(I+M_1)\omega(z),\quad z\in\Gamma, \tag{28}$$

and

$$Q_1 iz=2A/S, \tag{29}$$

where $A$ is the area of the body.

## 5 A Modified Muskhelishvili Equation

An interesting extension for the interior stress problem is to let it involve a problem exterior to $D$. The exterior problem is one where the prescribed traction on $\Gamma$ is zero and the stress at infinity is zero. We shall seek $\Phi$ and $\Psi$ such that the two problems are solved simultaneously. Clearly, $\Phi$ and $\Psi$ are zero outside $D$. This follows from the uniqueness of the solution to the second fundamental exterior problem in the plane ([4]), and implies that $\Omega$ of (10) and $\Xi$ of (11) are boundary values of analytic functions in $D$. Now we choose $\Omega$ to be the value of $\Phi$ on $\Gamma$, and choose $\Xi$ in such a way that the traction is zero outside $\Gamma$ and jumps a quantity $t$ as $\Gamma$ is crossed. The jump condition makes $\Psi$ of (11) take the form

$$\Psi(z)=-\frac{1}{2\pi i}\int_\Gamma \frac{\overline{\Phi(\tau)}d\overline{\tau}}{(\tau-z)}-\frac{1}{2\pi i}\int_\Gamma \frac{\overline{\tau}\Phi(\tau)d\tau}{(\tau-z)^2}-\frac{1}{2\pi i}\int_\Gamma \frac{\overline{n}\overline{t}d\tau}{(\tau-z)},\quad z\in D. \tag{30}$$

The requirement that the traction outside $\Gamma$ is zero leads to the following integral equation for $\Phi$ on $\Gamma$

$$(I-M_3)\Phi(z)=\frac{\overline{n}t(z)}{2}+\frac{\overline{n}}{n}\frac{1}{2\pi i}\int_\Gamma \frac{ntd\overline{\tau}}{(\overline{\tau}-\overline{z})},\quad z\in\Gamma, \tag{31}$$

accompanied with the solvability condition

$$Q_2\overline{n}t=0, \tag{32}$$

where

$$Q_2 f=-\frac{1}{2A}\Re\mathrm{e}\left\{\int_\Gamma f(z)\overline{z}dz\right\}. \tag{33}$$

In (31) $M_3$ is a compact integral operator given by

$$M_3\Phi(z)=\frac{1}{2\pi i}\left[\int_\Gamma \frac{\Phi(\tau)d\tau}{(\tau-z)}+\frac{\overline{n}}{n}\int_\Gamma \frac{\Phi(\tau)d\tau}{(\overline{\tau}-\overline{z})}+\int_\Gamma \frac{\overline{\Phi(\tau)}d\overline{\tau}}{(\overline{\tau}-\overline{z})}+\frac{\overline{n}}{n}\int_\Gamma \frac{(\tau-z)\overline{\Phi(\tau)}d\overline{\tau}}{(\overline{\tau}-\overline{z})^2}\right]. \tag{34}$$

Equation (31) can be viewed as the derivative of the conjugate of Eq. (7) in paragraph 54 of Mikhlin [4]. That equation was originally derived by Muskhelishvili.

A quantity of physical interest, which can be computed once (31) is solved for $\Phi$, is the complex tangential derivative of the displacement on $\Gamma$

$$\frac{d}{dz}(u_x+iu_y)=\left(\frac{1}{\kappa}+\frac{1}{\mu}\right)\Phi(z)-\frac{\overline{n}t}{2\mu},\quad z\in\Gamma. \tag{35}$$

The following Lemma will be useful for proving Eq. (39) below.

LEMMA 5.1.

$$Q_2 i=1, \tag{36}$$

$$Q_2(I-M_3)\Phi=0. \tag{37}$$

*Proof.* Equation (36) is proven by applying Gauss' theorem. Equation (37) is proven by expressing $(I-M_3)\Phi$ explicitly in terms of analytic potentials and then applying Cauchy's theorem. □

Equation (31) with the solvability condition (32) does not have a unique solution. The operator on the left-hand side of (31) is rank-one deficient. Imaginary constants are null-solutions. For any solution $\Phi$ we can form a new solution as $\Phi+i\alpha$, where $\alpha$ is a real constant. See paragraph 54 of Mikhlin [4] for a similar result for the null-space of an operator derived for the potential $\phi$. Here we propose the uniqueness condition

$$Q_2\Phi=0. \tag{38}$$

We see, from (33) and (35), that the condition (38) has the same physical interpretation as the condition (23).

We are now in the position to propose a new formulation for (31) and (38), assuming that (32) holds.

THEOREM 1. *Given the solvability condition (32), equation (31) and the uniqueness condition (38) are equivalent to the following Fredholm equation of the second kind*

$$(I-M_3+iQ_2)\Phi(z)=\frac{\overline{n}t(z)}{2}+\frac{\overline{n}}{n}\frac{1}{2\pi i}\int_\Gamma \frac{ntd\overline{\tau}}{(\overline{\tau}-\overline{z})},\quad z\in\Gamma. \tag{39}$$

*Proof.* Equation (39) trivially follows from (31) and (38). To prove the converse, we apply $Q_2$ from the left in (39) and use the relations (36)–(37) and (32). This gives (38). Subtraction of (38) from (39) gives back (31). □

Uniqueness of the solution to (39) can be proven using the method of paragraph 54 in Mikhlin [4]. First one proves that an assumed homogeneous solution, $\Phi_0$, to (39) has to be an imaginary constant. Then (38), which is implied by (39), gives that this constant is zero.

## 6 Numerical Comparison Between Formulations

In this section we undertake a comparison between algorithms for the classic Sherman-Lauricella Eq. (20) with the choice (19) for $B$, for the modified formulation (27), and for the modified Muskhelishvili Eq. (39). The algorithms are of Nyström type based on composite 16-point Gaussian quadrature and the GMRES iterative solver ([10]). The iterations are terminated when the residual is as small as it can get, which typically means $2\cdot 10^{-15}$. Compensated summation ([11,12]) is used for the computation of matrix-vector multiplications and inner products in the GMRES iterative solver. For details on how to regularize the Cauchy-type singular operator $M_1$ of (22), see Helsing and Jonsson [13].

For setups with smooth boundaries and analytical solutions, such as loaded circular or elliptic disks, it is hard to say which equation leads to the best algorithm. Algorithms based on the three equations all require only a few GMRES iterations for full convergence. Nontrivial examples are needed in order to detect differences in performance.

When comparing the performance of the algorithms below, we need a reference quantity $q_{\mathrm{ref}}$ to measure accuracy against. We have decided to use the $L^2$ norm of the hydrostatic stress on $\Gamma$, that is

$$q_{\mathrm{ref}}=\left(\int_\Gamma (\sigma_{xx}(z)+\sigma_{yy}(z))^2 ds\right)^{1/2}, \tag{40}$$

as such a reference quantity.

*Example 1: A Symmetric Starfish.* We first consider a body in the shape of a nine-armed starfish parameterized by

$$z(t)=(1+0.36\cos 9t)e^{it},\quad 0\leqslant t<2\pi. \tag{41}$$

The load is chosen as

$$g(z)=z^2. \tag{42}$$

We start out with testing the algorithm for the classic Eq. (20) with $B$ as in (19). The starfish of (41) is symmetric with respect to the origin. A natural choice for the arbitrary point is therefore $z^*=0$. With this choice for $z^*$ the algorithm for the classic Eq. (20) requires 2000 discretization points to reach a relative error in $q_{\mathrm{ref}}$ of $10^{-12}$. Upon increased resolution the quality of the solution slowly gets worse. (See Fig. 1.) The number of GMRES iterations required is 25. The sensitivity to the placement of the arbitrary point $z^*$ turns out to be quite large in this example. When the position of $z^*=0$ is changed a tiny distance to $z^*=0.01i$, which

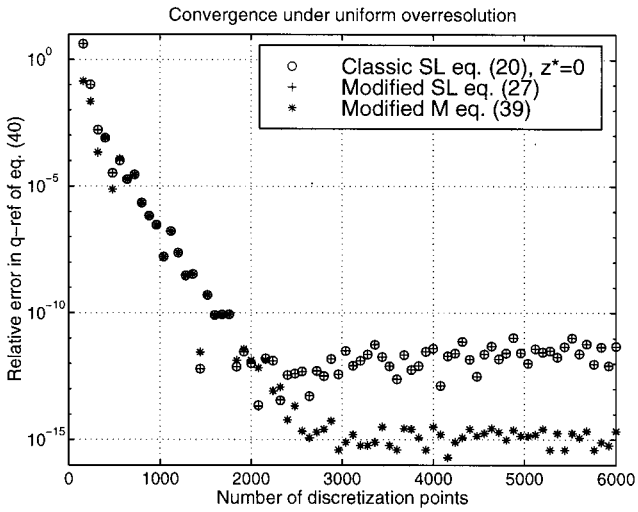



**Fig. 1 Example 1. Convergence of the reference quantity $q_{\mathrm{ref}}$ of (40), defined as the Euclidean norm of the hydrostatic stress on the boundary, for algorithms based on the classic Eq. (20) with $z^*=0$, the modified formulation (27), and the modified Eq. (39). The correct value, $q_{\mathrm{ref}}=71.79088302407723$ was computed using quadruple precision arithmetic.**

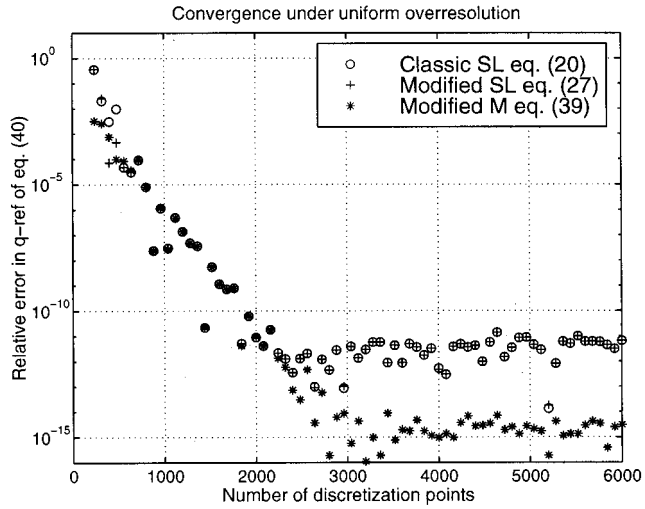



**Fig. 3 Example 2. Convergence of the reference quantity $q_{\mathrm{ref}}$ of (40) for algorithms based on the classic Eq. (20) with $z^*$ placed at the center of gravity, the modified formulation (27), and the modified Eq. (39). The correct value, $q_{\mathrm{ref}}=73.45130087430866$ was computed using quadruple precision arithmetic.**

still is far away from the boundary of the starfish contour, the number of GMRES iterations needed for convergence more than doubles (see Fig. 2).

The algorithm for the modified formulation (27) gives results which are almost identical to those of (20) with the optimal choice $z^*=0$. (See Fig. 1.) The algorithm for the modified formulation requires also about 2000 discretization points for a relative error in result $q_{\mathrm{ref}}$ of about $10^{-12}$. Solving the system of linear equations takes 24 GMRES iterations, which is one iteration less than for (20) with the optimal choice $z^*=0$. (See Fig. 2.)

The algorithm for the modified Eq. (39) has, by far, the best stability properties in this example. Like the previous formulations it gives a relative error in $q_{\mathrm{ref}}$ of about $10^{-12}$ for 2000 discretization points, but as the resolution is increased the relative error in $q_{\mathrm{ref}}$ decreases further and stabilizes on about $2\cdot 10^{-15}$. Equation (39) also gives the best results for underresolved calculations. (See Fig. 1.)

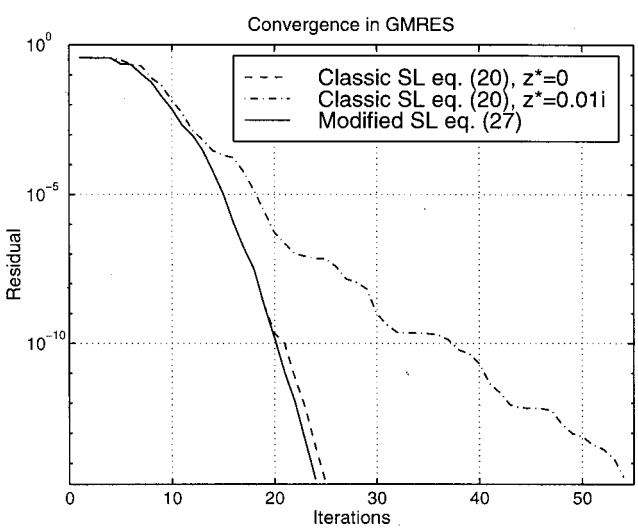



**Fig. 2 Example 1. The iteration history from GMRES for the classic Eq. (20) and the modified formulation (27). The number of discretization points is 2080. The iterations are terminated when the residual is less than $5\cdot 10^{-15}$.**

*Example 2: An Irregular Starfish*. The starfish of (41) has an obvious symmetry point $z=0$ which, as we have seen, is the optimal choice for the arbitrary point $z^*$. To investigate the properties of the three algorithms under more general conditions we perturb the geometry in the previous example so that all arms of the starfish have different shapes

$$z(t)=(1+0.1\sin t+0.36\cos 9t)e^{it},\quad 0\leqslant t<2\pi. \tag{43}$$

The load is changed to

$$g(z)=z^2-\frac{izSQ_1z^2}{2A}, \tag{44}$$

so that the solvability conditions (15) and (32) still are satisfied.

The convergence properties for the three algorithms in this example turn out to be similar to those in Example 1. (See Fig. 3.) The main difference is that GMRES now requires 105 iterations for full convergence instead of the 25 iterations in Example 1. An increase in the number of iterations could be expected since the irregular starfish of (43) constitutes a more difficult geometry than the symmetric starfish of (41).

The modified Eq. (39) still gives the most stable convergence and still performs best for underresolved calculations. The algorithm for the classic Eq. (20), with reasonable placements of the arbitrary point $z^*$, still gives results which are similar to, or only slightly worse than, those of the modified formulation (27). (See Fig. 3.) One difference is worth pointing out: The number of GMRES iterations needed for convergence with the classic Eq. (20), and for fully resolved calculations, is less sensitive to the placement of $z^*$ in Example 2 than in Example 1. There seems to be no interior point that is obviously optimal.

## 7 Conclusions and Outlook

We conclude that an algorithm based on the modified formulation (27) has shown to be equally or more efficient for interior

stress problems than an algorithm based on the classic Sherman-Lauricella Eq. (20). The chief advantage with (27) over (20) is that (27) omits the need for the arbitrary point $z^*$. A nonoptimal choice for $z^*$ can in certain situations greatly deteriorate the performance of iterative algorithms based on the classic Eq. (20). In addition, removal of the arbitrary point $z^*$ is an advantage from a coding viewpoint.

We further conclude that an algorithm based on the modified Muskhelishvili Eq. (39) shows superior stability properties compared to both the classic (20) and the modified (27) Sherman-Lauricella equations and that (39) is best for underresolved calculations. This is so, since in (39) we solve for the stress potential $\Phi$ on $\Gamma$ directly, while in (27) we solve for the density $\omega$. The density $\omega$ is related to $\Phi$ on $\Gamma$ via a transform containing differentiation. Differentiation is in itself an ill-conditioned operation.

In a forthcoming paper we intend to apply (39) to the problem of computing so-called notch intensity factors of loaded rectangular specimens.

## Acknowledgment

This work was supported by NFR, TFR, and The Knut and Alice Wallenberg Foundation under TFR contract 99-380.

## References

[1] Sherman, D. I., 1940, ''Sur la Solution du Second Problème Fondamental de la Théorie Statique Plane de l'Élasticité,'' Comptes Rendus (Doklady) l'Acad. Sci. l'URSS, **28**, pp. 25–32.
[2] Muskhelishvili, N. I., 1953, *Some Basic Problems of the Mathematical Theory of Elasticity*, P. Noordhoff, Ltd., Groningen.
[3] Sokolnikoff, I. S., 1956, *Mathematical Theory of Elasticity*, McGraw-Hill, New York.
[4] Mikhlin, S. G., 1957, *Integral equations*, Pergamon Press, London.
[5] Parton, V. Z., and Perlin, P. I., 1982, *Integral Equation Methods in Elasticity*, MIR, Moscow.
[6] Lauricella, G., 1909, ''Sur l'Intégration de l'Équation Relative à l'Équilibre des Plaques Élastiques Encastrées,'' Acta Math., **32**, pp. 201–256.
[7] Greenbaum, A., Greengard, L., and Mayo, A., 1992, ''On the Numerical Solution of the Biharmonic Equation in the Plane,'' Physica D, **60**, pp. 216–225.
[8] Greengard, L., Kropinski, M. C., and Mayo, A., 1996, ''Integral Equation Methods for Stokes Flow and Isotropic Elasticity in the Plane,'' J. Comput. Phys., **125**, pp. 403–414.
[9] Strandberg, M., 1999, ''A Numerical Study of the Stress Fields Arising From Sharp and Blunt V-Notches in a SENT-Specimen,'' Int. J. Fract., in press.
[10] Saad, Y., and Schultz, M. H., 1986, ''GMRES: A Generalized Minimum Residual Algorithm for Solving Nonsymmetric Linear Systems,'' SIAM J. Sci. Stat. Comput., **7**, pp. 856–869.
[11] Kahan, W., 1965, ''Further Remarks on Reducing Truncation Errors,'' Commun. ACM, **8**, p. 40.
[12] Higham, N. J., 1996, *Accuracy and Stability of Numerical Algorithms*, SIAM, Philadelphia, PA, pp. 92–97.
[13] Helsing, J., and Jonsson, A., 1999, ''Elastostatics for Plates With Holes,'' *Department of Solid Mechanics KTH report 99-250*.

**W. D. Armstrong**[1]
Assistant Professor,
Thomas J. Watson School of Engineering
and Applied Science,
Department of Mechanical Engineering,
State University of New York,
Binghamton, NY 13902-6000
Assoc. Mem. ASME

**V. Kumar**
Associate Professor,
Department of Mechanical Engineering,
University of Washington,
Seattle, WA 98195
Mem. ASME

# A Stress and Relative Density-Dependent Dynamic Compliance Spectra Model of the Creep Response of Microcellular Polycarbonate

*The present paper reports a new discrete complex compliance spectra method in which each frequency component is a direct function of stress and relative density. Comparisons between model calculations and experimental measurements show that the model exhibits excellent quantitative agreement at all experimental stress-relative density states and significantly smoothes the experimental input data. It is anticipated that design engineers will use the present method to accurately predict the creep strain histories resulting from a broad range of specific stress-relative density combinations.* [S0021-8936(00)01804-3]

## 1 Introduction

Microcellular foams are generally closed-cell foams with average cell size of order 10 micrometers ([1]). Such foams are characterized by a high density of cells, usually exceeding 100 million cells per cubic centimeter. The high cell density makes it possible to achieve significant reductions in density of the polymer while keeping the average cell size small. Such foams have been created in polycarbonate using carbon dioxide as the gas for bubble nucleation ([2]). The polycarbonate-carbon dioxide system has been a model microcellular system, and a number of investigations have focused on the properties of this system. For example, tensile, fatigue, and fracture toughness studies have uncovered some unique properties of these novel foams ([3–6]). A key advantage offered by microcellular foams is the potential to reduce the amount of material needed by replacing the solid polymer by a microcellular polymer of reduced density. As polycarbonate is used in many load-bearing applications, a knowledge of the creep behavior of microcellular polycarbonate is of special interest. An experimental investigation of the time dependent response of microcellular polycarbonate foams was conducted by Wing [7] who found that creep strains of the foams contain a significantly higher viscoplastic component compared to the unfoamed material. Wing obtained creep data on foams of relative densities (density of foam devised by the density of polycarbonate) of 1.00, 0.95 and 0.81. The specimens were subjected to various levels of constant tension for eight hours.

Wing et al. [8] have presented a model for the creep response of microcellular polycarbonate based on the Schapery's theory of nonlinear viscoelasticity. The present paper reexamines the experimental measurements of Wing [7] with a new stress and relative density-dependent discrete dynamic compliance spectra method. Comparison between model calculations and experimental measurements show that the new spectra model exhibits excellent quantitative agreement at all experimental stress-relative density states. A materials scientist may use the present model interactively to design experiments which will allow the accurate modeling of nonlinear creep properties of the microcellular material over a broad range of stress and relative density. A designer may use the present model to identify the creep strain as a function of time resulting from a given stress-relative density combination, or to identify relative density range over which the material provides relatively stable creep properties.

The present model requires a set of creep curves measured under variable stress and material relative density as input, and then predicts the creep strain under general conditions of time, stress, and relative density as output. The method in essence uses many thousands of creep parameters, a process made mathematically tractable by the use of fast Fourier transforms. This method was selected in order to construct a single mathematical entity which may be accurately applied to the prediction of the response of a highly nonlinear viscoelastic media over very broad stress and relative density conditions. The method is quite general, and may be applied to many other multivariable time-dependent problems in engineering.

## 2 A Stress and Relative Density-Dependent Discrete Dynamic Compliance

The present method will make use of fast Fourier transforms ([9]). The first analysis task is, therefore, to fit time-sampled experimental creep data with an appropriate set of continuous functions which may be evaluated at a regular interval. The present calculation uses an $n=0$ to 2 series of Chebyshev polynomials of the first kind for this purpose, the Chebyshev polynomials of the first kind, $T_n$, being defined by

$$T_n(\mathrm{Cos}\,\theta)=\mathrm{Cos}(n\theta) \tag{1}$$

and obeying the orthogonality condition.

$$\int_{-1}^{1} T_m(x)T_n(x)(1-x^2)^{-1/2}dx=0, \quad \{m\neq n\} \tag{2}$$

A discrete creep strain frequency spectra for each of the individual creep tests may now be constructed by fast Fourier transforming a strain list with an initial element equal to zero and subsequent elements obtained for the creep strain time function ([10–16]). The present calculations set the first strain time sample equal to zero and then evaluate the fitted creep strain functions every 10 seconds from 1 to 651 seconds to produce a 67-element

[1]To whom correspondence should be addressed.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Aug 25, 1998; final revision, July 5, 2000. Associate Technical Editor: I. M. Daniel. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

 **Copyright © 2000 by ASME** 

time-sample list. The experimental discrete strain frequency components may then be obtained by fast Fourier transformation ([15])

$$\tilde{E}\left(\frac{n}{NT}\right)=\frac{1}{\sqrt{N}}\sum_{k=0}^{N-1}\varepsilon(kT)\exp\left[2\pi i\frac{nk}{N}\right] \tag{3}$$

where $N$ is the number of elements in the time sample list and $T$ is the time-sampling interval. The same numerical procedure is used for the stress, assuming an ideal creep stress step time sequence.

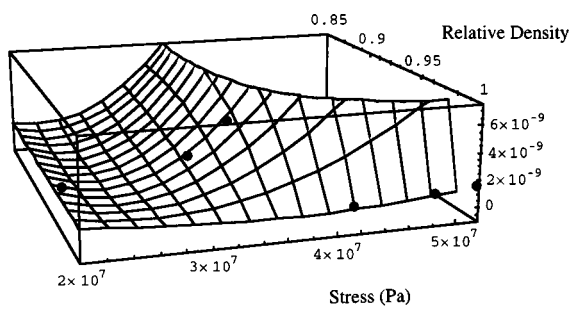
$$\sigma(kT)=\{0,\sigma_0,\sigma_0,\ldots,\sigma_0\} \tag{4}$$

A discrete dynamic compliance may then be obtained by complex division of stress frequency components with corresponding strain frequency components.

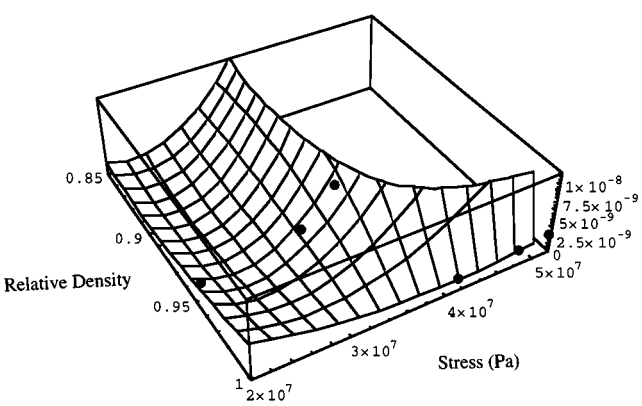
$$\tilde{J}=\left\{\frac{\tilde{E}\left(\frac{0}{NT}\right)}{\tilde{S}\left(\frac{0}{NT}\right)},\frac{\tilde{E}\left(\frac{1}{NT}\right)}{\tilde{S}\left(\frac{1}{NT}\right)},\frac{\tilde{E}\left(\frac{2}{NT}\right)}{\tilde{S}\left(\frac{2}{NT}\right)},\ldots,\frac{\tilde{E}\left(\frac{N-1}{NT}\right)}{\tilde{S}\left(\frac{N-1}{NT}\right)}\right\} \tag{5}$$

We now assume that the real and imaginary parts of each discrete frequency component smoothly vary with applied stress and relative density. We may then assemble a stress and relative density-dependent discrete dynamic compliance spectra by regression for the real and imaginary parts of each experimental discrete dynamic compliance component. The regression basis used for the present calculations is given by

$$\sum_{i=0}^{2}\sum_{k=0}^{2}T_i(\rho_{\mathrm{rel}})T_k\left(\frac{\sigma}{60\,10^6}\right). \tag{6}$$




(a)




(b)

**Fig. 1 The stress-relative density dependencies of the real, (*a*), and imaginary, (*b*), parts of the third discrete dynamic compliance component. Dots indicate experimental measurements.**

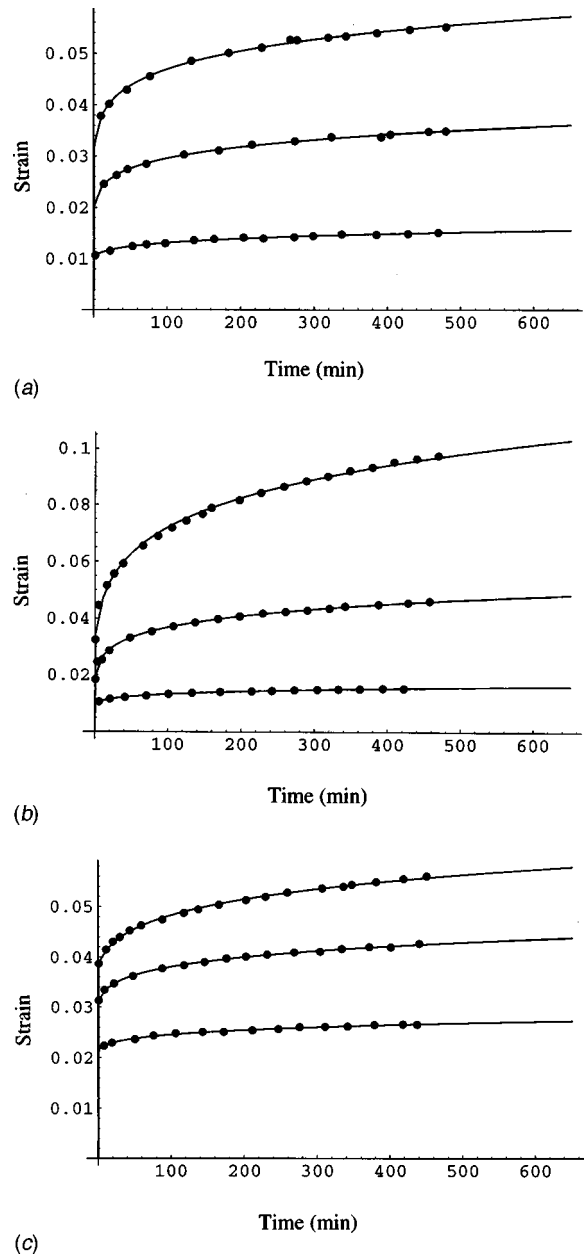



**Fig. 2 Comparison between experimental creep measurements and model calculations; (*a*) relative density=0.81, *s* =27.6, 24.1, and 13.8 Mpa, (*b*) relative density=0.95, *s*=34.5, 31.0, and 20.7 Mpa, and (*c*) relative density=1.00, *s*=51.7, 48.2, and 41.3 Mpa.**

Figure 1 shows the real, (*a*), and imaginary, (*b*), response surface of the third discrete dynamic compliance frequency component as functions of relative density and applied stress. Experimental stress-relative density states are identified with black dots. The figure shows that the real and complex parts of the dynamic compliance appear to be smooth, easily fit functions of stress and relative density within the experimental range. The present creep behavior is clearly under sampled at low relative density and high stress. An additional measurement at a relative density of approximately 0.81 and a stress of approximately 34 MPa would very likely have improved model predictions in this stress-relative density region.

## 3 Evaluation of the Creep Model

The calculation of creep processes begins by the evaluation of the real and complex parts of the stress-dependent dynamic com-

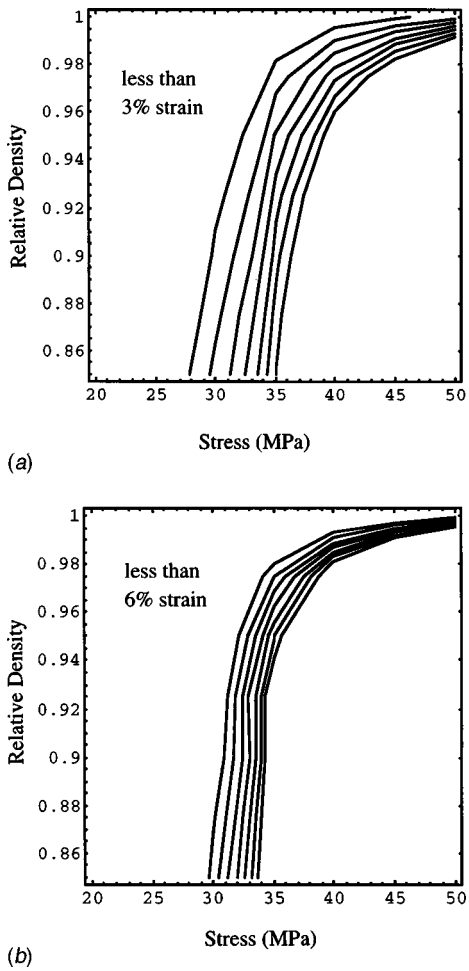



**Fig. 3 Predicted creep strain as a function of applied stress and relative density; (*a*) creep strain after 31 minutes with one percent creep contours ranging evenly from 3 percent strain to 9 percent strain, and (*b*) creep strain after 501 minutes with one percent creep contours ranging evenly from six percent strain to 12 percent strain.**

pliance spectra for a particular level of applied stress and relative density. The discrete creep strain frequency spectrum may then be obtained by multiplying dynamic compliance spectra components by corresponding stress frequency components. The creep strain as a function of time is then obtained by inverse discrete Fourier transformation.

$$\varepsilon(kT)=\frac{1}{\sqrt{N}}\sum_{n=0}^{N-1}\tilde{E}\left(\frac{n}{NT}\right)\exp\left[-2\pi i\frac{nk}{N}\right] \tag{7}$$

Figure 2 compares experimental creep measurements with model calculations at each experimental applied stress-relative density condition. The figure shows that the model exhibits excellent quantitative agreement at all experimental stress-relative density states. The model also provides further benefit by significantly smoothing the experimental data. The ability of the model to precisely match all experimental measurements along with the smoothness of the dynamic compliance response surfaces indicate that the model should provide accurate interpolation and some range of accurate extrapolation to nonexperimental stress-relative density states when the dynamic compliance surfaces are well sampled. A materials scientist may therefore use the present model interactively to design experiments which will allow the accurate modeling of the nonlinear creep of the microcellular materials over a broad range of stress and relative density.

The model may be further used to calculate the creep strain corresponding to any time sample coordinate as a generalized function of stress and relative density. Figure 3 shows two examples of how the resulting isochronal creep prediction may be displayed as a contour plot. A designer may use such a plot to identify stress-relative density combinations which will not result in an excessive accumulation of creep strain, or to identify relative density ranges over which the microcellular material provides stable creep properties. For example, Fig. 3(*a*) shows that a microcellular polycarbonate sample with a relative density of 0.86 under a 30 Mpa stress will strain nearly the same amount after 501 seconds as a 0.99 relative density sample under a 40 Mpa stress. The plot further shows that the creep of microcellular polycarbonate is very sensitive to changes in relative density when the relative density is between 1 and 0.96, and much less sensitive to changes in relative density when the relative density is between 0.95 and 0.85. Therefore, it would appear to be very difficult to obtain consistent creep measurements from materials with relative densities above 0.95.

## 4 Summary and Conclusions

The present paper reports a new stress and relative density-dependent discrete dynamic compliance spectra method which may be used to model the creep behavior of nonlinear viscoelastic microcellular materials. Comparisons between model calculations and experimental measurements show that the model exhibits excellent quantitative agreement at all experimental stress-relative density states. The model also significantly smoothes the experimental data. A materials scientist may use the present model interactively to design experiments which will allow the accurate modeling of the nonlinear creep properties of the microcellular material over a broad range of stress and relative density. A designer may use the present model to identify the creep strain resulting from a given stress-relative density combination after a given amount of time, or to identify relative density ranges over which the material provides relatively stable creep properties.

## Acknowledgment

The support of the Department of the Navy, Office of Naval Research under contract N00014-99-1-0276 and the direction of Program Officer Jan Lindberg is gratefully acknowledged; the content of this paper does not necessarily reflect the position or the policy of the Government, and no official endorsement should be inferred.

## References

[1] Kumar, V., 1993, ''Microcellular Polymers: Novel Materials for the 21st Century,'' Prog. Rubber Plast. Tech., **9**, No. 1, pp. 54–70.
[2] Kumar, V., and Weller, J. E., 1994, ''Production of Microcellular Polycarbonate Using Carbon Dioxide for Bubble Nucleation,'' ASME J. Eng. Ind., **116**, pp. 413–420.
[3] Kumar, V., et al., 1994, ''Characterization of Tensile Behavior of Microcellular Polycarbonate,'' ASME J. Eng. Mater. Technol., **116**, pp. 439–445.
[4] Kumar, V., and Weller, J. E., 1994, ''A Model for the Unfoamed Skin on Microcellular Foams,'' Polym. Eng. Sci., **34**, pp. 169–173.
[5] Seeler, K. A., and Kumar, V., 1993, ''Tension-Tension Fatigue of Microcellular Polycarbonate: Initial Results,'' J. Reinf. Plast. Compos., **12**, No.3, pp. 359–376.
[6] Barlow, C., 1997, Masters thesis, ''The Effects of Microstructure on the Fracture Behavior of Microcellular Polycarbonate,'' Department of Materials Science and Engineering, University of Washington, Seattle.
[7] Wing, G., 1993, Masters thesis, ''Time Dependent Behavior of Microcellular Polycarbonate,'' Department of Materials Science and Engineering, University of Washington, Seattle.
[8] Wing, G., et al., 1995, ''Time Dependent Response of Polycarbonate and Microcellular Polycarbonate,'' Polym. Eng. Sci., **35**, No. 8, pp. 673–679.
[9] Brigham, E. O., 1974, *The Fast Fourier Transform*, Prentice-Hall, Englewood Cliffs, NJ.
[10] Arridge, R. G. C., and Barham, P. J., 1986, ''Fourier Transform Mechanical Spectroscopy,'' J. Phys. D: Appl. Phys., **19**, pp. L89–L96.
[11] Holly, E. E., et al., 1988, ''Fourier Transform Mechanical Spectroscopy of Viselastic Materials With Transient Structure,'' J. Non-Newtonian Fluid Mech., **27**, pp. 17–26.

[12] Kamath, V. M., and Mackely, M. R., 1989, ''Determination of Polymer Relaxation Moduli and Memory Functions Using Integral Transforms,'' J. Non-Newtonian Fluid Mech., **32**, pp. 119–144.
[13] Aspden, R. M., 1991, ''Aliasing Effects in Fourier Transforms of Monotonically Decaying Functions and the Calculation of Viscoelastic Moduli by Combining Transforms Over Different Time Periods,'' J. Phys. D: Appl. Phys., **24**, pp. 803–808.
[14] Holmes, A. D., and Jing, F., 1994, ''The Experimental Nature of Stress Relaxation and Recovery of High-Frequency Information,'' J. Phys. D: Appl. Phys., **27**, pp. 2475–2479.
[15] Armstrong, W. D., 1998, ''A Stress Dependent Dynamic Compliance Spectra Approach to the Nonlinear Viscoelastic Response of Polymers,'' J. Polym. Sci., Part B: Polym. Phys., **36**, pp. 2301–2309.
[16] Armstrong, W. D., and Kumar, V., 2000, ''A Discrete Complex Compliance Spectra Model of the Nonlinear Viscoelastic Creep and Recovery of Microcellular Polymers,'' J. Polym. Sci., Part B: Polym. Phys., **38**, pp. 691–697.

**V. A. Levin**
Mem. ASME,
Department of Mechanics and Mathematics,
Moscow State University,
Moscow 119899, Russia

**V. V. Lokhin**
Institute of Mechanics,
Moscow State University,
Moscow 117234, Russia

**K. M. Zingerman**
Department of Applied Mathematics
and Cybernetics,
Tver State University,
Tver 170000, Russia

# Effective Elastic Properties of Porous Materials With Randomly Dispersed Pores: Finite Deformation

*A method is developed for the analysis of the effective properties of porous nonlinear elastic materials with randomly distributed interacting pores under finite deformations. The method is based on the solution of the problems of nonlinear elasticity for a representative region using Signorini's expansion. The constitutive equations for the matrix material and for the comparison material are written in a form corresponding to Murnaghan's potential. The technique, which is used for ensemble averaging, approximately simulates the uniform distribution of pores. The computations are performed for plane strain, when pores are equal in size, and a circular cylindrical shape in the undeformed state is assumed.* [S0021-8936(00)01802-X]

## 1 Introduction

There are two well-known approaches to the analysis of effective mechanical properties of inhomogeneous materials that can be applied to linear as well as to nonlinear problems. One of them is concerned with bounds on the effective characteristics of inhomogeneous materials using variational principles. The variational principle for linear elasticity has been proposed by Hashin and Shtrikman [1]. The boundaries of the effective properties of nonlinear elastic materials under small deformations have been considered by Willis [2,3] and Ponte Castaneda [4]. Ponte Castaneda and Zaidman [5] analyzed the effect of evolution of microstructure on the effective properties of nonlinear viscous porous materials under quasi-static strains.

Another approach for estimation of effective properties is concerned with the solution of a boundary value problem for a representative region or a cell of inhomogeneous material, followed by homogenization. The effective properties of materials with periodically arranged pores have been analyzed with the use of this approach by Horvay [6] and Filshtinskiy [7] for linear elasticity and by Abeyaratne and Triantafyllidis [8] for nonlinear elasticity. The effective characteristics of linear elastic materials with randomly distributed interacting cracks and pores of different shapes have been investigated by Kachanov [9], Kachanov et al. [10], Mauge and Kachanov [11], and Shafiro and Kachanov [12]. In particular, Mauge and Kachanov [10] performed computations for a range of sample crack arrays positioning cracks with the help of a random number generator.

Levin et al. [13,14] considered a method of estimation of effective properties of porous bodies with uniform and statistically isotropic distribution of pores. In those papers, it is assumed that the centers of pores are located on the nodes of a mesh prescribed in the representative region; the number of pores in this region is fixed; and all possible arrangements of pores are equiprobable.

In the present paper, the approach of Levin et al. [13,14] is further developed. In contrast to those papers, the number of pores in the representative region is not fixed, and the probability of each arrangement of pores is assumed to be a function of porosity. This modification of the approach considered permits one to avoid the repetition of computations for different values of porosity.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, September 9, 1998; final revision, January 3, 2000. Associate Technical Editor: L. T. Wheeler. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler. Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

## 2 Averaging Over a Representative Region

Let us concentrate now on two-dimensional problems. At first, we consider the scheme of averaging over a representative area. Elastic moduli of the comparison material are determined from the following assumption, considered by Horvay [6] and Filshtinskiy [7]: If the comparison material occupies the representative region including pores, the average displacements on each side of this region are equal to the ones for the given porous material at the same loads.

Let the shape of pores be prescribed in the undeformed configuration. Let the Cartesian coordinates with axes $x_1$ and $x_2$ be chosen in this configuration. It is assumed that the representative area $S$ is a parallelogram of sides $a$ and $b$ that make angles of $\alpha$ and $\beta$ with the $x_1$-axis, correspondingly (Fig. 1).

Let A, B, C, D be the vertices of $S$, $\Gamma$ be the boundary of $S$, $\gamma_k$ be the boundaries of pores within $S$ (it is assumed that $\gamma_k$ never intersects $\Gamma$). Let the uniform stresses $\sigma=\sigma_{ij}e_ie_j$ be given on $\Gamma$; $u(x)=u_m(x)e_m$ is the vector of displacements which are brought about by the given stresses. Let $u^e=u^e_me_m$, $\Psi^e=\Psi^e_{ij}e_ie_j$ respectively denote the displacement vector and the deformation gradient of the comparison material at the same loads. Because the loads on the boundary $\Gamma$ are constant and the comparison material is uniform, the strains of the comparison material are affine. Then

$$u^e_m=(\Psi^e_{nm}-\delta_{nm})x_n+c_m\,. \tag{1}$$

Taking into account the definition of the comparison material, we have

$$\int_{AD}u^e_m dl=\int_{AD}u_m dl;\ \int_{BC}u^e_m dl=\int_{BC}u_m dl;$$

$$\int_{AB}u^e_m dl=\int_{AB}u_m dl;\ \int_{DC}u^e_m dl=\int_{DC}u_m dl. \tag{2}$$

It follows from (1), (2) that

 Copyright © 2000 by ASME 

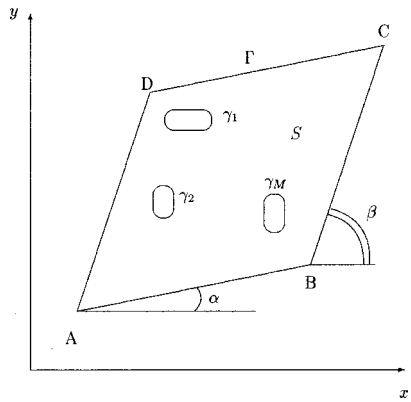



**Fig. 1 Representative area**

$$(\Psi^e_{11}-1)\cos\beta+\Psi^e_{21}\sin\beta=\frac{1}{b}[\langle u_1\rangle|_{\mathrm{DC}}-\langle u_1\rangle|_{\mathrm{AB}}],$$

$$(\Psi^e_{11}-1)\cos\alpha+\Psi^e_{21}\sin\alpha=\frac{1}{a}[\langle u_1\rangle|_{\mathrm{BC}}-\langle u_1\rangle|_{\mathrm{AD}}],$$

$$\Psi^e_{12}\cos\beta+(\Psi^e_{22}-1)\sin\beta=\frac{1}{b}[\langle u_2\rangle|_{\mathrm{DC}}-\langle u_2\rangle|_{\mathrm{AB}}],$$

$$\Psi^e_{12}\cos\alpha+(\Psi^e_{22}-1)\sin\alpha=\frac{1}{a}[\langle u_1\rangle|_{\mathrm{BC}}-\langle u_1\rangle|_{\mathrm{AD}}]. \tag{3}$$

Here $\langle u\rangle|_{\mathrm{PQ}}=(1/|\mathrm{PQ}|)\int_{\mathrm{PQ}}u\,dl$.

Solving the set of linear Eqs. (3), one can obtain $\Psi^e_{ij}$. It should be noted that in the special case, when the representative area is a square whose sides are parallel to the coordinate axes, the set of Eqs. (3) can be reduced to the equations considered by Levin et al. [13,14].

In order to determine the effective moduli for a given arrangement of pores in $S$, we use the technique described by Levin et al. [14]. The nonlinear problem of elasticity is solved using Signorini's expansion technique up to the second order ([15–18]). The solution can be represented in the form

$$u_m=\overset{(0)}{u}_{mij}\sigma_{ij}+\overset{(1)}{u}_{mijkl}\sigma_{ij}\sigma_{kl} \tag{4}$$

Substituting (4) into (3) and solving (3), we have

$$\Psi^e_{mn}=\delta_{mn}+\overset{(0)}{S}_{mnij}\sigma_{ij}+\overset{(1)}{S}_{mnijkl}\sigma_{ij}\sigma_{kl}. \tag{5}$$

If we substitute (5) into the expression for the Green strain tensor

$$\mathbf{E}^e=\frac{1}{2}(\mathbf{\Psi}^e\mathbf{\Psi}^{eT}-\mathbf{I}), \tag{6}$$

and retain in the obtained expression only linear and quadratic terms, we have

$$E^e_{mn}=\overset{(0)}{T}_{mnij}\sigma_{ij}+\overset{(1)}{T}_{mnijkl}\sigma_{ij}\sigma_{kl}. \tag{7}$$

And finally, solving the system (7) by the perturbation technique up to the second order, we can write the effective constitutive equations in a form

$$\Sigma^0_{mn}=\overset{(0)}{C}_{mnij}E^e_{ij}+\overset{(1)}{C}_{mnijkl}E^e_{ij}E^e_{kl}, \tag{8}$$

where $\Sigma^0_{mn}$ are components of the second Piola-Kirchhoff stress tensor

$$\mathbf{\Sigma}^0=(\det\mathbf{\Psi})(\mathbf{\Psi}^T)^{-1}\cdot\mathbf{\Sigma}^0\cdot\mathbf{\Psi}^{-1}. \tag{9}$$

The solution of nonlinear elasticity problems for the representative area with pores is arrived at with the use of a specialized computer algebra system. In addition, we proceed from the assumption of Mori-Tanaka's scheme ([19]). Following Kachanov et al. [10], we solve the problem for given far-field stresses that are assumed to be equal to the average stresses in a matrix $\sigma^M_{ij}$, determined as

$$\sigma^M_{ij}=(1-p)^{-1}\sigma_{ij},$$

where $\sigma_{ij}$ are the average stresses in a porous medium and $p$ is the porosity of a given material.

The interactions between pores within the representative area are taken into account during the solution of the problem of nonlinear elasticity.

## 3 Ensemble Averaging

Let us concentrate now on the scheme of ensemble averaging (which approximately simulates a statistically uniform distribution of pores). We suppose that all the pores assume a circular shape in the undeformed state and are uniform in size. Let the representative area $S$ a square of side $a$, whose sides are parallel to the coordinate axes. The uniform mesh with nodes $(x_i,y_j)$ is constructed in $S$:

$$x_i=(i+1/2)h,\quad y_j=(j+1/2)h, \tag{10}$$

where $i=0,\ldots,N-1$, $j=0,\ldots,N-1$, $h=a/N$; $N>0$ is a given integer.

The arrangements of nodes at $N=2$ and at $N=3$ are shown in Fig. 2, as an example.

It is assumed that centers of pores are located exactly on the nodes of the mesh. For each arrangement of pores satisfying this assumption, the effective moduli $\overset{(0)}{C}_{mnij}$, $\overset{(1)}{C}_{mnijkl}$, are computed using the scheme presented above. Then the averaging over all these arrangements is performed with regard for their probabilities, which depend on porosity. Note that the number of pores in the representative area is not fixed, unlike in Levin et al. [13–14]. Note also that the condition $R<h/2$ is sufficient to avoid intersections of boundaries of pores $\gamma_k$ with one another and with $\Gamma$ ($R$ is the pore radius).

Let $\mathcal{P}$ denote the probability of existence of a pore, whose center is at a certain node. It is assumed that $\mathcal{P}$ doesn't depend on the existence of pores whose centers are in other nodes. Taking into account these assumptions, one can obtain an expression for porosity:

$$p=\mathcal{P}\frac{\pi R^2}{h^2}. \tag{11}$$

From (11) one can express $\mathcal{P}$ in terms of porosity:

$$\mathcal{P}=p\frac{h^2}{\pi R^2}. \tag{12}$$

Let $\mathcal{P}^{(k)}$ denote the probability that the region $S$ contains exactly $k$ pores. With regard for the assumptions made, $\mathcal{P}^{(k)}$ can be represented as

$$\mathcal{P}^{(k)}=C^k_n\mathcal{P}^k(1-\mathcal{P})^{n-k} \tag{13}$$

($n=N^2$ is the number of nodes in $S$; $C^k_n$ are binomial coefficients).

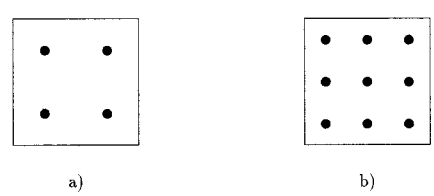



**Fig. 2 Arrangement of nodes in the representative area; (*a*) *N*=2, (*b*) *N*=3**

Let $\langle C\rangle$ be one of the moduli $\langle \overset{(0)}{C}_{mnij}\rangle$, $\langle \overset{(1)}{C}_{mnijkl}\rangle$, obtained by all possible arrangements of pores in $S$, and $\langle C\rangle|_k$ be the correspondent modulus, obtained by averaging by the arrangements, for the region $S$ which contains exactly $k$ pores. Then

$$\langle C\rangle=\sum_{k=0}^{n}\langle C\rangle|_k\mathcal{P}^{(k)}. \tag{14}$$

Note that when the probability $\mathcal{P}$ is small enough (but not necessarily infinitesimal), one can compute the moduli $\langle C\rangle|_k$ at $k>n_0$ ($n_0<n$ is a given integer) neglecting the interaction between pores in $S$ without significant loss of accuracy.

After the computation of moduli $\langle \overset{(0)}{C}_{mnij}\rangle$, $\langle \overset{(1)}{C}_{mnijkl}\rangle$, the averaging by all possible orientations of coordinate axes in the plane of deformation is performed, following Shermergor [20] and Levin et al. [14]. This averaging permits one to represent the effective constitutive equations in a form proposed by Murnaghan [21] (see also Lurie, [16]).

$$\mathbf{\Sigma}^0=\lambda^e(\mathbf{E}^e:\mathbf{I})\mathbf{I}+2G^e\mathbf{E}^e+3\mathcal{C}_3^e(\mathbf{E}^e:\mathbf{I})^2\mathbf{I}+\mathcal{C}_4^e[(\mathbf{E}^e)^2:\mathbf{I}]\mathbf{I}$$
$$+2\mathcal{C}_6^e(\mathbf{E}^e:\mathbf{I})\mathbf{E}^e+3\mathcal{C}_5^e(\mathbf{E}^e)^2 \tag{15}$$

It was noted by Levin et al. [14] that one cannot uniquely determine all four moduli $\mathcal{C}_3^e$, $\mathcal{C}_4^e$, $\mathcal{C}_5^e$, $\mathcal{C}_6^e$ in (16) restricting oneself to the plane-strain problems. In that case one of these moduli may be given arbitrarily. After that, another three moduli are determined uniquely, and the value of the arbitrarily prescribed modulus has no effect on the behavior of the comparison material under plane strain.

Note that, as follows from (11), use of the scheme considered is limited by the maximal value of porosity $p_{\max}=\pi R^2/h^2$.

## 4 Computations and Numerical Results

In the present paper the computations are performed for the case when the mechanical properties of the matrix are described by Murnaghan potential ([21]) and the constitutive equations have the form

$$\mathbf{\Sigma}^0=\lambda(\mathbf{E}:\mathbf{I})\mathbf{I}+2G\mathbf{E}+3\mathcal{C}_3(\mathbf{E}:\mathbf{I})^2\mathbf{I}+\mathcal{C}_4(\mathbf{E}^2:\mathbf{I})\mathbf{I}+2\mathcal{C}_4(\mathbf{E}:\mathbf{I})\mathbf{E}$$
$$+3\mathcal{C}_5\mathbf{E}^2. \tag{16}$$

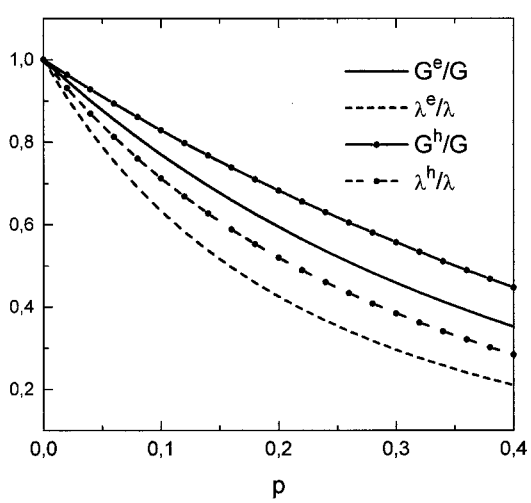



**Fig. 3 Effective linear elastic moduli $\lambda^e$, $G^e$ and upper Hashin-Shtrikman bounds $\lambda^h$, $G^h$ referred to the correspondent matrix moduli versus porosity**

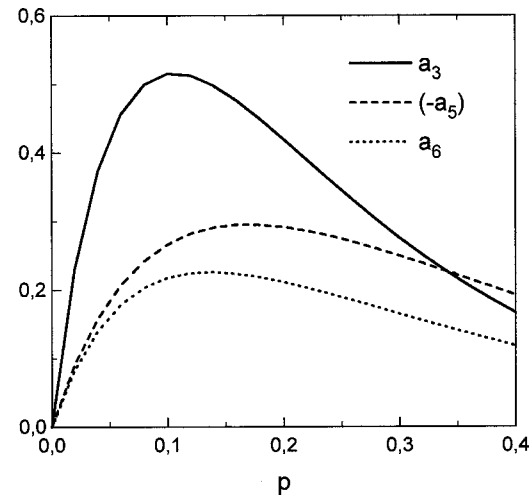



**Fig. 4 Coefficients $a_3$, $a_5$, $a_6$ versus porosity**

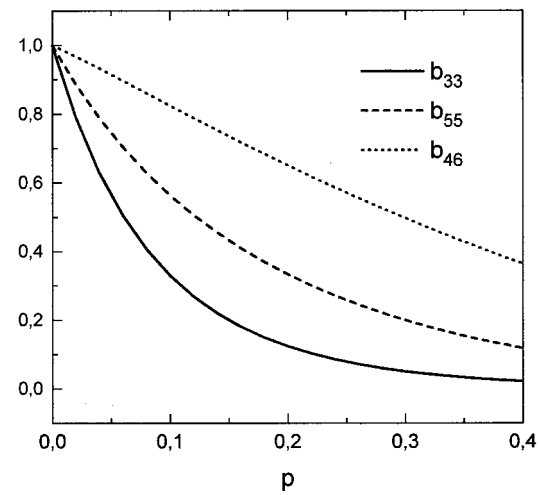



**Fig. 5 Coefficients $b_{33}$, $b_{55}$, $b_{46}$ versus porosity**

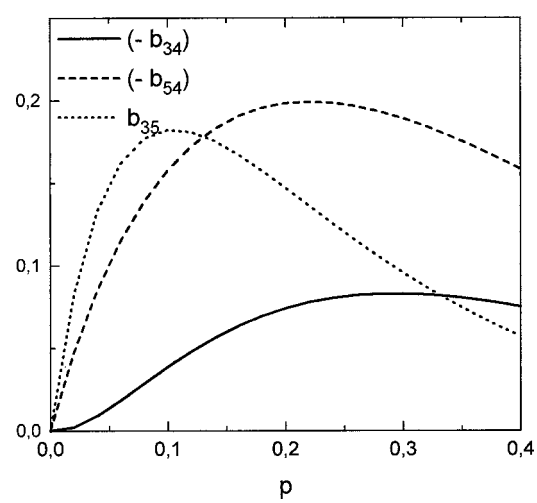



**Fig. 6 Coefficients $b_{34}$, $b_{54}$, $b_{35}$ versus porosity**

Because the solution of the problem of nonlinear elasticity, obtained by the perturbation technique up to the second order, depends linearly on the material constants $\mathcal{C}_3$, $\mathcal{C}_4$, $\mathcal{C}_5$, the effective moduli $\mathcal{C}_i^e$, computed by the scheme considered, also depend linearly on these constants:

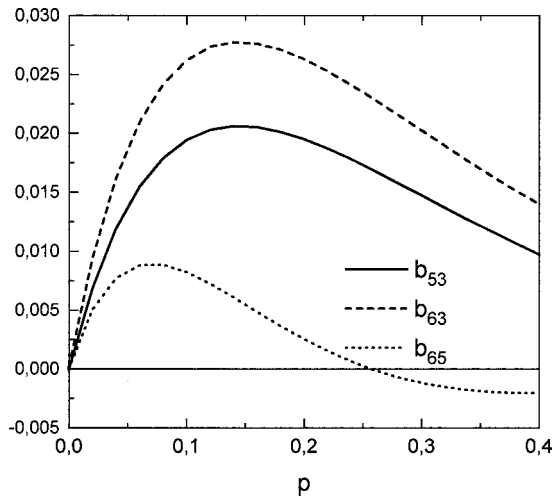



**Fig. 7 Coefficients $b_{53}$, $b_{63}$, $b_{65}$ versus porosity**

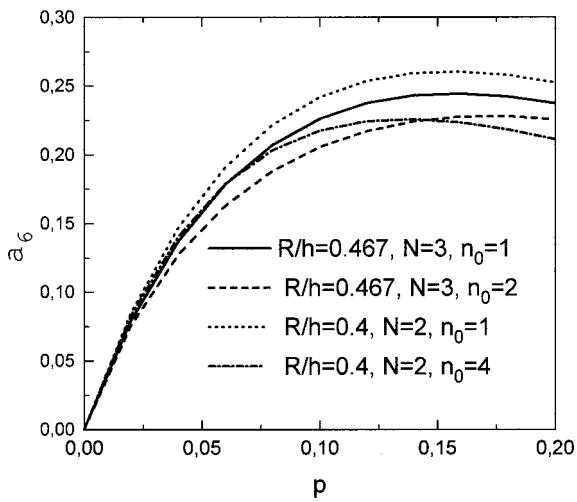



**Fig. 8 Effect of the computational parameters on the plot of the coefficient $a_6$ versus porosity**

$$\mathcal{C}_i^e = a_i + \sum_{j=3}^{5} b_{ij}\mathcal{C}_j, \quad (i=3,5,6). \tag{17}$$

It is assumed that the effective modulus $\mathcal{C}_4^e$ is arbitrarily prescribed and

$$\mathcal{C}_4^e = \mathcal{C}_4(1-p)^2. \tag{18}$$

The plots of the effective moduli versus porosity are presented below for the matrix materials with $\lambda/G=2.24$ at $R/h=0.4$, $N=2$, $n_0=N^2=4$. The ratios $\lambda^e/\lambda$, $G^e/G$ are plotted in Fig. 3, the coefficients $a_i$ are plotted in Fig. 4, and the coefficients $b_{ij}$ are plotted in Figs. 5–7. If the matrix constants $\mathcal{C}_3$, $\mathcal{C}_4$, $\mathcal{C}_5$ are known, one can obtain the effective moduli $\mathcal{C}_3^e$, $\mathcal{C}_5^e$, $\mathcal{C}_6^e$ from Eq. (17) with $\mathcal{C}_4^e$ given by (18).

The computations performed show also that the correction for the effects of interaction between pores within the representative area do not exceed three percent for the moduli $\lambda^e$ and $G^e$ and 20 percent for the coefficients $a_i$, $b_{ij}$ for $p \leqslant 0.4$. Note that Mauge and Kachanov [11] discovered the same effect analyzing the effective properties of linear elastic solids with arbitrarily located microcracks. The plots of the coefficient $a_6$ versus porosity, computed at the various values of $R/h$, $N$, and $n_0$, are presented in Fig. 8, as a reference (at $n_0=1$, the interaction between pores located within the representative region is neglected).

## References


[1] Hashin, Z., and Shtrikman, S., 1963, ''A Variational Approach to the Theory of the Elastic Behavior of Multiphase Materials,'' J. Mech. Phys. Solids, **11**, pp. 127–140.

[2] Willis, J. R., 1983, ''The Overall Response of Composite Materials,'' ASME J. Appl. Mech., **50**, pp. 1202–1209.

[3] Willis, J. R., 1991, ''On Methods for Bounding the Overall Properties of Nonlinear Composites,'' J. Mech. Phys. Solids, **39**, No. 1, pp. 73–86.

[4] Ponte Castaneda, P., 1991, ''The Effective Mechanical Properties of Nonlinear Isotropic Solids,'' J. Mech. Phys. Solids, **39**, No. 1, pp. 45–71.

[5] Ponte Castaneda, P., and Zaidman, M., 1994, ''Constitutive Models for Porous Materials With Evolving Microstructure,'' J. Mech. Phys. Solids, **42**, No. 9, pp. 1459–1497.

[6] Horvay, G., 1952, ''The Plane-Stress Problem of Perforated Plates,'' ASME J. Appl. Mech., **19**, pp. 355–360.

[7] Filshtinskiy, L. A., 1964, ''Stresses and Strains in Elastic Plane Weakened by Doubly Periodic System of Equal Circular Holes,'' Prikl. Mat. Mekh., **28**, pp. 430–441 (in Russian).

[8] Abeyaratne, R., and Triantafyllidis, N., 1984, ''An Investigation of Localization in Porous Elastic Material Using Homogenization Theory,'' ASME J. Appl. Mech., **51**, pp. 481–486.

[9] Kachanov, M., 1992, ''Effective Elastic Properties of Cracked Solids: Critical Review of Some Basic Concepts,'' Appl. Mech. Rev., **45**, pp. 304–335.

[10] Kachanov, M., Tsukrov, I., and Shafiro, B., 1994, ''Effective Moduli of a Solid With Holes and Cavities of Various Shapes,'' Appl. Mech. Rev., **47**, No. 1, Part 2, pp. S151–S174.

[11] Mauge, C., and Kachanov, M., 1994, ''Effective Elastic Properties of an Anisotropic Material With Arbitrary Oriented Interacting Cracks,'' J. Mech. Phys. Solids, **42**, No. 4, pp. 561–584.

[12] Shafiro, B., and Kachanov, M., 1997, ''Materials With Fluid-Filled Pores of Various Shapes: Effective Elastic Properties and Fluid Pressure Polarization,'' Int. J. Solids Struct., **34**, No. 27, pp. 3517–3540.

[13] Levin, V. A., Lokhin, V. V., and Zingerman, K. M., 1996, ''Method of Estimation of Effective Properties of Porous Bodies Undergoing Finite Deformation,'' Int. J. Fract., **81**, pp. R79–R82.

[14] Levin, V. A., Lokhin, V. V., and Zingerman, K. M., 1997, ''A Method of Estimation of Effective Characteristics of Porous Bodies Under Finite Deformations,'' Mekh. Tver. Tela, No. 4, pp. 45–50 (in Russian).

[15] Green, A. E., and Adkins, J. E., 1960, *Large Elastic Deformations and Non-Linear Continuum Mechanics*, Clarendon Press, Oxford, UK.

[16] Lurie, A. I., 1990, *Nonlinear Theory of Elasticity*, North-Holland, Amsterdam.

[17] Levin, V. A., 1998, ''Theory of Repeated Superposition of Large Deformations: Elastic and Viscoelastic Bodies,'' Int. J. Solids Struct., **35**, No. 20, pp. 2585–2600.

[18] Levin, V. A., and Zingerman, K. M., 1998, ''Interaction and Microfracturing Pattern for Succesive Origination (Introduction) of Pores in Elastic Bodies: Finite Deformation,'' ASME J. Appl. Mech., **65**, No. 2, pp. 431–435.

[19] Mori, T., and Tanaka, K., 1973, ''Average Stress in Matrix and Average Elastic Energy of Materials With Misfitting Inclusions,'' Acta Metall., **21**, pp. 571–574.

[20] Shermergor, T. D., 1977, *Theory of Elasticity of Microinhomogeneous Media*, Nauka, Moscow (in Russian).

[21] Murnaghan, F. D., 1951, *Finite Deformation of an Elastic Solids*, John Wiley and Sons, New York.

# Development of One-Dimensional Models for Elastic Waves in Heterogeneous Beams

**H. Murakami**[1]
Professor of Applied Mechanics,
Department of Mechanical
and Aerospace Engineering,
University of California,
9500 Gilman Drive,
La Jolla, CA 92093-0411
ASME Member

**J. Yamakawa**
Assistant Professor,
Department of Mechanical Engineering,
The National Defense Academy
Yokosuka 239-8686, Japan
Assoc. Mem. ASME

*In order to develop an accurate one-dimensional model for wave propagation in heterogeneous beams with uniform cross sections, a Hamilton-type principle is developed by incorporating Reissner's semi-complimentary energy function. Trial displacement and transverse stress fields are constructed from the solutions of micro-boundary value problems (MBVP's) defined over the cross section. The MBVP's are developed from asymptotic expansions that assume a small diameter cross section compared to the axial length and a typical signal wavelength. Saint Venant's semi-inverse torsion and flexure problems are included in the system of MBVP's. By utilizing the displacement and transverse stress fields constructed from the numerical solutions of the MBVP's, the constitutive relations are developed. The model generalizes the Mindlin-Hermann rod model and the Timoshenko beam model for anisotropic heterogeneous beams. The accuracy of the model is assessed by comparing the predicted phase velocity spectra to those computed by using a semi-analytical finite element method. Numerical results are shown for reinforced concrete beams with exterior composite layers.* [S0021-8936(01)00601-8]

## Introduction

The existing composite beam models were constructed based upon elementary kinematics, a Taylor series expansion, and the minimum potential energy theorem with appropriate trial displacements. Discussion of these models may be found in a review article by Kapania and Raciti [1] for laminated composite beams and in papers by Bank and Kao [2] and Song and Librescu [3] for thin and thick-walled beams.

Kosmatka and Dong [4] developed an anisotropic beam model for quasi-static deformation of prismatic composite beams by utilizing the displacement field obtained from Saint Venant's semi-inverse analyses. Pecastaings, Sun, and Davenne [5] also employed Saint Venant's warping functions to improve the accuracy of elementary solutions for heterogeneous beams with isotropic constituents. Their work shows the advantage of constructing beam models from the solutions of the semi-inverse analyses.

For homogeneous isotropic prismatic bars with uniform cross section, the Timoshenko beam theory ([6]) and the Mindlin-Herrmann rod theory ([7]) were constructed to mimic the wave propagation in an isotropic cylinder with uniform circular cross section. These problems are known as the Pochhammer problem ([8–10]). In order to examine the accuracy of the elementary theory as applied to heterogeneous beams with isotropic constituents, phase velocity spectra were computed by the authors employing a semi-analytical finite element method ([11]). Strong wave dispersion was observed for both longitudinal and flexural wave propagation, which is consistent with Pochhammer's solution for the problem of isotropic circular cylinders ([12,13]). Therefore, in order to capture wave dispersion, it is necessary to construct high-order theories for heterogeneous rods, beams, and shafts; hereafter they are all referred to as heterogeneous beams.

To date, a high-order one-dimensional theory has not been reported for dynamic response of heterogeneous beams. The work presented was performed in response to a need for an economical beam model that captures stress and strain concentrations around reinforcing bars under initial seismic loading of reinforced concrete columns. Elastic stress concentrations cause the onset of major crack formation emanating from the bars.

The objective of this paper is to develop a one-dimensional theory for heterogeneous beams with uniform cross section. In what follows, Reissner's semi-complementary energy function, originally developed for plates ([14]), was modified for heterogeneous beams. The resulting semi-complementary energy function for the beam was then incorporated into a Hamilton-type principle to establish a dynamic extension of the mixed variational equation based upon displacements and transverse stresses ([15]). (In beam problems, the transverse stresses are stress components excluding axial normal stress.) In order to furnish trial displacement and transverse stress fields, asymptotic expansion, assuming a small diameter cross section in relation to the axial length and a typical signal wavelength, was employed to define micro-boundary value problems (MBVP's) over the cross section. Saint Venant's semi-inverse torsion and flexure problems for heterogeneous cross sections are included in the systems of MBVP's. Numerical examples are shown for composite-retrofitted reinforced concrete columns with circular and square cross sections. The model accuracy is assessed by comparing the predicted phase velocity spectra to those obtained from the three-dimensional theory of elasticity.

## Formulation of the Problem

Figure 1 illustrates the circular and square cross sections of the retrofitted reinforced concrete columns with exterior composite layers. Let the $x_3$-axis be in the direction of the beam axis, and the $x_1,x_2$-plane be the plane of the cross section. The $x_1$, $x_2$, and $x_3$-axes form a rectangular Cartesian coordinate system. The axial region of the beam with length $L$ is $0 \leqslant x_3 \leqslant L$. The cross section $\Omega$ consists of the nonoverlapping subregions of its constituents. In the examples shown in Fig. 1, $\Omega$ consists of concrete $\Omega^{(2)}$, the exterior composite layer $\Omega^{(3)}$, and the longitudinal steel bars, $\Omega_m^{(1)}$, $m=1,2,\ldots,N$ where $N$ denotes the number of longitudinal bars:

$$\Omega=\Omega^{(1)}\cup\Omega^{(2)}\cup\Omega^{(3)}, \quad \Omega^{(1)}=\bigcup_{m=1}^{N}\Omega_m^{(1)}. \tag{1}$$

[1]To whom correspondence should be addressed.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, November 6, 1998; final revision, June 6, 2000. Associate Technical Editor: A. K. Mal. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

 Copyright © 2000 by ASME 

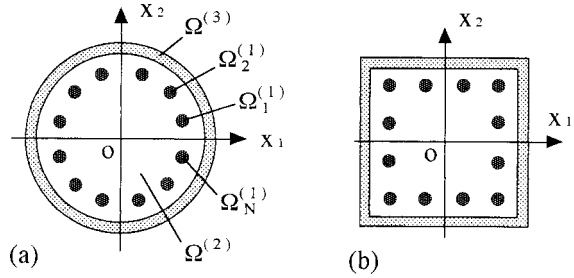



**Fig. 1 Heterogeneous beams with (*a*) circular and (*b*) square cross sections**

For notational simplicity, the superscripts $(\alpha)$ denote variables associated with constituents $\alpha$: material $\alpha=1$ (steel), 2 (concrete), and 3 (composite). The boundary of each subregion is denoted by $\partial\Omega^{(\alpha)}$ whose unit outward normal is $n_i^{(\alpha)}$.

Cartesian indicial notation is employed along with the summation convention for repeated indices. The governing relations for the displacement vector $u_i$ and the symmetric stress and strain tensor fields, $\sigma_{ij}$ and $e_{ij}$, in each material region are as follows:
(a) Equations of motion:

$$\sigma_{ji,j}^{(\alpha)}+\rho^{(\alpha)}b_i^{(\alpha)}=\rho^{(\alpha)}\ddot{u}_i^{(\alpha)} \quad \text{in } \Omega^{(\alpha)}, \tag{2}$$

where $(\ )_{,j}\equiv\partial(\ )/\partial x_j$, $\rho$ is the mass density, $b_j$ is the constant body force, and superposed dots mean differentiation with time $t$.
(b) Hooke's law:

$$\sigma_{ij}^{(\alpha)}=C_{ijmn}^{(\alpha)}e_{mn}^{(\alpha)} \quad \text{in } \Omega^{(\alpha)}, \tag{3a}$$

$$\text{or} \quad e_{mn}^{(\alpha)}=D_{mnij}^{(\alpha)}\sigma_{ij}^{(\alpha)} \quad \text{in } \Omega^{(\alpha)}, \tag{3b}$$

where $C_{ijmn}$ and $D_{ijmn}(=C_{ijmn}^{-1})$ are, respectively, the elastic modulus and compliance tensors. The exterior composite layer (material 3) is characterized by a cylindrically orthotropic material that becomes monoclinic with respect to the $x_i$-coordinate system ([16]). In compact notation, the strain-stress relation of monoclinic materials is given as

$$\begin{Bmatrix} e_{11} \\ e_{22} \\ e_{33} \\ 2e_{23} \\ 2e_{31} \\ 2e_{12} \end{Bmatrix}^{(3)} = \begin{bmatrix} D_{11} & D_{12} & D_{13} & 0 & 0 & D_{16} \\ D_{12} & D_{22} & D_{23} & 0 & 0 & D_{26} \\ D_{13} & D_{23} & D_{33} & 0 & 0 & D_{36} \\ 0 & 0 & 0 & D_{44} & D_{45} & 0 \\ 0 & 0 & 0 & D_{45} & D_{55} & 0 \\ D_{16} & D_{26} & D_{36} & 0 & 0 & D_{66} \end{bmatrix}^{(3)} \times \begin{Bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{23} \\ \sigma_{31} \\ \sigma_{12} \end{Bmatrix}^{(3)} \quad \text{in } \Omega^{(3)}, \tag{3c}$$

where indices 1–6 for $D_{ij}(=D_{ji})$ are interpreted as

$$[1\ \ 2\ \ 3\ \ 4\ \ 5\ \ 6]\equiv[11\ \ 22\ \ 33\ \ 23\ \ 31\ \ 12].$$

For isotropic material 1 (steel) and material 2 (concrete), the nonzero components of the compliance tensor can be written in terms of Young's modulus $E$ and Poisson's ratio $\nu$.

In matrix form, Eq. (3*c*) becomes

$$\{e\}^{(\alpha)}=[D]^{(\alpha)}\{\sigma\}^{(\alpha)}=[C]^{(\alpha)^{-1}}\{\sigma\}^{(\alpha)} \quad \alpha=1,2,3. \tag{3d}$$

(c) Strain-displacement relations:

$$e_{mn}^{(\alpha)}=\frac{1}{2}(u_{m,n}^{(\alpha)}+u_{n,m}^{(\alpha)}) \quad \text{in } \Omega^{(\alpha)}. \tag{4}$$

(d) Interfacial boundary conditions:

$$u_i^{(1)}=u_i^{(2)}, \quad \sigma_{ji}^{(1)}n_j^{(1)}=\sigma_{ji}^{(2)}n_j^{(1)} \quad \text{on } \partial\Omega_m^{(1)}, \quad m=1,2,\ldots,N, \tag{5}$$

$$u_i^{(2)}=u_i^{(3)}, \quad \sigma_{ji}^{(2)}n_j^{(2)}=\sigma_{ji}^{(3)}n_j^{(2)} \quad \text{on } \partial\Omega_o^{(2)}, \tag{6}$$

which represent the continuity of displacements and tractions on each interface between the steel bars and concrete $\partial\Omega_m^{(1)}$ and on the interface between the concrete and composite layer $\partial\Omega_o^{(2)}$.
(e) Lateral traction boundary conditions:

$$\sigma_{ji}^{(3)}n_j^{(3)}=p_i \quad \text{on } \partial\Omega_o, \tag{7}$$

where $p_i$ is a prescribed function of time and $\partial\Omega_o(\equiv\partial\Omega_o^{(3)})$ is the exterior boundary of the cross section.
(f) End conditions at $x_3=0$ and $L$:

$$\begin{aligned} &\text{either } \sigma_{3i}^{(\alpha)}=p_i^{0(\alpha)}(x_1,x_2,t) \text{ or } u_i^{(\alpha)} \text{ prescribed at } x_3=0, \\ &\text{either } \sigma_{3i}^{(\alpha)}=p_i^{L(\alpha)}(x_1,x_2,t) \text{ or } u_i^{(\alpha)} \text{ prescribed at } x_3=L, \end{aligned} \tag{8}$$

where $p_i^0$ and $p_i^L$ $(i=1,2,3)$ are prescribed functions.
(g) Initial conditions:

$$u_i^{(\alpha)} \text{ and } \dot{u}_i^{(\alpha)} \text{ prescribed at } t=0. \tag{9}$$

The objective of the subsequent analysis is to derive a set of one-dimensional equations whose solutions can be utilized to approximate the solution of the problem defined by (a)–(g). To this end, a dynamic extension of Reissner's mixed variational principle for displacements and transverse stresses will be developed for heterogeneous beams. In beam problems, transverse stresses are stress components excluding axial normal stress, i.e., $\sigma_{ij}$, $(i,j)\neq(3,3)$. Transverse stress components appear in the traction conditions on the cross section, (5)–(7). In what follows, the transverse stresses will be denoted by a superscript $t$ as $\sigma_{ij}^t$.

## Reissner's Semi-Complementary Energy Function

In heterogeneous beams, axial normal stress experiences discontinuity at material interfaces. In the mixed variational formulation by Reissner [17], trial stress fields are required even when they are discontinuous. However, it is extremely difficult to construct trial fields for the discontinuous stresses encountered in heterogeneous beams and plates. To improve this situation, Reissner [14] proposed a semi-complementary energy function for plate problems. Reissner's semi-complementary energy function offers an attractive mixed variational formulation in which trial stresses are required only for transverse stresses but not for discontinuous axial normal stress. In this section, a semi-complementary energy function is derived for beam problems in which all stress components except axial normal stress become transverse stresses $\sigma_{ij}^t$. For notational simplicity, superscripts for material $\alpha$ (=1,2,3) will not be shown in this section. In order to incorporate the constitutive relations of transverse stresses, Reissner augmented the strain-energy density function $U(e_{ij})$ with the transverse strain-displacement relations:

$$\Pi_R\equiv U(e_{ij})+\sum_{(i,j)\neq(3,3)}\sigma_{ij}^t\left\{\frac{1}{2}(u_{i,j}+u_{j,i})-e_{ij}^t\right\}, \tag{10}$$

where $e_{ij}^t$ denotes transverse strains and

$$\sigma_{ij}=\frac{\partial U}{\partial e_{ij}}. \tag{11}$$

The strain energy density function is decomposed as

$$U(e_{ij})=U^a(e_{33})+U^t(e_{ij}^t,e_{33}). \tag{12}$$

Substituting (12) into (10), we obtain

$$\Pi_R = U^a(e_{33}) + \sum_{(i,j)\neq(3,3)} \frac{1}{2}(u_{i,j}+u_{j,i})\sigma^t_{ij} - \tilde{U}^t(\sigma^t_{ij}, e_{33}), \tag{13a}$$

where a partial Legendre transformation

$$\tilde{U}^t(\sigma^t_{ij}, e_{33}) \equiv \sum_{(i,j)\neq(3,3)} e^t_{ij}\sigma^t_{ij} - U^t(e^t_{ij}, e_{33}) \tag{13b}$$

was performed to switch independent variables from transverse strains $e^t_{ij}$ to transverse stresses $\sigma^t_{ij}$. In Reissner's mixed variational principle, the strain-displacement relation $e_{33} \equiv u_{3,3}$ is treated as a constraint condition. Since the above transformation is complicated for beam problems, the functional $\Pi_R$ is modified by using the complementary energy density function $W_c$ that is obtained from the full Legendre transformation

$$W_c(\sigma_{ij}) \equiv e_{ij}\sigma_{ij} - U(e_{ij}), \tag{14}$$

where

$$e_{ij} = \frac{\partial W_c}{\partial \sigma_{ij}}. \tag{15}$$

Rewriting (10) with (14), one obtains

$$\Pi_R = e_{33}\sigma_{33} + \sum_{(i,j)\neq(3,3)} \frac{1}{2}(u_{i,j}+u_{j,i})\sigma^t_{ij} - W_c(\sigma^t_{ij}, \sigma_{33}). \tag{16}$$

Reissner's semi-complementary energy function is now defined as

$$W_R(\sigma^t_{ij}, e_{33}) \equiv W_c(\sigma^t_{ij}, \sigma_{33}) - e_{33}\sigma_{33}, \tag{17}$$

where a partial Legendre transformation was performed to switch independent variables from $\sigma_{31}$ to $e_{33}$. From (15) the function furnishes the constitutive relations in semi-inverted form:

$$e^t_{ij} = \frac{\partial W_R}{\partial \sigma^t_{ij}}, \tag{18a}$$

$$\sigma_{33} = -\frac{\partial W_R}{\partial e_{33}}. \tag{18b}$$

The augmented strain-energy density function with (16) and (17) becomes

$$\Pi_R = \sum_{(i,j)\neq(3,3)} \frac{1}{2}(u_{i,j}+u_{j,i})\sigma^t_{ij} - W_R(\sigma^t_{ij}, u_{3,3}), \tag{19}$$

where the constraint condition $e_{33} \equiv u_{3,3}$ was used. In what follows, Eq. (17) is applied to monoclinic materials defined by (3*c*).

The partial Legendre transformation in (17) can be easily performed by using the constitutive equation for $e_{33}$ ($\equiv u_{3,3}$) in (3*c*) to solve for $\sigma_{33}$. This is given by

$$\sigma_{33}(\sigma^t_{ij}, e_{33}) = \frac{e_{33}}{D_{33}} - [D''_{13}\ \ D''_{23}\ \ 0\ \ 0\ \ D''_{36}]\{\sigma_t\}, \tag{20a}$$

where

$$D''_{i3} \equiv \frac{D_{i3}}{D_{33}}, \quad i,j = 1,2,6, \tag{20b}$$

$$\{\sigma_t\}^T \equiv [\sigma_{11}\ \ \sigma_{22}\ \ \sigma_{23}\ \ \sigma_{31}\ \ \sigma_{12}], \tag{20c}$$

and $\{\}^T$ denotes the transpose of $\{\}$.

In matrix form $\delta W_R$ is expressed with (18) as

$$\delta W_R = \lfloor \delta\sigma_{11}\ \ \delta\sigma_{22}\ \ -\delta e_{33}\ \ \delta\sigma_{23}\ \ \delta\sigma_{31}\ \ \delta\sigma_{12} \rfloor [D']\{\sigma_R\} \tag{21a}$$

where

$$\{\sigma_R\}^T \equiv [\sigma_{11}\ \ \sigma_{22}\ \ e_{33}\ \ \sigma_{23}\ \ \sigma_{31}\ \ \sigma_{12}], \tag{21b}$$

and $[D']$ is the semi-inverted matrix directly obtained from (3*c*) by using (20*a*,*b*):

$$\{e_R\} = [D']\{\sigma_R\}. \tag{22a}$$

In expanded form

$$\begin{Bmatrix} e_{11} \\ e_{22} \\ \sigma_{33} \\ 2e_{23} \\ 2e_{31} \\ 2e_{12} \end{Bmatrix} = \begin{bmatrix} D'_{11} & D'_{12} & D''_{13} & 0 & 0 & D'_{16} \\ D'_{12} & D'_{22} & D''_{23} & 0 & 0 & D'_{26} \\ -D''_{13} & -D''_{23} & 1/D_{33} & 0 & 0 & -D''_{36} \\ 0 & 0 & 0 & D_{44} & D_{45} & 0 \\ 0 & 0 & 0 & D_{45} & D_{55} & 0 \\ D'_{16} & D'_{26} & D''_{36} & 0 & 0 & D'_{66} \end{bmatrix} \times \begin{Bmatrix} \sigma_{11} \\ \sigma_{22} \\ e_{33} \\ \sigma_{23} \\ \sigma_{31} \\ \sigma_{12} \end{Bmatrix}, \tag{22b}$$

where $e_{33} \equiv u_{3,3}$ and

$$D'_{ij} \equiv D_{ij} - \frac{D_{i3}D_{3j}}{D_{33}}, \quad i,j = 1,2,6. \tag{22c}$$

In Hamilton's principle, the augmented strain-energy density function (19) will be used instead of the strain-energy density function (11).

## Hamilton-Type Principle and Derivation of Beam Equations

In order to construct a dynamic extension of the mixed variational principle for displacements and transverse stresses, Hamilton's principle is employed with the augmented strain-energy density function for two arbitrary times $t_1$ and $t_2$ ($>t_1$),

$$\delta \int_{t_1}^{t_2} \left\{ \int_0^L \left[ \iint_\Omega \left\{ \sum_{(i,j)\neq(3,3)} \frac{1}{2}(u_{i,j}+u_{j,i})\sigma^t_{ij} - W_R(\sigma^t_{ij}, u_{3,3}) - \rho b_i u_i - \frac{\rho}{2}\dot{u}_i\dot{u}_i \right\} dx_1\, dx_2 - \oint_{\partial\Omega_o} p_i u_i\, ds \right] dx_3 - \iint_\Omega [p^L_i(x_1,x_2,t)u_i(x_1,x_2,L,t) - p^0_i(x_1,x_2,t)u_i(x_1,x_2,0,t)]\, dx_1\, dx_2 \right\} dt = 0, \tag{23a}$$

where all variations are assumed to vanish at $t_1$ and $t_2$,

$$\delta u_i = \delta\sigma^t_{ij} = 0 \quad \text{at } t = t_1 \text{ and } t_2, \tag{23b}$$

and if displacements are prescribed at $x_3 = 0$ and L as described in (8), the corresponding virtual displacements vanish.

Recasting Eq. (23*a*) in the form of the principle of virtual work gives

$$\int_0^L \Bigg\{ \iint_\Omega \Bigg[ \sum_{(i,j)\neq(3,3)} \frac{1}{2}\sigma_{ij}^t(\delta u_{i,j}+\delta u_{j,i}) - \frac{\partial W_R}{\partial u_{3,3}}\delta u_{3,3} - \rho(b_i-\ddot{u}_i)\delta u_i + \sum_{(i,j)\neq(3,3)} \left\{ \frac{1}{2}(u_{i,j}+u_{j,i}) - \frac{\partial W_R}{\partial \sigma_{ij}^t} \right\} \delta\sigma_{ij}^t \Bigg] dx_1\, dx_2 - \oint_{\partial\Omega_o} p_i \delta u_i\, ds \Bigg\} dx_3 - \iint_\Omega \{ p_i^L(x_1,x_2,t)\,\delta u_i(x_1,x_2,L,t) - p_i^0(x_1,x_2,t)\,\delta u_i(x_1,x_2,0,t)\} dx_1\, dx_2 = 0. \tag{24}$$

It can be shown that for admissible displacements which satisfy the displacement end conditions (8) the Euler-Lagrange equations of (24) consist, as they should, of the system (2) with (18*b*) or (20*a*), (18*a*) with (4), and the traction boundary conditions (7) and (8).

In order to derive a system of one-dimensional equations from (24), one must obtain approximate displacement and transverse stress fields. In the next section, a scheme is presented for constructing both trial displacement and transverse stress fields based upon asymptotic expansion. The asymptotic analysis furnishes a displacement trial field of the form

$$u_1(x_1,x_2,x_3,t) \approx U_1(x_3,t) - \Psi_3(x_3,t)x_2 + \Phi(x_3,t)\hat{u}_{1(1)}(x_1,x_2) + \cdots \tag{25a}$$

$$u_2(x_1,x_2,x_3,t) \approx U_2(x_3,t) + \Psi_3(x_3,t)x_1 + \Phi(x_3,t)\hat{u}_{2(1)}(x_1,x_2) + \cdots \tag{25b}$$

$$u_3(x_1,x_2,x_3,t) \approx U_3(x_3,t) - \Psi_2(x_3,t)x_1 + \Psi_1(x_3,t)x_2 + \Delta_3(x_3,t)\hat{u}_{3(2)}(x_1,x_2) + \cdots \tag{25c}$$

where $U_i$ denotes the average displacement over the cross section. The $\Psi_i$-terms represent the rotational degrees-of-freedom of the cross section. The $U_i$ and $\Psi_i$-terms are employed for the kinematic description of elementary theories. The remaining terms with $\Phi$ and $\Delta_3$ describe the in-plane displacement field due to the axial strain $U_{3,3}$ and longitudinal warping, respectively. The $\Phi$-displacement mode $\hat{u}_{i(1)}$ will be computed from the solution of a generalized plane-strain problem over the cross section, while the $\Delta_3$-displacement mode $\hat{u}_{3(2)}$ describes the longitudinal warping obtained from an antiplane problem. The trial transverse stress field will be presented after introducing the micro-boundary value problems over the cross section.

In (25) the higher-order terms are constructed to be orthogonal to lower-order terms with respect an integration over the cross section. Among possible schemes, the orthogonality with respect to the integration weighted by mass density $\rho$ simplifies the kinetic energy expression. To this end, the $x_1$ and $x_2$-axes are selected to be centroidal axes with respect to the mass-weighted integration over the cross section:

$$\iint_\Omega \rho x_1\, dx_1\, dx_2 = \iint_\Omega \rho x_2\, dx_1\, dx_2 = 0. \tag{26}$$

It is assumed that cross sections are symmetric with respect to the $x_1$ and $x_2$-axes. (In order to treat nonsymmetric cross sections, one must carefully define the centers of shear and twist for nonsymmetric cross sections.)

By substituting a trial displacement field into (24), one-dimensional equations of any order can be obtained. From the displacement variables shown in (25), the resulting variational equation becomes

$$\int_0^L \{(\delta U_{1,3}-\delta\Psi_2)N_1 + (\delta U_{2,3}+\delta\Psi_1)N_2 + \delta U_{3,3}N_3^I + \delta\Delta_{3,3}N_3^{II} + \delta\Phi H^I + \delta\Phi_{,3}H^{II} + \delta\Delta_3 H^{III} + \delta\Psi_{1,3}M_1 + \delta\Psi_{2,3}M_2 + \delta\Psi_{3,3}M_3 - \delta U_1 m(B_1-\ddot{U}_1) - \delta U_2 m(B_2-\ddot{U}_2) - \delta U_3 m(B_3-\ddot{U}_3) + \delta\Phi J_p\ddot{\Phi} + \delta\Delta_3 J_d\ddot{\Delta}_3 + \delta\Psi_1 J_1\ddot{\Psi}_1 + \delta\Psi_2 J_2\ddot{\Psi}_2 + \delta\Psi_3 J_3\ddot{\Psi}_3 - \delta U_1 P_1^I - \delta U_2 P_2^I - \delta U_3 P_3^I - \delta\Phi P_p^{II} - \delta\Delta_3 P_3^{III} - \delta\Psi_1 P_1^{II} - \delta\Psi_2 P_2^{II} - \delta\Psi_3 P_3^{II}\} dx_3 + BC \text{ terms} = 0, \tag{27}$$

where the BC terms include the virtual work due to the end forces at $x_3=0$ and $L$.

In (27) the stress variables are defined as

$$[N_1\ \ N_2\ \ N_3^I\ \ N_3^{II}] \equiv \iint_\Omega [\sigma_{31}\ \ \sigma_{32}\ \ \sigma_{33}\hat{u}_{3(2)}\sigma_{33}]dx_1\, dx_2, \tag{28a}$$

$$[M_1\ \ M_2\ \ M_3] \equiv \iint_\Omega [x_2\sigma_{33}\ \ -x_1\sigma_{33}\ \ (x_1\sigma_{32}-x_2\sigma_{31})] \times dx_1\, dx_2, \tag{28b}$$

$$H^I \equiv \iint_\Omega \{\hat{u}_{1(1),1}\sigma_{11} + \hat{u}_{2(1),2}\sigma_{22} + (\hat{u}_{1(1),2}+\hat{u}_{2(1),1})\sigma_{12}\} dx_1\, dx_2, \tag{28c}$$

$$[H^{II}\ \ H^{III}] \equiv \iint_\Omega [\hat{u}_{1(1)}\sigma_{31} + \hat{u}_{2(1)}\sigma_{32} \times \hat{u}_{3(2),1}\sigma_{31} + \hat{u}_{3(2),2}\sigma_{32}]dx_1\, dx_2, \tag{28d}$$

the mass variables are

$$[m\ \ J_1\ \ J_2\ \ J_3] \equiv \iint_\Omega [1\ \ x_2^2\ \ x_1^2\ \ x_1^2+x_2^2]\rho\, dx_1\, dx_2, \tag{29a}$$

$$[J_p\ \ J_d] \equiv \iint [\hat{u}_{1(1)}^2 + \hat{u}_{2(1)}^2\ \ \hat{u}_{3(2)}^2]\rho\, dx_1\, dx_2, \tag{29b}$$

and the force variables are defined as

$$B_i \equiv \frac{1}{m}\iint_\Omega \rho b_i\, dx_1\, dx_2, \quad P_i^I \equiv \oint_{\partial\Omega_o} p_i\, ds, \tag{30a}$$

$$[P_1^{II}\ \ P_2^{II}\ \ P_3^{II}] \equiv \oint_{\partial\Omega_o} [x_2p_3\ \ -x_1p_3\ \ x_1p_2 - x_2p_1]ds, \tag{30b}$$

$$[P_p^{II}\ \ P_3^{III}] \equiv \oint_{\partial\Omega_o} [\hat{u}_{1(1)}p_1 + \hat{u}_{2(1)}p_2\ \ \hat{u}_{3(2)}p_3]ds. \tag{30c}$$

The Euler-Lagrange equations of (27) yield for $0<x_3<L$,

$$N_{1,3}+P_1^I+mB_1-m\ddot{U}_1=0, \tag{31a}$$

$$N_{2,3}+P_2^I+mB_2-m\ddot{U}_2=0, \tag{31b}$$

$$N_{3,3,}^I+P_3^I+mB_3-m\ddot{U}_3=0, \tag{31c}$$

$$N_{3,3}^{II}-H^{III}+P_3^{III}-J_d\ddot{\Delta}_3=0, \tag{31d}$$

$$H_{,3}^{II}-H^I+P_p^{II}-J_p\ddot{\Phi}=0, \tag{31e}$$

$$M_{1,3}-N_2+P_1^{II}-J_1\ddot{\Psi}_1=0, \tag{31f}$$

$$M_{2,3}+N_1+P_2^{II}-J_2\ddot{\Psi}_2=0, \tag{31g}$$

$$M_{3,3}+P_3^{II}-J_3\ddot{\Psi}_3=0, \tag{31h}$$

with boundary conditions at $x_3=0$ and $L$:

$$U_i \text{ or } N_i \text{ prescribed for } i=1,2 \text{ at } x_3=0 \text{ and } L, \tag{32a}$$

$$U_3 \text{ or } N_3^I \text{ prescribed at } x_3=0 \text{ and } L, \tag{32b}$$

$$\Delta_3 \text{ or } N_3^{II} \text{ prescribed at } x_3=0 \text{ and } L, \tag{32c}$$

$$\Phi \text{ or } H^{II} \text{ prescribed at } x_3=0 \text{ and } L, \tag{32d}$$

$$\Psi_i \text{ or } M_i \text{ prescribed for } i=1,2,3 \text{ at } x_3=0 \text{ and } L. \tag{32e}$$

It is noted in (31) that the acceleration terms are uncoupled due to the use of mass-weighted orthogonalization of displacements over the cross section. The substitution of the trial transverse stress field into the remaining terms of (24):

$$\int\int_{\Omega} \sum_{(i,j)\neq(3,3)} \left\{\frac{1}{2}(u_{i,j}+u_{j,i})-\frac{\partial W_R}{\partial \sigma_{ij}^t}\right\}\delta\sigma_{ij}^t\, dx_1\, dx_2=0 \tag{33}$$

furnishes several one-dimensional constitutive relations. The system of equations, (31) and (32), furnishes a Timoshenko-type model for flexural motion and a Mindlin-Herrmann-type model for longitudinal motion. For flexural wave motion, a Rayleigh-type model can be obtained from (27) by eliminating the transverse shear strains with

$$\Psi_1=-U_{2,3}, \quad \Psi_2=U_{1,3}. \tag{34a}$$

Furthermore, by neglecting the rotary inertia effect, the Euler-Bernoulli equations are obtained. For longitudinal wave motion, a Love-Rayleigh-type model can be obtained from (27) by using

$$\Phi=U_{3,3}, \quad \Delta_3=0. \tag{34b}$$

In what follows, the trial displacement and transverse stress fields are constructed from asymptotic analyses.

## Asymptotic Analyses

To begin, both independent and dependent variables are nondimensionalized by using the following quantities:

| | |
|---|---|
| $L$ | the length of the beam or a typical signal wavelength |
| $d$ | the diameter of the cross section |
| $E^{(2)}$, $\rho^{(2)}$ | Young's modulus and mass density of material 2 |
| $C_0^{(2)}\equiv\sqrt{E^{(2)}/\rho^{(2)}}$ | the bar velocity of material 2 |
| $t_{(m)}\equiv L/C_0^{(2)}$ | a typical signal travel time |
| $\varepsilon\equiv d/L$ | the ratio of cross section to axial dimensions. |

Using the above notation, nondimensional variables identified by overbars are introduced according to

$$[\overline{x_1^*}\ \ \overline{x_2^*}\ \ \overline{x_3}]\equiv\left[\frac{x_1}{d}\ \frac{x_2}{d}\ \frac{x_3}{L}\right], \quad \overline{u_i}\equiv\frac{u_i}{d}, \quad \overline{t}\equiv\frac{t}{t_{(m)}},$$
$$[\overline{C}_{ijmn}\overline{\sigma}_{ij}]\equiv\frac{1}{E^{(2)}}[C_{ijmn}\ \ \sigma_{ij}], \quad \overline{\rho}\equiv\frac{\rho}{\rho^{(2)}}. \tag{35}$$

Field variables are now considered to be functions of $\overline{x_3}$, $\overline{t}$, and microcoordinates $\overline{x_1^*}$ and $\overline{x_2^*}$ [18,19]. In the beam model for elastic wave motion, the body force $b_i$ and lateral traction $p_i$ will be neglected in the construction of displacement modes over the cross section.

Using these nondimensional variables, Eqs. (2)–(4) become

$$\frac{1}{\varepsilon}(\overline{\sigma}_{1i,1*}^{(\alpha)}+\overline{\sigma}_{2i,2*}^{(\alpha)})+\overline{\sigma}_{3i,3}^{(\alpha)}=\overline{\rho}^{(\alpha)}\ddot{\overline{u}}_i^{(\alpha)}, \tag{36}$$

where $(\ )_{,i*}\equiv\partial(\ )/\partial\overline{x_i^*}$ and $(\dot{\ })\equiv\partial(\ )/\partial\overline{t}$.

The inverse of (3*e*) using (4) furnishes in compact notation

$$\{\overline{\sigma}\}^{(\alpha)}=[\overline{C}]^{(\alpha)}\{e(\overline{\mathbf{u}})\}^{(\alpha)}, \tag{37a}$$

where

$$\{e(\overline{\mathbf{u}})\}^T\equiv\frac{1}{\varepsilon}[\overline{u}_{1,1*}\ \ \overline{u}_{2,2*}\ \ \varepsilon\overline{u}_{3,3}\ \ \varepsilon\overline{u}_{2,3}+\overline{u}_{3,2*}\ \ \overline{u}_{3,1*}$$
$$+\varepsilon\overline{u}_{1,3}\ \ \overline{u}_{1,2*}+\overline{u}_{2,1*}]. \tag{37b}$$

The continuity and boundary conditions (5)–(8) are also expressed in nondimensional form. From the premise that the diameter of cross section $d$ is much smaller than the axial length $L$, $\varepsilon\ll 1$, one expands the dependent variables in an asymptotic series:

$$\overline{u}_i^{(\alpha)}(\overline{x_1^*},\overline{x_2^*},\overline{x_3},\overline{t})\approx\overline{u}_{i(0)}^{(\alpha)}(\overline{x_1^*},\overline{x_2^*},\overline{x_3},\overline{t})+\varepsilon\overline{u}_{i(1)}^{(\alpha)}(\overline{x_1^*},\overline{x_2^*},\overline{x_3},\overline{t})$$
$$+\varepsilon^2\overline{u}_{i(2)}^{(\alpha)}(\overline{x_1^*},\overline{x_2^*},\overline{x_3},\overline{t})+\cdots, \tag{38a}$$

$$\overline{\sigma}_{ij}^{(\alpha)}(\overline{x_1^*},\overline{x_2^*},\overline{x_3},\overline{t})\approx\overline{\sigma}_{ij(0)}^{(\alpha)}(\overline{x_1^*},\overline{x_2^*},\overline{x_3},\overline{t})+\varepsilon\overline{\sigma}_{ij(1)}^{(\alpha)}(\overline{x_1^*},\overline{x_2^*},\overline{x_3},\overline{t})$$
$$+\overline{\sigma}_{ij(2)}^{(\alpha)}(\overline{x_1^*},\overline{x_2^*},\overline{x_3},\overline{t})+\cdots. \tag{38b}$$

In (38*a*) higher-order displacement terms are constructed to be orthogonal to lower-order terms with respect to the integration weighted by mass density over the cross section. Substituting (38) into (36), (37), and the nondimensional form of (5)–(8), and setting each coefficient having the power of $\varepsilon$ to zero, one obtains a system of micro-boundary value problems (MBVP's) defined on the cross section $\overline{\Omega}$:

$$O(\varepsilon^{-1}): \qquad \overline{\sigma}_{1i(0),1*}^{(\alpha)}+\overline{\sigma}_{2i(0),2*}^{(\alpha)}=0, \tag{39}$$

$$[\overline{C}]^{(\alpha)}\{e_{(-1)}\}^{(\alpha)}=\{0\}, \tag{40a}$$

where $\{0\}$ is a $6\times 1$ zero column matrix and

$$\{e_{(-1)}\}^{(\alpha)T}\equiv[\overline{u}_{1(0),1*}\ \ \overline{u}_{2(0),2*}\ \ 0\ \ \overline{u}_{3(0),2*}\ \ \overline{u}_{3(0),1*}\ \ \overline{u}_{1(0),2*}$$
$$+\overline{u}_{2(0),1*}]^{(\alpha)}. \tag{40b}$$

$O(\varepsilon^k)$ for $k=0,1,2,\ldots$:

$$\overline{\sigma}_{1i(k+1),1*}^{(\alpha)}+\overline{\sigma}_{2i(k+1),2*}^{(\alpha)}+\overline{\sigma}_{3i(k),3}^{(\alpha)}=\overline{\rho}^{(\alpha)}\ddot{\overline{u}}_{i(k)}^{(\alpha)}, \tag{41}$$

$$\{\overline{\sigma}_{(k)}\}^{(\alpha)}=[\overline{C}]^{(\alpha)}\{e_{(k)}\}^{(\alpha)}, \tag{42a}$$

where

$$\{\overline{\sigma}_{(k)}\}^{(\alpha)T}\equiv[\overline{\sigma}_{11(k)}\ \ \overline{\sigma}_{22(k)}\ \ \overline{\sigma}_{33(k)}\ \ \overline{\sigma}_{23(k)}\ \ \overline{\sigma}_{31(k)}\ \ \overline{\sigma}_{12(k)}]^{(\alpha)}, \tag{42b}$$

$$\{e_{(k)}\}^{(\alpha)T}\equiv[\overline{u}_{1(k+1),1*}\ \ \overline{u}_{2(k+1),2*}\ \ \overline{u}_{3(k),3}\ \ \overline{u}_{2(k),3}$$
$$+\overline{u}_{3(k+1),2*}\ \ \overline{u}_{3(k+1),1*}+\overline{u}_{1(k),3}\ \ \overline{u}_{1(k+1),2*}$$
$$+\overline{u}_{2(k+1),1*}]^{(\alpha)}. \tag{42c}$$

The continuity of displacement and traction vectors at material interfaces, (5) and (6), becomes

$$\bar{u}^{(1)}_{i(k)}=\bar{u}^{(2)}_{i(k)}, \quad \bar{\sigma}^{(1)}_{ji(k)}n^{(1)}_j=\bar{\sigma}^{(2)}_{ji(k)}n^{(1)}_j \quad \text{on } \partial\bar{\Omega}^{(1)}_m, \tag{43}$$

$$\bar{u}^{(2)}_{i(k)}=\bar{u}^{(3)}_{i(k)}, \quad \bar{\sigma}^{(2)}_{ji(k)}n^{(2)}_j=\bar{\sigma}^{(3)}_{ji(k)}n^{(2)}_j \quad \text{on } \partial\bar{\Omega}^{(2)}_o. \tag{44}$$

The traction-free condition (7) on the exterior surface $\partial\bar{\Omega}_o$ is

$$\bar{\sigma}^{(3)}_{ji(k)}n^{(3)}_j=0 \quad \text{on } \partial\bar{\Omega}_o. \tag{45}$$

Due to the positive definiteness of $[\bar{C}]$, Eq. (40*a*) reduces to

$$\{e_{(-1)}\}^{(\alpha)}=\{0\}. \tag{46}$$

By imposing the boundary conditions (43)–(45) on (46) and assuming that the rigid-body motion in the cross section $\bar{\Omega}$ is $O(\varepsilon)$, one finds $\bar{u}^{(\alpha)}_{i(0)}$ independent of $\bar{x}^*_1$ and $\bar{x}^*_2$:

$$\bar{u}^{(\alpha)}_{i(0)}=\bar{u}_{i(0)}(\bar{x}_3,\bar{t}). \tag{47}$$

The rigid-body mode in the $\bar{x}^*_1$ and $\bar{x}^*_2$-plane describes torsional deformation:

$$\bar{u}^{(\alpha)}_{1(1)}=-\bar{\Psi}_3(\bar{x}_3,\bar{t})\bar{x}^*_2, \quad \bar{u}^{(\alpha)}_{2(1)}=\bar{\Psi}_3(\bar{x}_3,\bar{t})\bar{x}^*_1, \tag{48}$$

where $\bar{\Psi}_3$ represents the angle of twist.

The subsequent boundary value problems for $\bar{u}^{(\alpha)}_{i(k+1)}$, $k=0,1,2,\ldots$, defined over the cross section $\bar{\Omega}$ consist of: (i) in-plane problems and (ii) antiplane problems. The in-plane problems are defined as

$$\bar{\sigma}^{(\alpha)}_{1i(k),1*}+\bar{\sigma}^{(\alpha)}_{2i(k),2*}=\bar{f}^{(\alpha)}_{i(k)} \quad \text{for } i=1,2 \text{ in } \bar{\Omega}^{(\alpha)}, \tag{49}$$

$$\begin{Bmatrix}\bar{\sigma}_{11(k)}\\ \bar{\sigma}_{22(k)}\\ \bar{\sigma}_{33(k)}\\ \bar{\sigma}_{12(k)}\end{Bmatrix}^{(\alpha)}=\begin{bmatrix}\bar{C}_{11} & \bar{C}_{12} & \bar{C}_{16}\\ \bar{C}_{12} & \bar{C}_{22} & \bar{C}_{26}\\ \bar{C}_{13} & \bar{C}_{23} & \bar{C}_{36}\\ \bar{C}_{16} & \bar{C}_{26} & \bar{C}_{66}\end{bmatrix}^{(\alpha)}\times\begin{Bmatrix}u_{1(k+1),1*}\\ u_{2(k+1),2*}\\ u_{1(k+1),2*}+u_{2(k+1),1*}\end{Bmatrix}^{(\alpha)}+\begin{Bmatrix}\bar{C}_{13}\\ \bar{C}_{23}\\ \bar{C}_{33}\\ \bar{C}_{36}\end{Bmatrix}^{(\alpha)}\bar{g}^{(\alpha)}_{33(k)} \quad \text{in } \Omega^{(\alpha)}, \tag{50}$$

where $\bar{f}^{(\alpha)}_{i(k)}$ and $g^{(\alpha)}_{33(k)}$ are forcing terms consisting of the derivatives of lower-order displacements:

$$\bar{f}^{(\alpha)}_{i(0)}\equiv 0, \quad \bar{f}^{(\alpha)}_{i(k)}\equiv\bar{\rho}^{(\alpha)}\ddot{\bar{u}}^{(\alpha)}_{i(k-1)}-\bar{\sigma}^{(\alpha)}_{3i(k-1),3} \quad \text{for } k=1,2,\ldots, \tag{51}$$

$$\bar{g}^{(\alpha)}_{33(k)}\equiv\bar{u}^{(\alpha)}_{3(k),3} \quad \text{for } k=0,1,2,\ldots. \tag{52}$$

At material interfaces, the continuity conditions are expressed by (43), (44), and (45) for $k+1$ and $i=1,2$.

The antiplane problems are defined as

$$\bar{\sigma}^{(\alpha)}_{13(k),1*}+\bar{\sigma}^{(\alpha)}_{23(k),2*}=\bar{f}^{(\alpha)}_{3(k)} \quad \text{in } \bar{\Omega}^{(\alpha)}, \tag{53}$$

$$\begin{Bmatrix}\bar{\sigma}_{23(k)}\\ \bar{\sigma}_{31(k)}\end{Bmatrix}^{(\alpha)}=\begin{bmatrix}\bar{C}_{44} & \bar{C}_{45}\\ \bar{C}_{45} & \bar{C}_{55}\end{bmatrix}^{(\alpha)}\begin{Bmatrix}\bar{u}_{3(k+1),2*}+\bar{g}_{23(k)}\\ \bar{u}_{3(k+1),1*}+\bar{g}_{31(k)}\end{Bmatrix}^{(\alpha)} \quad \text{in } \bar{\Omega}^{(\alpha)}, \tag{54}$$

where $\bar{f}^{(\alpha)}_{3(k)}$ has been defined in (51) for $i=3$ and $\bar{g}^{(\alpha)}_{23(k)}$ and $\bar{g}^{(\alpha)}_{31(k)}$ are forcing terms

$$\bar{g}^{(\alpha)}_{23(k)}\equiv\bar{u}^{(\alpha)}_{2(k),3}, \quad \bar{g}^{(\alpha)}_{31(k)}\equiv\bar{u}^{(\alpha)}_{1(k),3}. \tag{55}$$

The solutions of the above problems furnish the warping $\bar{u}^{(\alpha)}_{3(k+1)}$ of the cross section. The corresponding boundary conditions are (43)–(45) for $i=3$. Since the solution for each MBVP is unique up to a constant field, the orthogonality of $\bar{u}^{(\alpha)}_{i(k+1)}$ to the lower-order displacements with respect to the integration weighted by mass density over the cross section is imposed.

In order to solve (49) and (50) or (53) and (54) for $\bar{u}^{(\alpha)}_{i(k+1)}$, one must impose the integrability condition:

$$\sum_{\alpha=1}^{3}\int\int_{\bar{\Omega}^{(\alpha)}}\bar{f}^{(\alpha)}_{i(k)}\,d\bar{x}^*_1\,d\bar{x}^*_2=0, \tag{56}$$

which is a global statement of the traction-free condition (45).

The antiplane problem (53) and (54) with $k=0$ is defined with the forcing terms

$$\bar{f}^{(\alpha)}_{3(0)}\equiv 0, \quad \bar{g}^{(\alpha)}_{23(0)}\equiv\bar{u}_{2(0),3}, \quad \bar{g}^{(\alpha)}_{31(0)}\equiv\bar{u}_{1(0),3}. \tag{57}$$

The solution yields pure bending modes

$$\bar{u}^{(\alpha)}_{3(1)}=-\bar{u}_{2(0),3}(\bar{x}_3,\bar{t})\bar{x}^*_2-\bar{u}_{1(0),3}(\bar{x}_3,\bar{t})\bar{x}^*_1, \tag{58a}$$

$$\bar{\sigma}^{(\alpha)}_{23(0)}=\bar{\sigma}^{(\alpha)}_{31(0)}=0. \tag{58b}$$

The pure bending displacement field (58*a*) has been incorporated in the trial displacement field (25*c*) through $\Psi_i$ $(i=1,2)$.

The remaining MBVP's are induced by either $\bar{u}_{i(0)}$ or $\bar{\Psi}_3$. By utilizing the superposition principle for linear problems, one can consider independently: (i) the torsional deformation due to $\bar{\Psi}_3$, (ii) the pure bending and flexural deformation induced by $\bar{u}_{1(0)}$ and $\bar{u}_{2(0)}$, and (iii) the longitudinal deformation excited by $\bar{u}_{3(0)}$.

## Constitutive Relation for Torsional Deformation

For the torsional mode, MBVP's are generated by $\bar{\Psi}_3$. The antiplane problem (MBVP1), (53) and (54) with $k=1$, gives Saint Venant's torsion problem for heterogeneous cross sections. The forcing terms are obtained from (51), (55), and (48) as

$$\bar{f}^{(\alpha)}_{3(1)}=0, \quad \bar{g}^{(\alpha)}_{23(1)}=\bar{\Psi}_{3,3}\bar{x}^*_1, \quad \bar{g}^{(\alpha)}_{31(1)}=-\bar{\Psi}_{3,3}\bar{x}^*_2. \tag{59}$$

To solve MBVP1 the torsional warping field is expressed in the form

$$\bar{u}^{(\alpha)}_{3(2)}=\bar{\Psi}_{3,3}(\bar{x}_3,\bar{t})\tilde{u}^{(\alpha)}_{3(2)}(\bar{x}^*_1,\bar{x}^*_2), \tag{60a}$$

$$[\bar{\sigma}_{23(1)} \quad \bar{\sigma}_{31(1)}]^{(\alpha)}=\bar{\Psi}_{3,3}(\bar{x}_3,\bar{t})[\tilde{\sigma}_{23(1)} \quad \tilde{\sigma}_{31(1)}]^{(\alpha)}(\bar{x}^*_1,\bar{x}^*_2). \tag{60b}$$

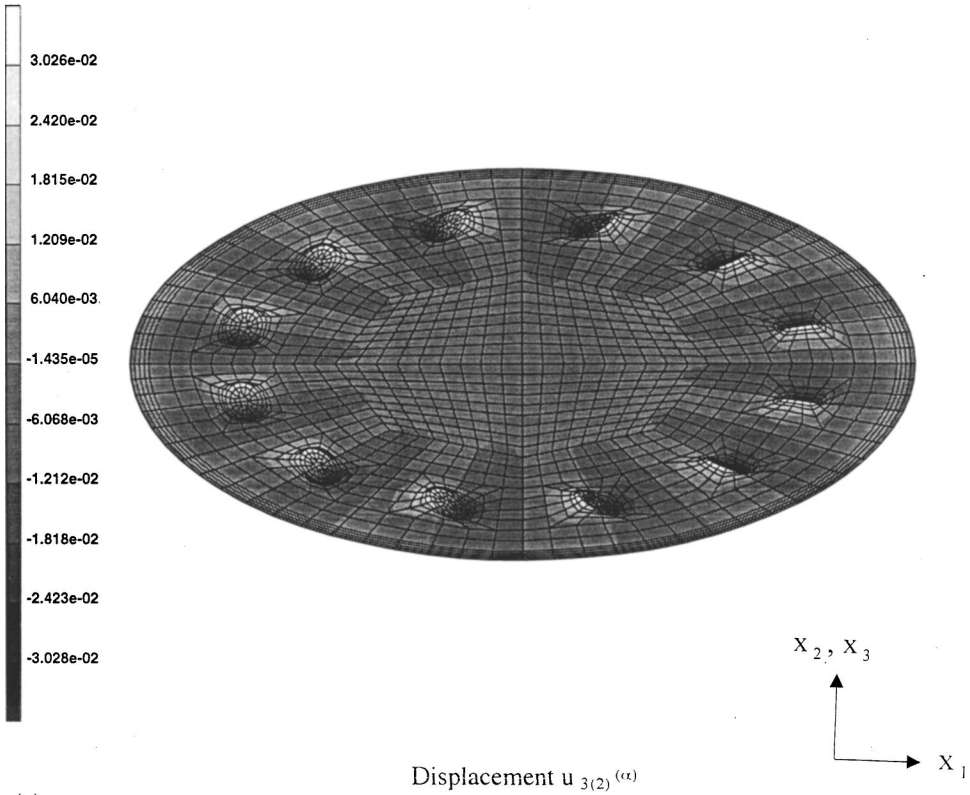
The above Saint Venant torsion problem (MBVP1) can be cast in a variational form for finite element analyses as

**Table 1 Material properties of steel and concrete**

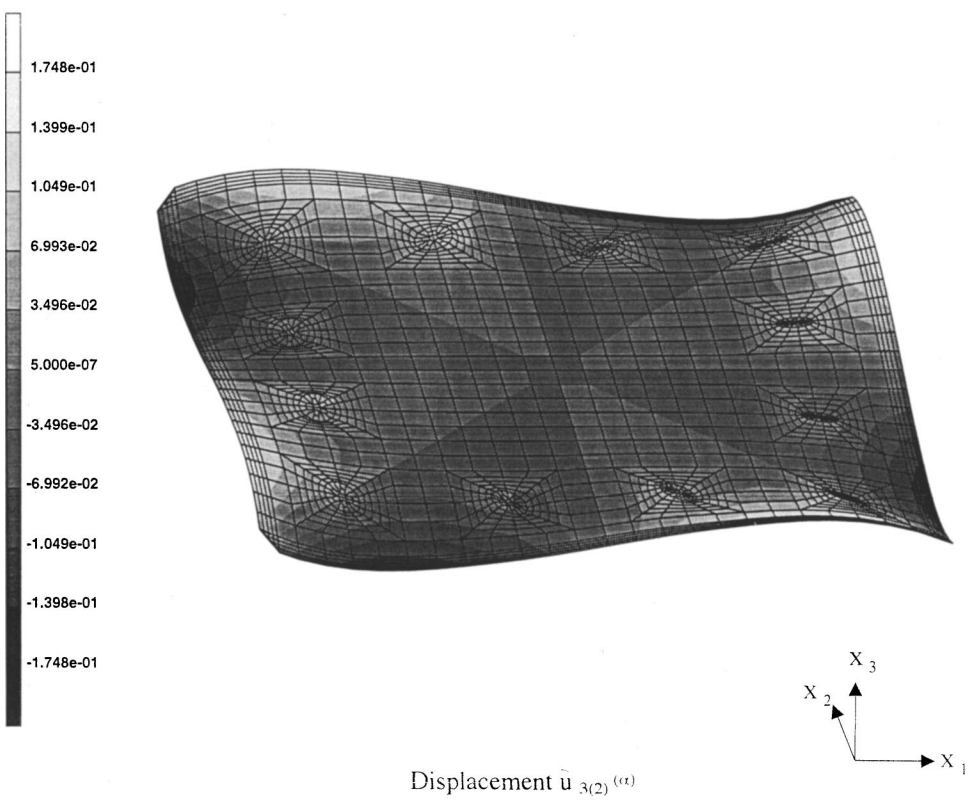
| | Steel (1) | Concrete (2) |
|---|---|---|
| $E$ | 210 GPa | 21 GPa |
| $G$ | 80.8 GPa | 9.1 GPa |
| $\nu$ | 0.3 | 0.15 |
| $\rho$ | 7800 kg/m$^3$ | 2300 kg/m$^3$ |

**Table 2 Material properties for the carbon/epoxy layer**

| $E_L$ | $E_T$ | $G_{LT}$ | $\nu_{LT}$ | $\nu_{TT}$ | $\rho$ |
|---|---|---|---|---|---|
| 148 GPa | 9.65 GPa | 4.55 GPa | 0.3 | 0.3 | 1500 kg/m$^3$ |




(*a*)




(*b*)

**Fig. 2** (*a*) **A contour plot of torsional warping $\tilde{u}_{3(2)}^{(\alpha)}$ for the circular cross section,** (*b*) **a contour plot of torsional warping $\tilde{u}_{3(2)}^{(\alpha)}$ for the square cross section**

$$\sum_{\alpha=1}^{3}\int\int_{\bar{\Omega}^{(\alpha)}}(\delta\tilde{u}^{(\alpha)}_{3(2),2*}\tilde{\sigma}^{(\alpha)}_{23(1)}+\delta\tilde{u}^{(\alpha)}_{3(2),1*}\tilde{\sigma}^{(\alpha)}_{31(1)})d\bar{x}^*_1\,d\bar{x}^*_2=0. \tag{61}$$

The circular and square cross sections, shown in Fig. 1, were discretized into four-node quadrilateral elements. Due to the symmetry of cross sections with respect to $\bar{x}^*_1$ and $\bar{x}^*_2$-axes, a mesh for only one quadrant is needed. Material properties are shown in Tables 1 and 2. The details of the cross sections will be deferred to the section on numerical results.

For the circular cross section, the torsional warping $\tilde{u}^{(\alpha)}_{3(2)}$ is shown in Fig. 2(*a*). Had this been a homogeneous circular cross section, no warping would have occurred ([20,21]). Figure 2(*b*) illustrates the contours of the torsional warping for the square cross section. The considerable amount of warping observed in Fig. 2(*b*) could be approximated by that for the homogeneous square cross-section of concrete.

The trial displacement field for torsional deformation is obtained from (25*a*,*b*) as

$$u^{(\alpha)}_1=-\Psi_3(x_3,t)x_2,\quad u^{(\alpha)}_2=\Psi_3(x_3,t)x_1,\quad u^{(\alpha)}_3=0. \tag{62}$$

By incorporating the torsional warping effect, a trial transverse-stress field is constructed in dimensional form as

$$\begin{gathered}\sigma^{(\alpha)}_{11}=\sigma^{(\alpha)}_{22}=\sigma^{(\alpha)}_{12}=0,\\ [\sigma_{23}\ \ \sigma_{31}]^{(\alpha)}=Q^a(x_3,t)[\tilde{\sigma}_{23(1)}\ \ \tilde{\sigma}_{31(1)}]^{(\alpha)}(x_1,x_2),\end{gathered} \tag{63}$$

where $Q^a(x_3,t)$ is a stress variable. From (63) the torque $M_3$ as defined in (28*b*) is expressed in terms of the stress variable $Q^a$ as

$$M_3(x_3,t)=q_aQ^a(x_3,t), \tag{64a}$$

where $q_a$ is the constant

$$q_a\equiv\sum_{\alpha=1}^{3}\int\int_{\Omega^{(\alpha)}}(x_1\tilde{\sigma}^{(\alpha)}_{23(1)}-x_2\tilde{\sigma}^{(\alpha)}_{31(1)})dx_1\,dx_2. \tag{64b}$$

In the above, by taking advantage of Reissner's mixed variational principle, the torsional warping is neglected in the trial displacement field (to avoid an additional beam displacement variable and the corresponding equation of motion). However, the torsional warping effect is included in the trial shear stress field to accurately predict the torsional rigidity.

The torsional constitutive relation is obtained from (33) with (62) and (63):

$$\begin{aligned}&\sum_{\alpha=1}^{3}\int\int_{\Omega^{(\alpha)}}[\delta\sigma^{(\alpha)}_{23}\{\Psi_{3,3}x_1-(D^{(\alpha)}_{44}\sigma^{(\alpha)}_{23}+D^{(\alpha)}_{45}\sigma^{(\alpha)}_{31})\}\\ &\quad+\delta\sigma^{(\alpha)}_{31}\{-\Psi_{3,3}x_2-(D^{(\alpha)}_{45}\sigma^{(\alpha)}_{23}+D^{(\alpha)}_{55}\sigma^{(\alpha)}_{31})\}]dx_1\,dx_2\\ &=\delta Q^a(q_a\Psi_{3,3}-q_bQ^a)=0,\end{aligned} \tag{65a}$$

where $q_b$ is the constant:

$$\begin{aligned}q_b\equiv&\sum_{\alpha=1}^{3}\int\int_{\Omega^{(\alpha)}}[\tilde{\sigma}_{23(1)}\ \ \tilde{\sigma}_{31(1)}]^{(\alpha)}\begin{bmatrix}D_{44} & D_{45}\\ D_{45} & D_{55}\end{bmatrix}^{(\alpha)}\\ &\times\begin{Bmatrix}\tilde{\sigma}_{23(1)}\\ \tilde{\sigma}_{31(1)}\end{Bmatrix}^{(\alpha)}dx_1\,dx_2.\end{aligned} \tag{65b}$$

As a result, the constitutive relation for the torque becomes

$$M_3=K_{00}\Psi_{3,3}=\frac{q_a^2}{q_b}\Psi_{3,3}. \tag{66}$$

If the elementary torsion theory were used (without incorporating the torsional warping), the errors in torsional rigidity would have been approximately 20 percent for circular cross sections and 40 percent for square cross sections ([22]).

## Constitutive Relations for Flexural Deformation

MBVP's for pure bending and flexural modes are generated by $\bar{u}_{1(0)}$ and $\bar{u}_{2(0)}$ in (47). In what follows, only the MBVP's induced by $\bar{u}_{1(0)}$ are presented. The in-plane problem defined by (49) and (50) with $k=0$ yields a trivial deformation field, $u^{(\alpha)}_{i(1)}=0$ ($i=1,2$), due to the lack of forcing terms, while the problem with $k=1$ furnishes MBVP2 for $\bar{u}^{(\alpha)}_{1(2)}$ and $\bar{u}^{(\alpha)}_{2(2)}$. The forcing terms for MBVP2 are obtained from (51) and (52) with (58*a*,*b*) as

$$\bar{f}^{(\alpha)}_{1(1)}=\bar{\rho}^{(\alpha)}\ddot{\bar{u}}_{1(0)},\quad \bar{f}^{(\alpha)}_{2(1)}=0, \tag{67a}$$

$$\bar{g}^{(\alpha)}_{33(1)}=-\bar{u}_{(0),33}\bar{x}^*_1. \tag{67b}$$

The integrability condition (56) yields

$$\ddot{\bar{u}}_{1(0)}=0. \tag{67c}$$

The solution of MBVP2 yields in-plane displacements for pure bending:

$$\bar{u}^{(\alpha)}_{i(2)}=\bar{u}_{1(0),33}(\bar{x}_3,\bar{t})\hat{u}^{(\alpha)}_{i(2)}(\bar{x}^*_1,\bar{x}^*_2),\quad i=1,2, \tag{68a}$$

$$\begin{aligned}&[\bar{\sigma}_{11(1)}\ \ \bar{\sigma}_{22(1)}\ \ \bar{\sigma}_{33(1)}\ \ \bar{\sigma}_{12(1)}]^{(\alpha)}\\ &\quad=\bar{u}_{1(0),33}(\bar{x}_3,\bar{t})[\hat{\sigma}_{11(1)}\ \ \hat{\sigma}_{22(1)}\ \ \hat{\sigma}_{33(1)}\ \ \hat{\sigma}_{12(1)}]^{(\alpha)}(\bar{x}^*_1,\bar{x}^*_2).\end{aligned} \tag{68b}$$

For homogeneous isotropic cross sections, $\hat{\sigma}_{33(1)}$ becomes proportional to $\bar{x}^*_1$, $\hat{\sigma}_{11(1)}=\hat{\sigma}_{22(1)}=\hat{\sigma}_{12(1)}=0$, and $\hat{u}_{1(2)}$ and $\hat{u}_{2(2)}$ are, respectively, proportional to $(\bar{x}^{*2}_1-\bar{x}^{*2}_2)/2$ and $\bar{x}^*_1\bar{x}^*_2$ ([20,21]). For the reinforced concrete cross sections with and without composite layers, the finite element solution of MBVP2 shows that the stress and displacement fields can be approximated with good accuracy by those for homogeneous isotropic cross sections.

The solution of the antiplane problem (53) and (54) with $k=1$ is a trivial field. A nontrivial stress field is obtained from the antiplane problem with $k=2$ (MBVP3). The forcing terms of MBVP3 come from (51), (55), (58*a*), (67*c*), and (68*a*,*b*):

$$\bar{f}^{(\alpha)}_{3(2)}=-\bar{u}_{1(0),333}\hat{\sigma}^{(\alpha)}_{33(1)}, \tag{69a}$$

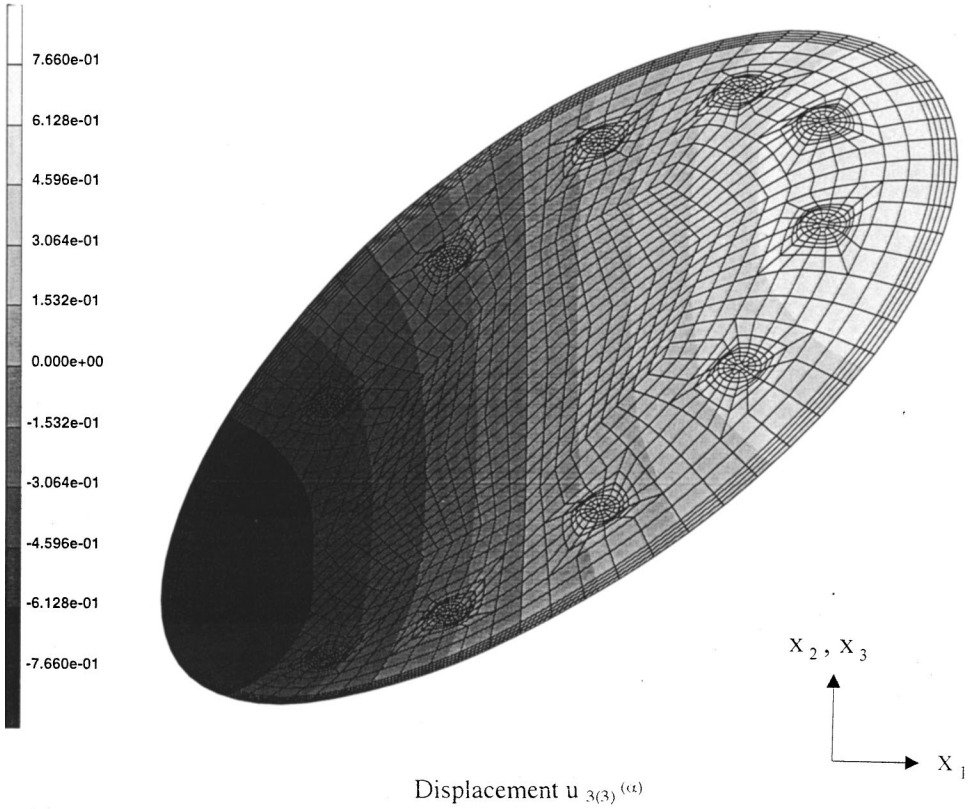
$$\bar{g}^{(\alpha)}_{23(2)}=\bar{u}_{1(0),333}\hat{u}^{(\alpha)}_{2(2)},\quad \bar{g}^{(\alpha)}_{31(2)}=\bar{u}_{1(0),333}\hat{u}^{(\alpha)}_{1(2)}. \tag{69b}$$

The integrability condition (56) for $i=3$ is automatically satisfied. The deformation field for the flexural warping is expressed in the form

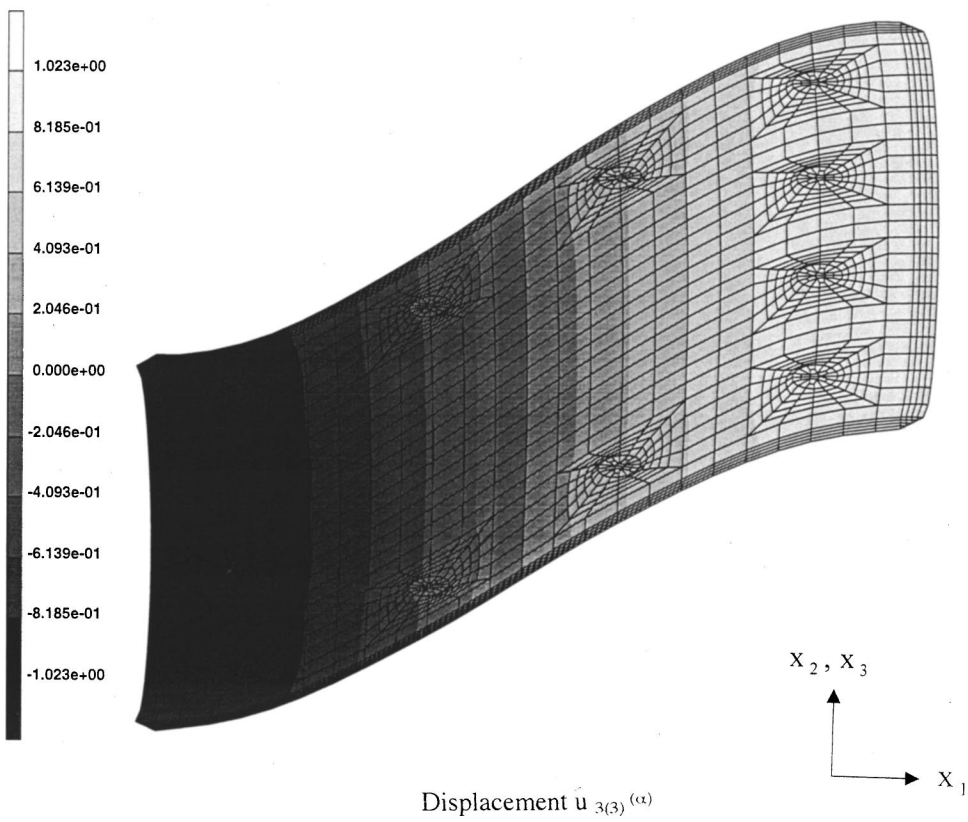
$$\bar{u}^{(\alpha)}_{3(3)}=\bar{u}_{1(0),333}\hat{u}^{(\alpha)}_{3(3)}(\bar{x}^*_1,\bar{x}^*_2), \tag{70a}$$

$$[\bar{\sigma}_{23(2)}\ \ \bar{\sigma}_{31(2)}]^{(\alpha)}=\bar{u}_{1(0),333}[\hat{\sigma}_{23(2)}\ \ \hat{\sigma}_{31(2)}]^{(\alpha)}(\bar{x}^*_1,\bar{x}^*_2). \tag{70b}$$

MBVP3 for $\hat{u}^{(\alpha)}_{3(3)}$ is known as Saint Venant's flexure problem ([20,21]). Contour plots of the flexural warping $\hat{u}^{(\alpha)}_{3(3)}$ are shown in Figs. 3(*a*) and 3(*b*) for the circular and square cross sections, respectively. (The contour plots are drawn on deformed cross sections.) Due to the forcing term $\hat{\sigma}^{(\alpha)}_{33(1)}$ in (69*a*), that is almost proportional to $\bar{x}^*_1$, the flexural warping $\hat{u}^{(\alpha)}_{3(3)}$ is dominated by the $\bar{x}^{*3}_1$-contribution. (The warping fields for the cross sections without composite layers are similar to the fields in Figs. 3(*a*) and 3(*b*). Comparisons of the contours in Figs. 3(*a*) and 3(*b*) with those for the homogeneous cross sections reveal that the heterogeneity induces only minute changes in warping. Therefore, in the construction of high-order beam models (without using Reissner's




(*a*)




(*b*)

**Fig. 3** (*a*) **A contour plot of flexural warping $\hat{u}^{(\alpha)}_{3(3)}$ for the circular cross section,** (*b*) **a contour plot of flexural warping $\hat{u}^{(\alpha)}_{3(3)}$ for the square cross section**

semi-complementary energy function), the above result justifies the use of a cubic-polynomial approximation for the axial displacement.

The trial displacement field is obtained from (25) as

$$u_1^{(\alpha)}=U_1(x_3,t),\quad u_2^{(\alpha)}=0,\quad u_3^{(\alpha)}=-\Psi_2(x_3,t)x_1. \tag{71}$$

By utilizing the stress fields defined by (68) and (70), a trial transverse stress field is constructed in dimensional form for flexural deformation as

$$[\sigma_{11}\ \ \sigma_{22}\ \ \sigma_{12}]^{(\alpha)}=S^b(x_3,t)[\hat{\sigma}_{11(1)}\ \ \hat{\sigma}_{22(1)}\ \ \hat{\sigma}_{12(1)}]^{(\alpha)}(x_1,x_2), \tag{72a}$$

$$[\sigma_{23}\ \ \sigma_{31}]^{(\alpha)}=Q^c(x_3,t)[\hat{\sigma}_{23(2)}\ \ \hat{\sigma}_{31(2)}]^{(\alpha)}(x_1,x_2), \tag{72b}$$

where $S^b(x_3,t)$ and $Q^c(x_3,t)$ are stress variables.

The stress variables $N_1$ and $M_2$ as defined in (28*a*,*b*) are computed by using (72) and the semi-inverted constitutive relation (22*b*) for $\sigma_{33}$ together with (71) as follows:

$$N_1=q_{1c}Q^c, \tag{73a}$$

$$M_2=q_{1d}\Psi_{2,3}+q_{1e}S^b, \tag{73b}$$

where

$$q_{1c}\equiv\sum_{\alpha=1}^{3}\int\int_{\Omega^{(\alpha)}}\hat{\sigma}_{31(2)}^{(\alpha)}\,dx_1\,dx_2, \tag{73c}$$

$$q_{1d}\equiv\sum_{\alpha=1}^{3}\int\int_{\Omega^{(\alpha)}}\frac{x_1^2}{D_{33}^{(\alpha)}}\,dx_1\,dx_2, \tag{73d}$$

$$q_{1e}\equiv\sum_{\alpha=1}^{3}\int\int_{\Omega^{(\alpha)}}x_1(D_{13}''^{(\alpha)}\hat{\sigma}_{11(1)}^{(\alpha)}+D_{23}''^{(\alpha)}\hat{\sigma}_{22(1)}^{(\alpha)}+D_{36}''^{(\alpha)}\hat{\sigma}_{12(1)}^{(\alpha)})\,dx_1\,dx_2. \tag{73e}$$

Substitution of (72) into (33) gives

$$\delta S^b(q_{1e}\Psi_{2,3}-q_{1f}S^b)+\delta Q^c\{q_{1c}(U_{1,3}-\Psi_2)-q_{1g}Q^c\}=0, \tag{74a}$$

where the constant $q_{1f}$ is defined as

$$q_{1f}\equiv\sum_{\alpha=1}^{3}\int\int_{\Omega^{(\alpha)}}\begin{Bmatrix}\hat{\sigma}_{11(1)}\\ \hat{\sigma}_{22(1)}\\ \hat{\sigma}_{12(1)}\end{Bmatrix}^{(\alpha)T}\begin{bmatrix}D_{11}' & D_{12}' & D_{16}'\\ D_{12}' & D_{22}' & D_{26}'\\ D_{16}' & D_{26}' & D_{66}'\end{bmatrix}^{(\alpha)}\times\begin{Bmatrix}\hat{\sigma}_{11(1)}\\ \hat{\sigma}_{22(1)}\\ \hat{\sigma}_{12(1)}\end{Bmatrix}^{(\alpha)}dx_1\,dx_2, \tag{74b}$$

and the constant $q_{1g}$ is obtained from (65*b*) by substituting $\hat{\sigma}_{23(2)}^{(\alpha)}$ and $\hat{\sigma}_{31(2)}^{(\alpha)}$ for $\tilde{\sigma}_{23(1)}^{(\alpha)}$ and $\tilde{\sigma}_{31(1)}^{(\alpha)}$.

For arbitrary variations of $S^b$ and $Q^c$-the following relations are obtained:

$$q_{1e}\Psi_{2,3}-q_{1f}S^b=0, \tag{75a}$$

$$q_{1g}Q^c=q_{1c}(U_{1,3}-\Psi_2). \tag{75b}$$

By solving (75*a*,*b*) for $S^b$ and $Q^c$ and substituting the result into (73*a*, *b*), the following constitutive relations are obtained:

$$N_1=K_{11}(U_{1,3}-\Psi_2), \tag{76a}$$

$$M_2=K_{99}\Psi_{2,3}, \tag{76b}$$

where

$$[K_{11}\ \ K_{99}]\equiv\left[\frac{q_{1c}^2}{q_{1g}}\quad q_{1d}+\frac{q_{1e}^2}{q_{1f}}\right]. \tag{76c}$$

The constitutive relations for $N_2$ and $M_1$ can also be obtained in a similar manner:

$$N_2=K_{22}(U_{2,3}+\Psi_1), \tag{77a}$$

$$M_1=K_{88}\Psi_{1,3}. \tag{77b}$$

Equations (76*a*) and (77*a*) incorporate the shear correction factors associated with the Timoshenko beam theory ([23,24]).

## Constitutive Relations for Longitudinal Deformation

The longitudinal mode is generated by $\overline{u}_{3(0)}$ in (47). The first problem is a generalized plane-strain problem (MBVP4) for $\overline{u}_{i(1)}^{(\alpha)}$ ($i=1,2$) defined by (49) and (50) for $k=0$ with the forcing terms $\overline{u}_{3(0),3}$ independent of $\overline{x}_1^*$ and $\overline{x}_2^*$.

$$\overline{f}_{i(0)}^{(\alpha)}=0,\quad \overline{g}_{33(0)}^{(\alpha)}=\overline{u}_{3(0),3}. \tag{78}$$

The rigid-body mode of MBVP4 has already been accounted for as the torsional deformation mode in (48). The solution of MBVP4 is expressed as

$$\overline{u}_{i(1)}^{(\alpha)}=\overline{u}_{3(0),3}(\overline{x}_3,\overline{t})\hat{u}_{i(1)}^{(\alpha)}(\overline{x}_1^*,\overline{x}_2^*),\quad i=1,2, \tag{79a}$$

$$[\overline{\sigma}_{11(0)}\ \ \overline{\sigma}_{22(0)}\ \ \overline{\sigma}_{33(0)}\ \ \overline{\sigma}_{12(0)}]^{(\alpha)}=\overline{u}_{3(0),3}(\overline{x}_3,\overline{t})[\hat{\sigma}_{11(0)}\ \ \hat{\sigma}_{22(0)}\ \ \hat{\sigma}_{33(0)}\ \ \hat{\sigma}_{12(0)}]^{(\alpha)}(\overline{x}_1^*,\overline{x}_2^*). \tag{79b}$$

For homogeneous beams $\hat{u}_{i(1)}$ ($i=1,2$) describes the deformation field due to Poisson's lateral contraction or expansion and the corresponding stresses become

$$\hat{\sigma}_{11(0)}=\hat{\sigma}_{22(0)}=\hat{\sigma}_{12(0)}=0,\quad \hat{\sigma}_{33(0)}=\overline{E}.$$

For circular cross section with and without composite layers, numerical results show that the displacement field $\hat{u}_{i(1)}^{(\alpha)}$ ($i=1,2$) is almost identical to that for the homogeneous circular cross sections and is proportional to $\overline{x}_i^*$. For the square cross section without the composite layers, $\hat{u}_{i(1)}^{(\alpha)}$ ($i=1,2$) is also proportional to $\overline{x}_i^*$. However, for the square cross section with the composite layer, shown in Fig. 4, the contours of $\hat{u}_{i(1)}^{(\alpha)}$ deviate considerably from vertical lines that indicate the proportionality to $\overline{x}_1^*$. This deviation from $\overline{x}_1^*$ is due to the constraint by the exterior composite layer on the Poisson contraction.

The trial displacement field employed for longitudinal deformation is in dimensional form:

$$u_1^{(\alpha)}=\Phi(x_3,t)\hat{u}_{1(1)}^{(\alpha)}(x_1,x_2),\quad u_2^{(\alpha)}=\Phi(x_3,t)\hat{u}_{2(1)}^{(\alpha)}(x_1,x_2),\quad u_3^{(\alpha)}=U_3(x_3,t)+\Delta_3(x_3,t)\hat{u}_{3(2)}^{(\alpha)}(x_1,x_2). \tag{80}$$

The trial transverse stress field is expressed through stress variables $S^a$, $S^c$, $Q^b$, and $Q^d$ as

$$[\sigma_{11}\ \ \sigma_{22}\ \ \sigma_{12}]^{(\alpha)}=S^a(x_3,t)[\hat{\sigma}_{11(0)}\ \ \hat{\sigma}_{22(0)}\ \ \hat{\sigma}_{12(0)}]^{(\alpha)}(x_1,x_2)+S^c(x_3,t)[\hat{\sigma}_{11(2)}\ \ \hat{\sigma}_{22(2)}\ \ \hat{\sigma}_{12(2)}]^{(\alpha)}(x_1,x_2), \tag{81a}$$

$$[\sigma_{23}\ \ \sigma_{31}]^{(\alpha)}=Q^b(x_3,t)[\hat{\sigma}_{23(1)}\ \ \hat{\sigma}_{31(1)}]^{(\alpha)}(x_1,x_2)+Q^d(x_3,t)[\hat{\sigma}_{23(3)}\ \ \hat{\sigma}_{31(3)}]^{(\alpha)}(x_1,x_2). \tag{81b}$$

The displacement mode associated with $\Phi$ in (80) and the transverse stress mode associated with $S^a$ in (81*a*) have been defined in MBVP4. The longitudinal warping $\hat{u}_{3(2)}^{(\alpha)}$ associated with $\Delta_3$ and the corresponding shear stresses $\hat{\sigma}_{23(1)}^{(\alpha)}$ and $\hat{\sigma}_{31(1)}^{(\alpha)}$ represented by $Q^b$ are obtained from the solution of MBVP5 defined by (55) and (56) for $k=1$. The integrability condition (56) for $i=3$ yields the elementary rod equation

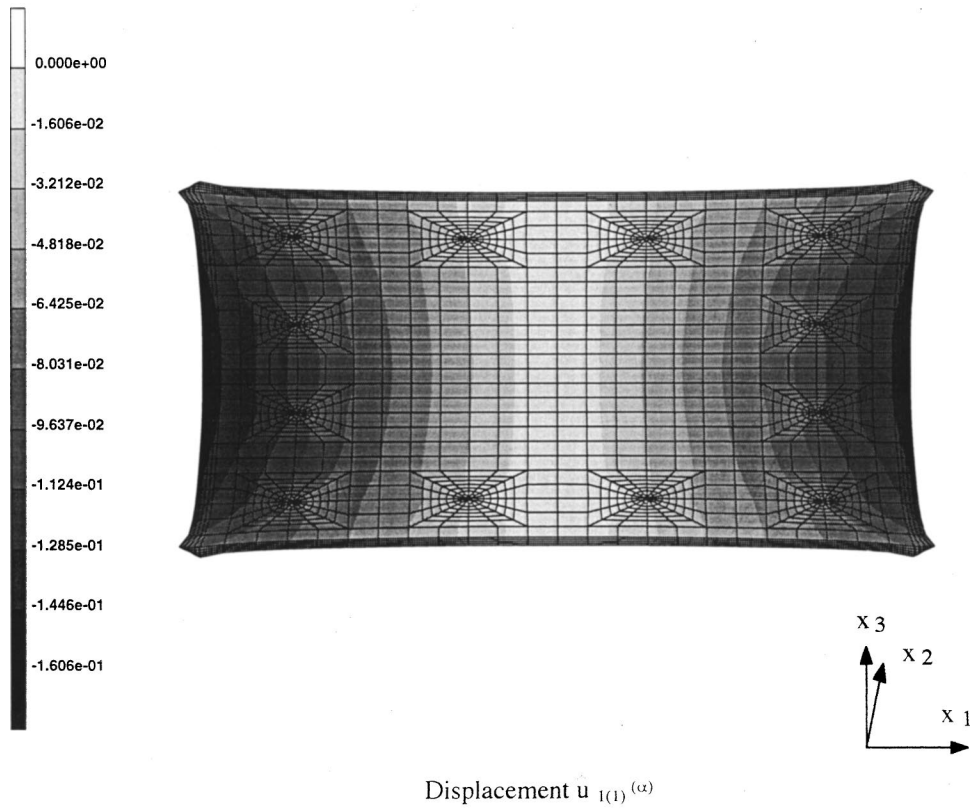



**Fig. 4 A contour plot of $\hat{u}^{(\alpha)}_{1(1)}$ for the square cross section**

$$\overline{E}^{(m)}\overline{A}\overline{u}_{3(0),33}-\overline{m}\ddot{\overline{u}}_{3(0)}=0, \tag{82a}$$

where

$$[\overline{A}\ \ \overline{m}\ \ \overline{E}^{(m)}\overline{A}]\equiv\sum_{\alpha=1}^{3}\int\int_{\overline{\Omega}^{(\alpha)}}[1\ \ \overline{\rho}^{(\alpha)}\ \ \hat{\sigma}^{(\alpha)}_{33(0)}]d\overline{x}_1^*\,d\overline{x}_2^*\,, \tag{82b}$$

in which $\overline{A}$, $\overline{m}$, and $\overline{E}^{(m)}$ are, respectively, the total area, the mass per unit axial length, and the effective Young modulus of the cross section.

The forcing terms for MBVP5 become

$$\overline{f}^{(\alpha)}_{3(1)}=\left(\overline{E}^{(m)}\overline{\rho}^{(\alpha)}\frac{\overline{A}}{\overline{m}}-\hat{\sigma}^{(\alpha)}_{33(0)}\right)\overline{u}_{3(0),33}\,, \tag{83a}$$

$$\overline{g}^{(\alpha)}_{23(1)}=\overline{u}_{3(0),33}\hat{u}^{(\alpha)}_{2(1)}\,,\quad \overline{g}^{(\alpha)}_{31(1)}=\overline{u}_{3(0),33}\hat{u}^{(\alpha)}_{1(1)}\,. \tag{83b}$$

Figures 5(*a*) and 5(*b*) show the contour plots of the longitudinal warping $\hat{u}^{(\alpha)}_{3(2)}$ for the heterogeneous circular and rectangular cross sections. The longitudinal warping $\hat{u}_{3(2)}$ for the homogeneous circular cross section can be found analytically to be proportional to $(2\overline{x}_1^{*2}+2\overline{x}_2^{*2}-1)/4$. The corresponding contours of $\hat{u}_{3(2)}$ become concentric circles. The contours of $\hat{u}_{3(2)}$ for homogeneous square cross sections are also approximated by concentric circles. Therefore, reinforcing bars alter the warping field moderately for the circular cross section in Fig. 5(*a*) and significantly for the square cross section in Fig. 5(*b*).

The solutions of MBVP4 and MBVP5 are sufficient for the construction of a Love-Rayleigh-type rod model for heterogeneous beams with the constraints (34*b*) ([25]). In this paper, to capture the longitudinal wave dispersion, a Mindlin-Herrmann-type model is developed in (80) and (81). The transverse stress mode $\hat{\sigma}^{(\alpha)}_{ij(2)}$ $(i,j=1,2)$ associated with $S^c$ is defined by (49) and (50) with $k=2$. The mode is excited by $\overline{u}_{3(0),333}$ (MBVP6). The integrability condition for MBVP6 is automatically satisfied. The transverse shear stress mode $\hat{\sigma}^{(\alpha)}_{j3(3)}$ $(j=1,2)$ associated with $Q^d$ results from the antiplane problem for $\overline{u}^{(\alpha)}_{3(4)}$ (MBVP7) defined by (53) and (54) with $k=3$. (The antiplane problem with $k=2$ yields $\overline{u}^{(\alpha)}_{3(3)}=0$.)

The constitutive relations for $N_3^I$, $N_3^{II}$, $H^I$, $H^{II}$, and $H^{III}$ are now obtained in two steps. First they are expressed in terms of stress variables $S$s and $Q$s as well as $U_{3,3}$ and $\Delta_{3,3}$ by using (28*a*,*b*,*c*,*d*) and (22*b*) for $\sigma_{33}$ with (81). Next the use of (33) with (80) and (81) yields the relation between the stress variables and strain variables: $U_{3,3}$, $\Delta_{3,3}$, $\Phi$, $\Phi_3$, and $\Delta_3$. The resulting constitutive relations become

$$\begin{Bmatrix} N_3^I \\ N_3^{II} \\ H^I \end{Bmatrix}=\begin{bmatrix} K_{33} & K_{34} & K_{35} \\ K_{34} & K_{44} & K_{45} \\ K_{35} & K_{45} & K_{55} \end{bmatrix}\begin{Bmatrix} U_{3,3} \\ \Delta_{3,3} \\ \Phi \end{Bmatrix}, \tag{84}$$

$$\begin{Bmatrix} H^{II} \\ H^{III} \end{Bmatrix}=\begin{bmatrix} K_{66} & K_{67} \\ K_{67} & K_{77} \end{bmatrix}\begin{Bmatrix} \Phi_{,3} \\ \Delta_3 \end{Bmatrix}. \tag{85}$$

The derivation of the beam constitutive relations is now complete. They are (66), (76), (77), (84), and (85).

## Phase Velocity Spectra of the Heterogeneous Beam Model

The phase velocity spectra of the fundamental modes have been used to assess the accuracy of one-dimensional beam and rod models for wave propagation ([6,7]). In this section, the derivation of the phase velocity spectra of the heterogeneous beam model is briefly described.

The displacements are written in the form of a harmonic wave mode propagating in the $x_3$-direction with the wave number $\xi$

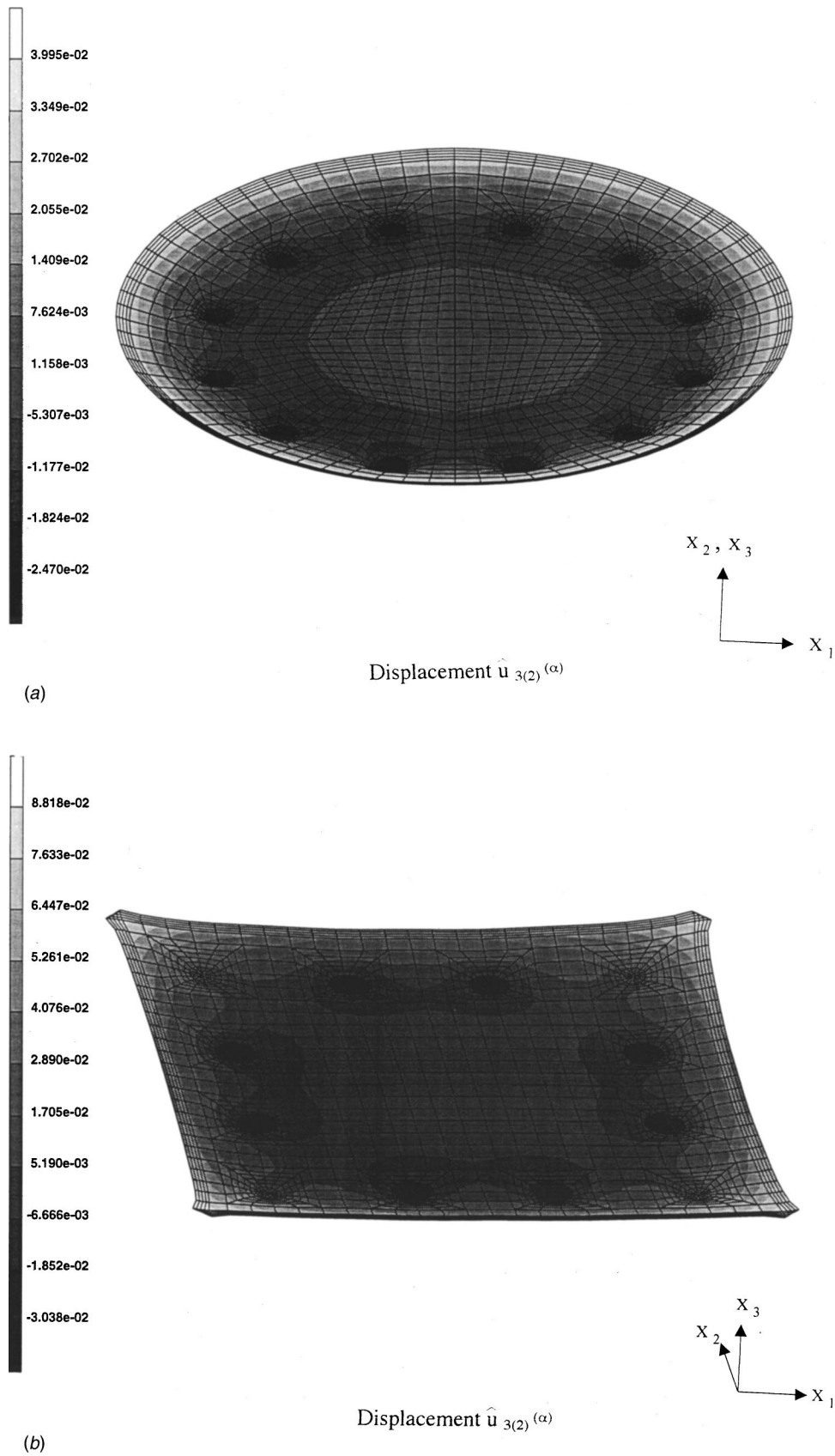



**Fig. 5** (*a*) **A contour plot of longitudinal warping $\hat{u}^{(\alpha)}_{3(2)}$ for the circular cross section,** (*b*) **a contour plot of longitudinal warping $\hat{u}^{(\alpha)}_{3(2)}$ for the square cross section**

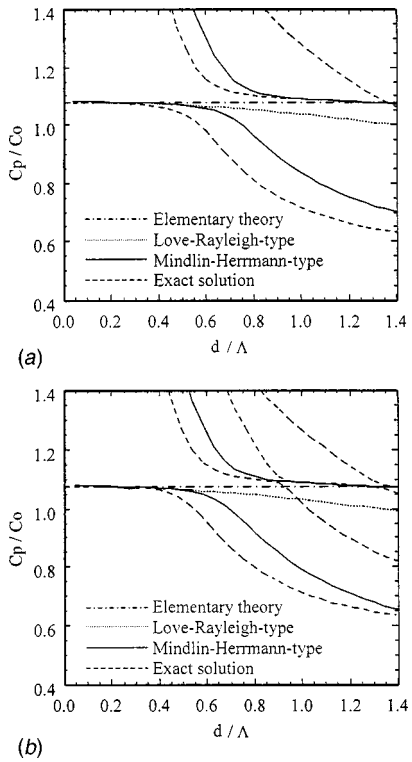



**Fig. 6 (*a*) Longitudinal phase velocity spectra for the beam with the circular cross section, (*b*) longitudinal phase velocity spectra for the beam with the square cross section**

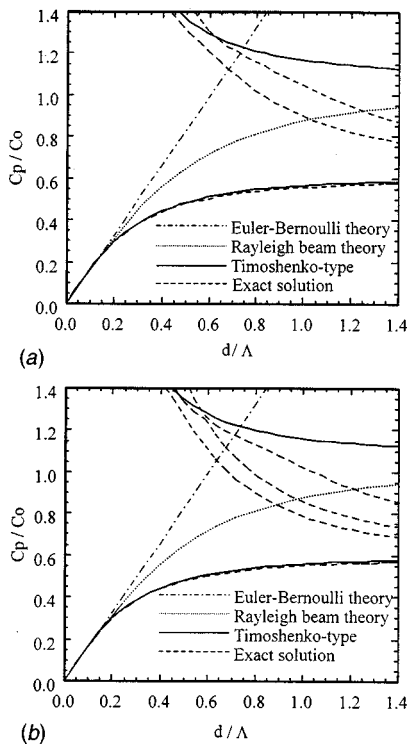



**Fig. 7 (*a*) Flexural phase velocity spectra for the beam with the circular cross section, (*b*) flexural phase velocity spectra for the beam with the square cross section**

($\equiv 2\pi$/wave length) and angular frequency $\omega$. For longitudinal wave motion the following harmonic waveform is assumed:

$$[U_3 \;\; \Delta_3 \;\; \Phi] = \exp\sqrt{-1}(\xi x_3 - \omega t)[\hat{U}_3 \;\; \hat{\Delta}_3 \;\; \hat{\Phi}]. \tag{86}$$

The substitution of (86) into (31*c*–*e*) for free vibration rewritten in terms of displacements via (84) and (85) yields an eigenvalue problem for a given wave number $\xi$:

$$\begin{bmatrix} m\omega^2 - \xi^2 K_{33} & -\xi^2 K_{34} & -\xi^2 K_{35} \\ -\xi^2 K_{34} & J_d\omega^2 - K_{77} - \xi^2 K_{44} & \xi^2(K_{67} - K_{45}) \\ -\xi^2 K_{35} & \xi^2(K_{67} - K_{45}) & \xi^2(J_p\omega^2 - K_{55} - \xi^2 K_{66}) \end{bmatrix} \times \begin{Bmatrix} \hat{U}_3 \\ \hat{\Delta}_3 \\ \hat{\Phi} \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \\ 0 \end{Bmatrix}. \tag{87}$$

Phase velocity $C_p$ and group velocity $C_g$ are obtained from $\xi$ and $\omega$ as ([10])

$$C_p = \frac{\omega}{\xi}, \quad C_g = \frac{d\omega}{d\xi}. \tag{88}$$

The characteristic Eq. (87) can be solved to find $C_p$ for a given $\xi$ by using a symbolic manipulator such as Mathematica.

For circular and square cross sections, it suffices to consider the flexural motion only in the $x_1$, $x_3$-plane due to the double symmetry. A harmonic flexural wave is considered in the form

$$[U_1 \;\; \Psi_2] = \exp\sqrt{-1}(\xi x_3 - \omega t)[\hat{U}_1 \;\; \sqrt{-1}\xi\hat{\Psi}_2]. \tag{89}$$

Substitution of (89) into (31*a*) and (31*g*) rewritten in terms of displacements via (76) yields an eigenvalue problem where one finds $\omega$s for a given $\xi$:

$$\begin{bmatrix} m\omega^2 - \xi^2 K_{11} & \xi^2 K_{11} \\ \xi^2 K_{11} & \xi^2(J_2\omega^2 - K_{11} - \xi^2 K_{99}) \end{bmatrix} \begin{Bmatrix} \hat{U}_1 \\ \hat{\Psi}_2 \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix}. \tag{90}$$

For torsional motion one substitutes

$$\Psi_3 = \hat{\Psi}_3 \exp\sqrt{-1}(\xi x_3 - \omega t) \tag{91}$$

into (31*h*) rewritten with (66) and finds that the phase velocity is independent of the wave number (nondispersive):

$$C_p = \sqrt{\frac{K_{00}}{J_3}}. \tag{92}$$

In the next section the predicted phase velocity spectra computed from (87), (90), and (92) will be compared to numerically obtained velocity spectra by using three-dimensional semi-analytical method ([11]).

## Numerical Results

The columns considered in the above MBVP's have exterior composite layers and circular and square cross sections of equal area. The diameter of the circular cross section is 500 mm. Each column is reinforced longitudinally by 12 steel bars of diameter 25 mm. The circular cross section has its bars arranged periodically every 30 degrees on a radius of 192.5 mm. For the column with the square cross section, longitudinal bars are arranged four per face as shown in Fig. 1. The material properties used for the calculations are shown in Table 1. The exterior composite layer has a thickness of 12.5 mm and is made of unidirectional carbon/epoxy composite with fibers in the circumferential direction. The material properties of the carbon/epoxy layer are shown in Table 2. The MBVP's were solved numerically by developing the finite element code that was validated by solving Saint Venent's problems for circular cross sections ([20,21]).

Figures 6(*a*) and 6(*b*) show the longitudinal phase velocity spectra for the columns with circular and square cross sections,

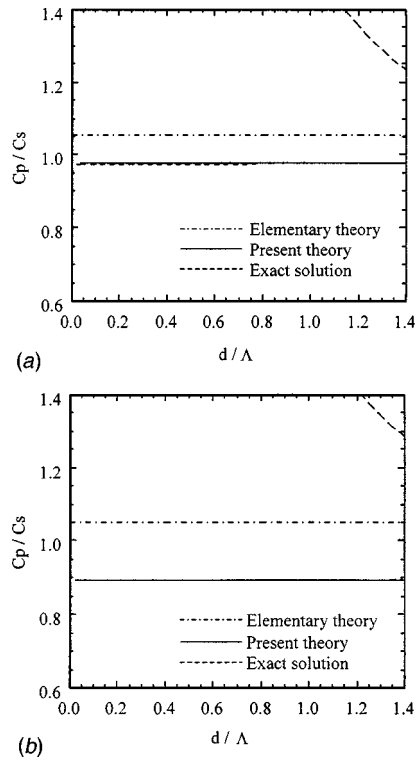



**Fig. 8 (*a*) Torsional phase velocity spectra for the beam with the circular cross section, (*b*) torsional phase velocity spectra for the beam with the square cross section**

respectively. In these figures solid lines represent predictions made by the Mindlin-Herrmann-type equation (87), while dashed lines show the exact phase velocity spectra obtained numerically by using the finite element waveguide analyses ([11]). The Mindlin-Herrmann-type rod model predicts dispersion of longitudinal waves unlike the nondispersive elementary rod model, shown by a dash-dotted line. In the figure the phase velocity spectra predicted by the Love-Rayleigh-type model are also shown by fine dotted lines. The Love-Rayleigh-type model predicts only minute dispersion due to the small Poisson's ratio of concrete, 0.15.

Figures 7(*a*) and 7(*b*) show the flexural phase velocity spectra for the beam with circular and square cross sections. The phase velocity spectra predicted by the Timoshenko-type beam model (90) are almost identical to the spectra obtained from the finite element waveguide analyses ([11]) shown by dashed lines. In the figures the predictions of the elementary Euler-Bernoulli theory and the Rayleigh theory are also plotted by dash-dotted and fine dotted lines, respectively.

Figures 8(*a*) and 8(*b*) include torsional phase velocity spectra for the beams with circular and square cross sections. The predicted phase velocity spectra based upon (92) shown by solid lines are extremely close to the spectra computed by using the finite element waveguide analyses ([22]). Since the torsional rigidity in (66) incorporates torsional warping effect, Eq. (92) accurately predicts the wave speed in the long wavelength range.

## Conclusions

By making a dynamic extension of Reissner's mixed variational equation, a one-dimensional beam model with a displacement and stress microstructure was developed for heterogeneous beams with uniform cross sections. Trial displacement and transverse stress fields were constructed from the solutions of the MBVP's defined over beam cross sections. Saint Venant's semi-inverse torsion and flexure problems for heterogeneous cross sections were included in the system of MBVP's. Unlike elementary beam theories, the heterogeneous beam model accurately predicts wave dispersion. Therefore, the proposed beam theory furnishes an economical, yet accurate, one-dimensional model including the stress concentration around reinforcing bars and the effect of composite retrofitting for dynamic reinforced-concrete frame analyses.


## Acknowledgment

The authors are thankful to the late Prof. Emeritus Eric Reissner for his advice on an earlier version of the manuscript.


## References


[1] Kapania, R. K., and Raciti, S., 1989, ''Recent Advances in Analysis of Laminated Beams and Plates, Part I: Shear Effects and Buckling,'' AIAA J., **27**, pp. 923–934.

[2] Bank, L. C., and Kao, C. H., 1990, ''Dynamic Response of Thin-Walled Composite Material Timoshenko Beams,'' ASME J. Energy Resour. Technol., **112**, pp. 149–154.

[3] Song, O., and Librescu, L., 1993, ''Free Vibration of Anisotropic Composite Thin-Walled Beams of Closed Cross-Section Contour,'' J. Sound Vib., **167**, pp. 129–147.

[4] Kosmatka, J. B., and Dong, S. B., 1991, ''Saint-Venant Solutions for Prismatic Anisotropic Beams,'' Int. J. Solids Struct., **28**, pp. 917–938.

[5] Pecastaings, F., Sun, Z., and Davenne, L., 1991, ''A Saint Venant's Heterogeneous Beam Theory,'' Eur. J. Mech. A/Solids, **10**, pp. 45–69.

[6] Timoshenko, S., 1921, ''On the Correction for Shear of the Differential Equation for Transverse Vibration of Prismatic Bars,'' Philos. Mag., **41**, pp. 744–746.

[7] Mindlin, R. D., and Herrmann, G., 1951, ''A One-Dimensional Theory of Compressional Waves in an Elastic Rod,'' *Proceedings of the 1st U.S. National Congress of Applied Mechanics*, ASME, New York, pp. 187–191.

[8] Pochhammer, L., 1876, ''Ueber die Fortpflanzungsgeschwindigkeiten Kleiner Schwingungen in Einem Unbegrenzten Isotropen Kreiscylinder,'' Z. Math., **81**, pp. 324–336.

[9] Love, A. E. H., 1927, *A Treatise on the Mathematical Theory of Elasticity*, 4th Ed., Cambridge University Press, London.

[10] Graff, K. F., 1975, *Wave Motion in Elastic Solids*, Ohio State University Press, Columbus, Ohio.

[11] Yamakawa, J., and Murakami, H., 1997, ''Longitudinal and Flexural Wave Propagation in Reinforced Concrete Columns,'' Int. J. Solids Struct., **34**, pp. 4357–4376.

[12] Onoe, M., McNiven, H. D., and Mindlin, R. D., 1962, ''Dispersion of Axially Symmetric Waves in Elastic Rods,'' ASME J. Appl. Mech., **29**, pp. 729–734.

[13] Pao, Y.-H., and Mindlin, R. D., 1960, ''Dispersion of Flexural Waves in an Elastic Circular Cylinder,'' ASME J. Appl. Mech., **27**, pp. 513–520.

[14] Reissner, E., 1984, ''On a Certain Mixed Variational Theorem and a Proposed Application,'' Int. J. Numer. Methods Eng., **20**, pp. 1366–1368.

[15] Murakami, H., Reissner, E., and Yamakawa, J., 1996, ''Anisotropic Beam Theories With Shear Deformation,'' ASME J. Appl. Mech., **63**, pp. 660–668.

[16] Lekhnitskii, S. G., 1963, *Theory of Elasticity of an Anisotropic Elastic Body*, Holdin-Day, San Francisco (translated by P. Fern).

[17] Reissner, E., 1950, ''On a Variational Theorem in Elasticity,'' J. Math. Phys., **29**, pp. 90–95.

[18] Bensoussan, A., Lions, J. L., and Papanicolaou, G., 1978, *Asymptotic Analysis for Periodic Structures*, North-Holland, Amsterdam.

[19] Murakami, H., and Hegemier, G. A., 1986, ''A Mixture Model for Unidirectionally Fiber-Reinforced Composites,'' ASME J. Appl. Mech., **53**, pp. 765–773.

[20] Timoshenko, S., and Goodier, J. N., 1970, *Theory of Elasticity*, 3rd Ed., McGraw-Hill, New York.

[21] Sokolnikoff, I. S., 1956, *Mathematical Theory of Elasticity*, 2nd Ed., McGraw-Hill, New York.

[22] Murakami, H., and Yamakawa, J., 1998, ''Torsional Wave Propagation in Reinforced Concrete Columns,'' Int. J. Solids Struct., **35**, pp. 2617–2637.

[23] Cowper, G. R., 1966, ''The Shear Coefficient in Timoshenko's Beam Theory,'' ASME J. Appl. Mech., **33**, pp. 335–340.

[24] Dhamarajan, S., and McCutchen, H., Jr., 1973, ''Shear Coefficients for Orthotropic Beams,'' J. Compos. Mater., **7**, pp. 530–535.

[25] Yamakawa, J., 1996, ''Mechanical Modeling of RC Beams,'' Ph.D. dissertation, University of California at San Diego, La Jolla, CA.

# Asymptotic Stability Analysis for Sheet Metal Forming—Part I: Theory

**W. Scherzinger**
Sandia National Laboratories,
P.O. Box 5800, MS 0847,
Albuquerque, NM 87185
Mem. ASME

**N. Triantafyllidis**
Department of Aerospace Engineering,
The University of Michigan,
Ann Arbor, MI 48109-2140
Fellow ASME

*In this paper is presented a general methodology for predicting puckering instabilities in sheet metal forming applications. A novel approach is introduced which does not use shell theory approximations. The starting point is Hill's stability functional for a three-dimensional rate-independent stressed solid which is modified for contact. By using a multiple scale asymptotic technique with respect to the small dimensionless thickness parameter $\varepsilon$, one can derive the two-dimensional version of the stability functional which is accurate up to $O(\varepsilon^4)$, thus taking into account bending effects. Loss of positive definiteness of this functional indicates possibility of a puckering instability in a sheet metal forming problem with a known stress and deformation state. An advantage of the proposed method is that the puckering investigation is independent of the algorithm used for calculating the deformed state of the sheet.* [S0021-8936(00)00804-7]

## 1 Introduction and Motivation

Stamping of sheet metal is one of the most widely used industrial manufacturing processes. There are three major problems that limit the formability of a stamped part which have to be accounted for in its design: springback, tearing, and puckering/wrinkling. Springback is the change in shape of the part that occurs after a part is removed from the blankholder/die assembly and is due to the elastic unloading of the part. Tearing is the splitting of the part in areas of high strain concentrations and is due to the localized necking of the sheet. Puckering, as defined by Devons [1] is a waviness of the sheet that is not in contact with the tooling surfaces and is a bifurcation buckling phenomenon due to the presence of compressive in-plane stresses in the sheet. When the same phenomenon occurs in areas that come into contact with the tooling, usually the flat part of the binder, the surface waviness phenomena is known as wrinkling.

In modeling a tearing problem, the difficulty is in the determination of the proper constitutive law since the phenomenon is local in nature. The difficulty in modeling springback is due to the geometry of the part and in the determination of its prestress state from which an elastic unloading takes place. Modeling of puckering requires both an accurate description of the constitutive behavior of the material and the solution of a boundary value problem. Moreover, experimental investigation of puckering faces the difficulty of the determination of the onset of the phenomenon, since imperfections in the form of minute amounts of surface waviness are always present in stamped parts.

Of interest here is the modeling of puckering instabilities for general stamping geometries. The standard approach thus far uses a shell-type analysis (usually in conjunction with a finite element method code) and follows the deformation of the part all the way to the formation of finite amplitude wrinkles (e.g., Taylor et al. [2]). More refined analyses use a linearized stability method to check for bifurcation in a part with a known prestress state obtained by using a shell-type analysis (e.g., Neal and Tugcu [3]). The obvious shortcoming of this approach is the stability results' dependence on the shell theory employed.

To overcome the inconsistencies associated with the use of a particular shell theory to calculate the onset of bifurcation in shell buckling problems, Triantafyllidis and Kwon [4] proposed reversing the order of the limiting process in the analysis, by first formulating the stability problem of the three-dimensional solid and then finding its critical load and buckling mode as the dimensionless thickness parameter, $\varepsilon$, goes to zero. This asymptotic methodology has recently been applied by Scherzinger and Triantafyllidis [5] to another similar problem, namely the buckling of slender beams with arbitrary cross sections (there the $\varepsilon$ parameter is the beam's slenderness defined as the square root of its sectional area over its length) where the interested reader can find another comprehensive application of the proposed methodology.

The departing point for our analysis is Hill's [6,7] stability functional for a three-dimensional elastoplastic solid, properly modified to account for contact with tooling surfaces. Since stability against puckering depends on the sign of the functional's minimum eigenvalue, the present work consists of a multiple scale asymptotic analysis to obtain the minimum eigenvalue and the corresponding eigenmode in terms of $\varepsilon$. The multiple scale analysis is a finite strain adaptation of the methodology proposed by Destuynder [8] for the consistent derivation of linear elastic shell theories from the corresponding three-dimensional equations of elasticity. The present method results in the calculation of the stability functional of a prestressed stamped sheet that is accurate to $O(\varepsilon^4)$. The functional includes bending stiffness effects and only requires a two-dimensional stress state and eigenmode (the degrees-of-freedom are the displacements of the midsurface). No shell theory assumptions are required and normality of plane sections and the plane stress assumption arise naturally as a part of the analysis.

The outline of this work is as follows: The presentation begins with a description of the kinematics for a shell of arbitrary shape in Section 2.1. The treatment of contact, essential for the stability of sheet metal forming problems, is presented in Section 2.2 followed by the statement of the variational problem to find the minimum eigenvalue in Section 2.3. The asymptotic analysis of the problem is presented in Section 3. Expansion of field quantities are given and substituted into the stability functional. The stability functional is evaluated in Section 3.2 and its two-dimensional form, suitable for sheet metal forming, is found accurate to $O(\varepsilon^4)$.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Sept. 24, 1999; final revision, Jan. 30, 2000. Associate Technical Editor: S. Kyriakides. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

 Copyright © 2000 by ASME 

## 2 Formulation

To begin the stability analysis for a sheet metal forming problem, one starts with Hill's [6] stability functional evaluated at a given equilibrium state[1] of a solid of volume $V$ and boundary $\partial V$

$$\mathcal{F}(\lambda,\mathbf{v})=\frac{1}{2}\int_V L^{ijkl}(\lambda)v_{k,l}v_{i,j}dV+\frac{1}{2}\int_{\partial V}K^{ik}v_k v_i dS. \quad (2.1)$$

Here $\lambda$ is a load parameter, which determines the current stress state and the internal variables of the solid, $v_i$ are the covariant components of a kinematically admissible perturbation away from the equilibrium state, $v_{i,j}$ denotes the covariant derivative with respect to the three-dimensional basis in the current configuration, $L^{ijkl}$ are the contravariant components of the incremental moduli that relate the rate of the first Piola-Kirchoff stress to the rate of the deformation gradient ($\dot{\mathbf{\Pi}}^T=\mathbf{L}:\dot{\mathbf{F}}$). The second integral in (2.1) is a penalty term that is introduced to account for the contact between the tooling surfaces and the sheet. The symmetry of the incremental moduli, $L^{ijkl}$, and the contact terms, $K^{ik}$, implies real eigenvalues for $\mathcal{F}$. Positive definiteness of (2.1) ensures stability of the structure at the given equilibrium state, in the sense that positive energy has to be externally supplied into the structure for any admissible perturbation. The idea proposed in this work is to take advantage of the slenderness of the structure to develop a two-dimensional form of the stability functional which is accurate to any desired order of the slenderness parameter.

**2.1 Geometric Preliminaries.** Consider a shell-like structure in its current configuration as shown in Fig. 1. The solid occupies a volume $V$ and has boundary $\partial V$ consisting of the top, bottom and lateral surfaces ($\partial V=\partial V_+\cup\partial V_-\cup\partial V_n$). Material points in the shell are identified by their convected coordinates, $\theta^i=(\theta^\alpha,\theta^3)$; a material point on the midsurface of the shell has position vector $\mathbf{r}(\theta^\alpha)$, and the domain of the midsurface of the sheet is denoted by $A$. The position vector, $\mathbf{p}$, for an arbitrary material point in the shell can be written as

$$\mathbf{p}(\theta^i)=\mathbf{r}(\theta^\alpha)+\theta^3\mathbf{n}(\theta^\alpha), \quad -\frac{h(\theta^\alpha)}{2}\leqslant\theta^3\leqslant\frac{h(\theta^\alpha)}{2}, \quad (2.2)$$

where $\mathbf{n}(\theta^\alpha)$ is the unit outward normal to the midsurface of the shell. At the material point on the midsurface, given by $\theta^\alpha$, the thickness of the shell is $h(\theta^\alpha)$.

The basis vectors for the midsurface, $\mathbf{a}_i$, and the three-dimensional solid, $\mathbf{g}_i$, are defined as follows:

$$\mathbf{a}_\alpha=\frac{\partial\mathbf{r}}{\partial\theta^\alpha}, \quad \mathbf{a}_3=\mathbf{n}; \quad \mathbf{g}_i=\frac{\partial\mathbf{p}}{\partial\theta^i}. \quad (2.3)$$

The midsurface and three-dimensional metrics are given by $a_{ij}=\mathbf{a}_i\cdot\mathbf{a}_j$ and $g_{ij}=\mathbf{g}_i\cdot\mathbf{g}_j$ with inverses $a^{ij}$ and $g^{ij}$. From (2.3) the relation between the $\mathbf{a}_i$ basis and the $\mathbf{g}_i$ basis is established with the help of $\mu_i^j$,

$$\mathbf{g}_i=\mu_i^j\mathbf{a}_j; \quad \mu_\beta^\alpha=\delta_\beta^\alpha-\theta^3 b_\beta^\alpha, \quad \mu_i^3=\mu_3^i=\delta_i^3,$$
$$\mathbf{a}_i=q_i^j\mathbf{g}_j; \quad q_i^j=(\mu_j^i)^{-1}, \quad (2.4)$$

where $b_\beta^\alpha$ are the mixed components of the curvature tensor for the midsurface. The evaluation of the gradient of the perturbation, introduced in (2.1), requires the derivatives of the midsurface basis vectors with respect to the convected coordinates

$$\frac{\partial\mathbf{a}_i}{\partial\theta^j}=t_{ij}^k\mathbf{a}_k. \quad (2.5)$$

[1]Here and subsequently Greek indices range from 1 to 2 and Latin indices range from 1 to 3.

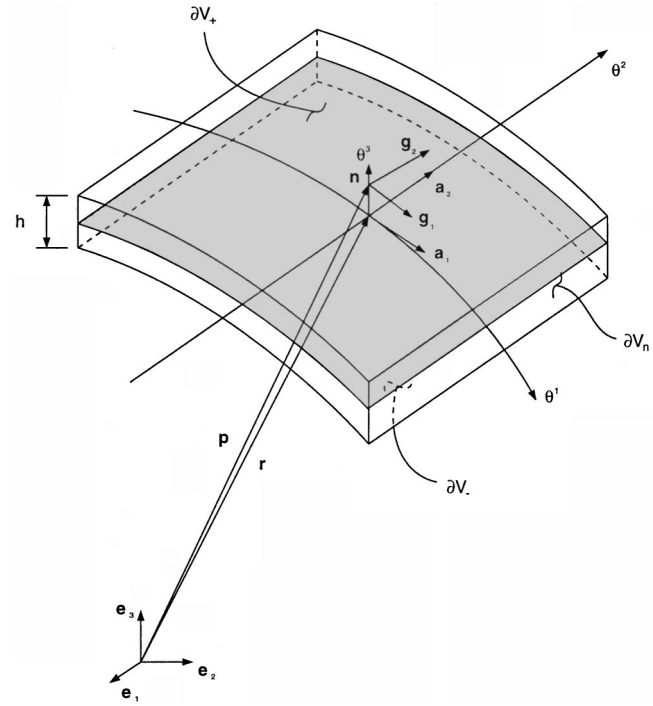



**Fig. 1 Three-dimensional kinematics for a shell-like structure. The midsurface is defined by the vector r while a point off the midsurface is defined by position vector p. The covariant midsurface basis is $\mathbf{a}_i$ and the covariant three-dimensional basis is $\mathbf{g}_i$.**

Since $\mathbf{a}_\alpha$ are the basis vectors for the midsurface, the coefficients $t_{\alpha\beta}^\gamma=\gamma_{\alpha\beta}^\gamma$ are the Christoffel symbols for the midsurface. From differential geometry (e.g., Goetz [9]), the only other nonzero components of $t_{ij}^k$ are

$$t_{\alpha\beta}^3=b_{\alpha\beta}=\alpha_{\alpha\gamma}b_\beta^\gamma, \quad t_{3\beta}^\gamma=-b_\beta^\gamma. \quad (2.6)$$

The components of any tensor in the subsequent analysis can be defined with respect to the three-dimensional basis, $\mathbf{g}_i$, or the midsurface basis, $\mathbf{a}_i$ (e.g., $\mathbf{v}=v^i\mathbf{g}_i=\bar{v}^i\mathbf{a}_i$).[2] The tensor components in (2.1) are referred to the three-dimensional basis; however, the following analysis naturally employs the midsurface basis. Therefore, relationships between the components of tensors referred to the two bases are needed. In particular, relationships are sought, with the help of (2.4) and (2.5), for the components of the incremental moduli, $\mathbf{L}$, and the gradient of the perturbation, $\mathbf{v}\nabla$,

$$L^{ijkl}=\bar{L}^{mnrs}q_m^i q_n^j q_r^k q_s^l; \quad v_{i,j}=\left(\frac{\partial\bar{v}_k}{\partial\theta^j}-t_{kj}^l\bar{v}_l\right)\mu_i^k=\bar{v}_{k,j}\mu_i^k. \quad (2.7)$$

**2.2 Treatment of Contact.** In metal forming applications, the sheet is in contact with the rigid surfaces of the punch, die, and blank holder, and this contact will constrain any kinematically admissible perturbation, $\mathbf{v}$. Therefore the effect of contact must be accounted for in the statement of the stability problem; this is done using a penalty-type formulation.

It is assumed, without loss of generality, that contact occurs between the tooling and the bottom surface of the sheet ($\partial V_-$). Denoting the energy due to contact as

[2]Here and subsequently components of tensors with respect to the midsurface basis will be denoted with a bar surmounting the symbol, $(\bar{\cdot})$.

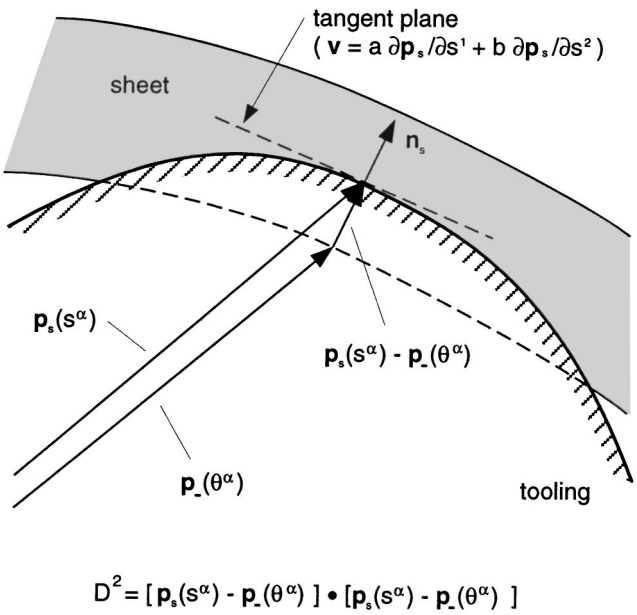



**Fig. 2 Contact between tooling surface and sheet. A point on the sheet's lower surface with position vector $\mathbf{p}_-$ is at a distance *D* from the rigid surface. The penetration distance is positive when $\mathbf{p}_- - \mathbf{p}_s$ has the same orientation as $\mathbf{n}_s$, the outward unit normal to the rigid surface.**

$$\mathcal{E} = \frac{1}{2}\int_{\partial V_-} kH(D)D^2 dA, \tag{2.8}$$

where $k$ is a foundation stiffness for the tooling, $D(\mathbf{u})$, a function of the displacement $\mathbf{u}$, is the interpenetration distance between the bottom surface of the sheet and the tooling surface, and $H(D)$ is the Heaviside step function ($H(D)=1$ if $D \geqslant 0$ and $H(D)=0$ if $D<0$), the contact term for the stability functional is augmented by

$$\frac{1}{2}(\mathcal{E}_{,uu}\delta_u)\delta u = \frac{1}{2}\int_{\partial V_-} kH(D)(\delta D\,\delta D + D\,\delta(\delta D))dA, \tag{2.9}$$

where $\mathcal{E}_{,uu}$ denotes the second functional derivative of $\mathcal{E}$ with respect to the displacement $\mathbf{u}$. Given a parameterization of the rigid surface in terms of the coordinates $s^\alpha$, the position vector for a point on this surface is given by $\mathbf{p}_s = \mathbf{p}_s(s^\alpha)$. The interpenetration distance between the sheet and the surface is defined as follows (see Fig. 2):

$$D^2 = [\mathbf{p}_s(s^\alpha) - \mathbf{p}_-(\theta^\alpha)]\cdot[\mathbf{p}_s(s^\alpha) - \mathbf{p}_-(\theta^\alpha)], \tag{2.10}$$

where $\mathbf{p}_-(\theta^\alpha) = \mathbf{p}(\theta^\alpha, -h(\theta^\alpha)/2)$ is the position vector of a material point on the bottom surface of the sheet that is in contact with the rigid surface. The distance $D$ is taken positive if there is a penetration of the sheet on the rigid surface, i.e., when $\mathbf{p}_s - \mathbf{p}_-$ is on the same direction as the outward normal to the rigid surface $\mathbf{n}_s$.

Given a material point $\mathbf{p}_-(\theta^\alpha)$, for a given equilibrium state, the point on the tooling surface, $\mathbf{p}_s(s^\alpha)$, is sought such that it minimizes $D$. This point is found to satisfy the following condition:

$$[\mathbf{p}_s(s^\alpha) - \mathbf{p}_-(\theta^\alpha)]\cdot\frac{\partial \mathbf{p}_s}{\partial s^\beta} = 0, \tag{2.11}$$

which states that the vector $\mathbf{p}_s - \mathbf{p}_-$ is orthogonal to the tangent plane of the surface at the point $\mathbf{p}_s$. With the above relationship, the penalty term used for modeling contact in the stability functional becomes

$$\frac{1}{2}\int_{\partial V_-} \mathbf{v}\cdot\mathbf{K}\cdot\mathbf{v}dA = \frac{1}{2}\int_{\partial V_-} kH(D)\left[\mathbf{v}\cdot\mathbf{v} - \left(\mathbf{v}\cdot\frac{\partial \mathbf{p}_s}{\partial s^\alpha}\right)\left(\mathbf{v}\cdot\frac{\partial \mathbf{p}_s}{\partial s^\beta}\right)\psi^{\alpha\beta}\right]dA, \tag{2.12}$$

where $\psi^{\alpha\beta}$ is

$$\psi^{\alpha\beta} = \left[\frac{\partial \mathbf{p}_s}{\partial s^\alpha}\cdot\frac{\partial \mathbf{p}_s}{\partial s^\beta} + (\mathbf{p}_s - \mathbf{p}_-)\cdot\frac{\partial^2 \mathbf{p}_s}{\partial s^\alpha \partial s^\beta}\right]^{-1}. \tag{2.13}$$

Finally, it is noted that the effect of friction is seen only through the principal solution of the sheet metal forming process; it is assumed that friction does not affect the stability calculations.

**2.3 Problem Statement.** Stability of the prestressed solid is guaranteed (see Hill [6]) if the functional (2.1) is positive definite, i.e., if its minimum eigenvalue, defined below, is positive

$$\beta = \min_{\mathbf{v}\in\mathcal{D}} \frac{2\mathcal{F}}{\langle \mathbf{v},\mathbf{v}\rangle}. \tag{2.14}$$

Here $\mathcal{D}$ is the space of kinematically admissible perturbations and $\langle \mathbf{v},\mathbf{v}\rangle$ is an appropriately chosen inner product. For the problem examined here, the following choice for the inner product is adopted:

$$\langle \mathbf{v},\mathbf{v}\rangle = \int_A \int_{-h/2}^{h/2} a^{ik}\overline{v}_k\overline{v}_i d\theta^3 dA. \tag{2.15}$$

The minimization problem in (2.14) can be reformulated as a variational problem

$$\beta\int_A\int_{-h/2}^{h/2} a^{ik}\overline{v}_k\delta\overline{v}_i d\theta^3 dA = \int_A\int_{-h/2}^{h/2} \overline{L}^{ijkl}\overline{v}_{k,l}\delta\overline{v}_{i,j}\mu d\theta^3 dA + \int_A [\overline{K}^{ik}\overline{v}_k\delta\overline{v}_i\mu]_{\theta^3=-h/2}dA. \tag{2.16}$$

In addition to the minimum eigenvalue, a mode uniqueness condition is also introduced:

$$\langle \mathbf{v},\mathbf{u}\rangle = \int_A\int_{-h/2}^{h/2} a^{ik}\overline{v}_k\overline{u}_i d\theta^3 dA = C \tag{2.17}$$

where $\mathbf{u}$ and $C$ are an appropriately chosen vector field and constant, respectively.

## 3 Asymptotics

The solution of the stability problem for a sheet metal forming application is found using the following approach: Given an equilibrium state for the solid, an asymptotic analysis is employed to deduce a stability criterion for the shell-like structure based on the minimum eigenvalue of Hill's functional for the three-dimensional solid. In this approach, the current geometry, stress state, and as a result, the incremental moduli are known a priori. Since the thickness is assumed to be small relative to the dimensions of the sheet and the minimum radius of curvature of the midsurface, the asymptotic analysis is performed using the ratio of the sheet thickness, $h$, relative to a characteristic length of the sheet, $l$, as the small parameter in the problem. Denoting this small parameter by $\varepsilon \equiv h/l$, asymptotic expansions are carried out in terms of $\varepsilon$. The scaled thickness coordinate, $\xi = \theta^3/\varepsilon$, is also used in the subsequent analysis and varies between $-\zeta \leqslant \xi \leqslant \zeta$, where $\zeta \equiv h/2\varepsilon$.

The minimum eigenvalue and all the field quantities are expanded in a regular perturbation series in terms of $\varepsilon^n$, $n=0,1,2,\ldots$ . By taking the expansions for these quantities, substituting them into (2.1), and evaluating the minimum eigenvalue of the stability functional, the stability of the sheet metal part is examined.

**3.1 Expansions.** For the analysis of the stability functional, the minimum eigenvalue, corresponding eigenmode, contact terms, and incremental moduli are expanded in a regular perturbation series in $\varepsilon$,

$$\beta(\varepsilon)=\overset{0}{\beta}+\varepsilon\overset{1}{\beta}+\varepsilon\overset{2}{\beta}+\cdots, \tag{3.1}$$

$$\overline{v_i}(\theta^j;\varepsilon)=\overset{0}{\overline{v_i}}(\theta^\alpha,\xi)+\varepsilon\overset{1}{\overline{v_i}}(\theta^\alpha,\xi)+\varepsilon^2\overset{2}{\overline{v_i}}(\theta^\alpha,\xi)+\cdots, \tag{3.2}$$

$$\overline{K}^{ik}(\theta^\alpha;\varepsilon)=\overset{0}{\overline{K}}{}^{ik}(\theta^\alpha)+\varepsilon\overset{1}{\overline{K}}{}^{ik}(\theta^\alpha)+\varepsilon^2\overset{2}{\overline{K}}{}^{ik}(\theta^\alpha)+\cdots, \tag{3.3}$$

$$\overline{L}^{ijkl}(\theta^\alpha;\varepsilon)=\overset{0}{\overline{L}}{}^{ijkl}(\theta^\alpha)+\varepsilon\overset{1}{\overline{L}}{}^{ijkl}(\theta^\alpha,\xi)+\varepsilon^2\overset{2}{\overline{L}}{}^{ijkl}(\theta^\alpha,\xi)+\cdots. \tag{3.4}$$

Similar expansions can be obtained for the geometric quantities $\mu_j^i$ and its inverse $q_j^i$ which relate the midsurface and three-dimensional covariant basis vectors (see (2.4)). Finally, it is noted that the foundation stiffness for the contact term in the problem must also be scaled with $\varepsilon$, $k\rightarrow\varepsilon k$. In order to physically motivate this rescaling, we recall from the theory of shells that the normal and shear stresses are on the order of $\varepsilon^2$ and $\varepsilon$, respectively. The rescaling is necessary to enforce this condition as $\varepsilon\rightarrow 0$. It should also be noted at this point that the adopted asymptotic expansions for the various field quantities are expected to be valid several thicknesses away from the midsurface boundary. Moreover, it is tacitly assumed that the corresponding boundary layer effects are inconsequential for the overall stability analysis of the structure.

**3.2 Evaluation of Stability Functional.** Using the above introduced rescaled expressions for $\theta^3$ and $k$, the variational equation for the minimum eigenvalue of the stability functional in (2.16) is rewritten as

$$\begin{aligned}
&\varepsilon\beta\int_A\int_{-\zeta}^{\zeta}a^{ik}\overline{v_k}\delta\overline{v_i}d\xi dA\\
&=\varepsilon\int_A\int_{-\zeta}^{\zeta}\left\{\frac{1}{\varepsilon^2}\overline{L}^{i3k3}\frac{\partial\overline{v_k}}{\partial\xi}\frac{\partial\delta\overline{v_i}}{\partial\xi}+\frac{1}{\varepsilon}(\overline{L}^{i3k\gamma}q_\gamma^\delta)\overline{v}_{k,\delta}\frac{\partial\delta\overline{v_i}}{\partial\xi}\right.\\
&\quad\left.+\frac{1}{\varepsilon}(\overline{L}^{i\alpha k3}q_\alpha^\beta)\frac{\partial\overline{v_k}}{\partial\xi}\delta\overline{v}_{i,\beta}+(\overline{L}^{i\alpha k\gamma}q_\alpha^\beta q_\gamma^\delta)\overline{v}_{k,\delta}\delta\overline{v}_{i,\beta}\right\}\mu d\xi dA\\
&\quad+\varepsilon\int_A[\overline{K}^{ik}\overline{v_k}\delta\overline{v_i}\mu]_{\xi=-\zeta}dA.
\end{aligned} \tag{3.5}$$

Substituting the expansions (3.1)–(3.4) into (3.5), and collecting terms of like order, the following governing equations are found that must be satisfied for the various orders of $\varepsilon$.

*Equations of $O(\varepsilon^{-1})$.* The lowest order governing equations are those of $O(\varepsilon^{-1})$, namely

$$\int_A\int_{-\zeta}^{\zeta}\overset{0}{\overline{L}}{}^{i3k3}\frac{\partial\overset{0}{\overline{v_k}}}{\partial\xi}\frac{\partial\delta\overline{v_i}}{\partial\xi}d\xi dA=0. \tag{3.6}$$

Integrating by parts, and in view of the arbitrariness of $\delta\overline{v_i}$, the governing differential equations and boundary conditions for $\overset{0}{\overline{v_k}}$ are found

$$\begin{gathered}
\frac{\partial}{\partial\xi}\left[\overset{0}{\overline{L}}{}^{i3k3}\frac{\partial\overset{0}{\overline{v_k}}}{\partial\xi}\right]=0\quad\text{in}\quad-\zeta\leqslant\xi\leqslant\zeta\\
\overset{0}{\overline{L}}{}^{i3k3}\frac{\partial\overset{0}{\overline{v_k}}}{\partial\xi}=0\quad\text{at}\quad\xi=\pm\zeta.
\end{gathered} \tag{3.7}$$

The solution of (3.7) is

$$\overset{0}{\overline{L}}{}^{i3k3}\frac{\partial\overset{0}{\overline{v_k}}}{\partial\xi}=0,\quad\Rightarrow\frac{\partial\overset{0}{\overline{v_k}}}{\partial\xi}=0. \tag{3.8}$$

Since $\overset{0}{\overline{L}}{}^{i3k3}$ is nonsingular for all applications (the material is assumed to be in the elliptic range of its response), the solution of (3.8) implies $\overset{0}{\overline{v_k}}=\overset{0}{\overline{v_k}}(\theta^\alpha)$, which states that the lowest order term in the expansion of the mode is only dependent on the midsurface coordinates.

*Equations of $O(\varepsilon^0)$.* Terms of $O(\varepsilon^0)$ are collected next and, making use of (3.8), the following governing equations are found:

$$\int_A\int_{-\zeta}^{\zeta}\left[\overset{0}{\overline{L}}{}^{i3k3}\frac{\partial\overset{1}{\overline{v_k}}}{\partial\xi}+\overset{0}{\overline{L}}{}^{i3k\delta}\overset{0}{\overline{v}}_{k,\delta}\right]\frac{\partial\delta\overline{v_i}}{\partial\xi}d\xi dA=0. \tag{3.9}$$

Following the same steps as in (3.6), we obtain the governing differential equations and boundary conditions for $\overset{1}{\overline{v_k}}$:

$$\begin{gathered}
\frac{\partial}{\partial\xi}\left[\overset{0}{\overline{L}}{}^{i3k3}\frac{\partial\overset{1}{\overline{v_k}}}{\partial\xi}+\overset{0}{\overline{L}}{}^{i3k\delta}\overset{0}{\overline{v}}_{k,\delta}\right]=0\quad\text{in}\quad-\zeta\leqslant\xi\leqslant\zeta\\
\overset{0}{\overline{L}}{}^{i3k3}\frac{\partial\overset{1}{\overline{v_k}}}{\partial\xi}+\overset{0}{\overline{L}}{}^{i3k\delta}\overset{0}{\overline{v}}_{k,\delta}=0\quad\text{at}\quad\xi=\pm\zeta.
\end{gathered} \tag{3.10}$$

Solving (3.10) gives the following result for $\partial\overset{1}{\overline{v_k}}/\partial\xi$:

$$\overset{0}{\overline{L}}{}^{i3k3}\frac{\partial\overset{1}{\overline{v_k}}}{\partial\xi}+\overset{0}{\overline{L}}{}^{i3k\delta}\overset{0}{\overline{v}}_{k,\delta}=0. \tag{3.11}$$

Recalling the plane stress assumption for the principal solution, $\tau^{i3}=0$ and $\dot{\tau}^{i3}=0$, and the resulting orthotropy of the incremental moduli, the Kirchoff-Love hypothesis for the deformation of thin shells is recovered for the eigenmode

$$\frac{\partial\overset{1}{\overline{v_\alpha}}}{\partial\xi}=-\overset{0}{\overline{v}}_{3,\alpha}. \tag{3.12}$$

*Equations of $O(\varepsilon^1)$.* The next lowest order equations are those of $O(\varepsilon^1)$. Making use of (3.8) and (3.11), the governing equations of $O(\varepsilon^1)$ are derived:

$$\begin{aligned}
&\int_A\int_{-\zeta}^{\zeta}\left\{\left[\overset{0}{\overline{L}}{}^{i3k3}\frac{\partial\overset{2}{\overline{v_k}}}{\partial\xi}+\overset{1}{\overline{L}}{}^{i3k3}\frac{\partial\overset{1}{\overline{v_k}}}{\partial\xi}+\overset{0}{\overline{L}}{}^{i3k\delta}\overset{1}{\overline{v}}_{k,\delta}\right.\right.\\
&\quad\left.+\left(\overset{1}{\overline{L}}{}^{i3k\delta}+\overset{0}{\overline{L}}{}^{i3k\gamma}\overset{1}{q}{}_\gamma^\delta\right)\overset{0}{\overline{v}}_{k,\delta}\right]\frac{\partial\delta\overline{v_i}}{\partial\xi}\\
&\quad\left.+\left(\overset{0}{\overline{L}}{}^{i\beta k3}\frac{\partial\overset{1}{\overline{v_k}}}{\partial\xi}+\overset{0}{\overline{L}}{}^{i\beta k\delta}\overset{0}{\overline{v}}_{k,\delta}\right)\delta\overline{v}_{i,\beta}\right\}d\xi dA\\
&\quad+\int_A[\overset{0}{\overline{K}}{}^{ik}\overset{0}{\overline{v_k}}\delta\overline{v_i}]_{\xi=-\zeta}dA\\
&=\overset{0}{\beta}\int_A\int_{-\zeta}^{\zeta}a^{ik}\overset{0}{\overline{v_k}}\delta\overline{v_i}\,d\xi dA.
\end{aligned} \tag{3.13}$$

From the results in (3.11), it becomes convenient to introduce the plane stress incremental moduli, $\overline{\mathcal{P}}^{i\beta k\delta}$,

$$\overline{\mathcal{P}}^{i\beta k\delta} \equiv [\overline{L}^{i\alpha k\gamma} - \overline{L}^{i\alpha m3}(\overline{L}^{n3m3})^{-1}\overline{L}^{n3k\gamma}]q^{\beta}_{\alpha}q^{\delta}_{\gamma}. \tag{3.14}$$

Using (3.14) and integrating (3.13) by parts, and assuming the thickness varies slowly throughout the sheet (i.e., $\partial\zeta/\partial\theta^{\alpha} = O(\varepsilon^n)$, $n \geqslant 2$), the following governing partial differential equations and boundary conditions are found:

$$\frac{\partial}{\partial\xi}\left[\overset{0}{\overline{L}}{}^{i3k3}\frac{\partial\overset{2}{\overline{v}}_k}{\partial\xi} + \overset{1}{\overline{L}}{}^{i3k3}\frac{\partial\overset{1}{\overline{v}}_k}{\partial\xi} + \overset{0}{\overline{L}}{}^{i3k\delta}\overset{1}{\overline{v}}_{k,\delta} + (\overset{1}{\overline{L}}{}^{i3k\delta} + \overset{0}{\overline{L}}{}^{i3k\gamma}\overset{1}{q}{}^{\delta}_{\gamma})\overset{0}{\overline{v}}_{k,\delta}\right]$$
$$+ \frac{\partial}{\partial\theta^{\beta}}\left[\overset{0}{\overline{\mathcal{P}}}{}^{i\beta k\delta}\overset{0}{\overline{v}}_{k,\delta}\right] + (\overset{0}{\overline{\mathcal{P}}}{}^{i\beta k\delta}t^{\gamma}_{\beta\gamma} + \overset{0}{\overline{\mathcal{P}}}{}^{m\beta k\delta}t^{i}_{m\beta})\overset{0}{\overline{v}}_{k,\delta} + \overset{0}{\beta}a^{ik}\overset{0}{\overline{v}}_k$$
$$= 0 \quad \text{in} \quad -\zeta \leqslant \xi \leqslant \zeta$$

$$\overset{0}{\overline{L}}{}^{i3k3}\frac{\partial\overset{2}{\overline{v}}_k}{\partial\xi} + \overset{1}{\overline{L}}{}^{i3k3}\frac{\partial\overset{1}{\overline{v}}_k}{\partial\xi} + \overset{0}{\overline{L}}{}^{i3k\delta}\overset{1}{\overline{v}}_{k,\delta} + (\overset{1}{\overline{L}}{}^{i3k\delta} + \overset{0}{\overline{L}}{}^{i3k\gamma}\overset{1}{q}{}^{\delta}_{\gamma})\overset{0}{\overline{v}}_{k,\delta}$$
$$= \overset{0}{\overline{K}}{}^{ik}\overset{0}{\overline{v}}_k \quad \text{at} \quad \xi = -\zeta$$

$$\overset{0}{\overline{L}}{}^{i3k3}\frac{\partial\overset{2}{\overline{v}}_k}{\partial\xi} + \overset{1}{\overline{L}}{}^{i3k3}\frac{\partial\overset{1}{\overline{v}}_k}{\partial\xi} + \overset{0}{\overline{L}}{}^{i3k\delta}\overset{1}{\overline{v}}_{k,\delta} + (\overset{1}{\overline{L}}{}^{i3k\delta} + \overset{0}{\overline{L}}{}^{i3k\gamma}\overset{1}{q}{}^{\delta}_{\gamma})\overset{0}{\overline{v}}_{k,\delta}$$
$$= 0 \quad \text{at} \quad \xi = \zeta. \tag{3.15}$$

Considering (3.13) with $\delta\overline{v}_i = \delta\overline{v}_i(\theta^{\alpha})$, one finds the following governing equations for $\overset{0}{\overline{v}}_i(\theta^{\alpha})$ in variational form

$$\int_A \left\{\overset{0}{\overline{\mathcal{P}}}{}^{i\beta k\delta}\overset{0}{\overline{v}}_{k,\delta}\delta\overline{v}_{i,\beta} + \frac{1}{2\zeta}\overset{0}{\overline{K}}{}^{ik}\overset{0}{\overline{v}}_k\delta\overline{v}_i\right\}dA = \overset{0}{\beta}\int_A a^{ik}\overset{0}{\overline{v}}_k\delta\overline{v}_i dA, \tag{3.16}$$

while the equations for $\partial\overset{2}{\overline{v}}_k/\partial\xi$ are found from (3.15) and (3.16):

$$\overset{0}{\overline{L}}{}^{i3k3}\frac{\partial\overset{2}{\overline{v}}_k}{\partial\xi} + \overset{1}{\overline{L}}{}^{i3k3}\frac{\partial\overset{1}{\overline{v}}_k}{\partial\xi} + \overset{0}{\overline{L}}{}^{i3k\delta}\overset{1}{\overline{v}}_{k,\delta} + (\overset{1}{\overline{L}}{}^{i3k\delta} + \overset{0}{\overline{L}}{}^{i3k\gamma}\overset{1}{q}{}^{\delta}_{\gamma})\overset{0}{\overline{v}}_{k,\delta}$$
$$= f(\xi)\overset{0}{\overline{K}}{}^{ik}\overset{0}{\overline{v}}_k, \tag{3.17}$$

where $f(\xi) \equiv (\zeta - \xi)/2\zeta$.

*Equations of $O(\varepsilon^2)$.* The next lowest order equations are those of $O(\varepsilon^2)$. Making use of (3.8), (3.11), (3.14), and (3.17), the governing equations of $O(\varepsilon^2)$ are found:

$$\int_A\int_{-\zeta}^{\zeta}\left\{\left[\overset{0}{\overline{L}}{}^{i3k3}\frac{\partial\overset{3}{\overline{v}}_k}{\partial\xi} + \overset{1}{\overline{L}}{}^{i3k3}\frac{\partial\overset{2}{\overline{v}}_k}{\partial\xi} + \overset{2}{\overline{L}}{}^{i3k3}\frac{\partial\overset{1}{\overline{v}}_k}{\partial\xi} + \overset{0}{\overline{L}}{}^{i3k\delta}\overset{2}{\overline{v}}_{k,\delta}\right.\right.$$
$$+ (\overset{1}{\overline{L}}{}^{i3k\delta} + \overset{0}{\overline{L}}{}^{i3k\gamma}\overset{1}{q}{}^{\delta}_{\gamma})\overset{1}{\overline{v}}_{k,\delta} + (\overset{2}{\overline{L}}{}^{i3k\delta} + \overset{1}{\overline{L}}{}^{i3k\gamma}\overset{1}{q}{}^{\delta}_{\gamma} + \overset{0}{\overline{L}}{}^{i3k\gamma}\overset{2}{q}{}^{\delta}_{\gamma})\overset{0}{\overline{v}}_{k,\delta}$$
$$\left. + \overset{1}{\mu}f(\xi)\overset{0}{\overline{K}}{}^{ik}\overset{0}{\overline{v}}_k\right]\frac{\partial\delta\overline{v}_i}{\partial\xi} + \left[\overset{0}{\overline{\mathcal{P}}}{}^{i\beta k\delta}\overset{1}{\overline{v}}_{k,\delta} + \overset{1}{\overline{\mathcal{P}}}{}^{i\beta k\delta}\overset{0}{\overline{v}}_{k,\delta} + \overset{1}{\mu}\overset{0}{\overline{\mathcal{P}}}{}^{i\beta k\delta}\overset{0}{\overline{v}}_{k,\delta}\right.$$
$$\left.\left. + f(\xi)\overset{0}{\overline{L}}{}^{i\beta m3}(\overset{0}{\overline{L}}{}^{n3m3})^{-1}\overset{0}{\overline{K}}{}^{nk}\overset{0}{\overline{v}}_k\right]\delta\overline{v}_{i,\beta}d\xi dA\right.$$
$$+ \int_A\left[(\overset{1}{\mu}\overset{0}{\overline{K}}{}^{ik}\overset{0}{\overline{v}}_k + \overset{1}{\overline{K}}{}^{ik}\overset{0}{\overline{v}}_k + \overset{0}{\overline{K}}{}^{ik}\overset{1}{\overline{v}}_k)\delta\overline{v}_i\right]_{\xi=-\zeta}dA$$
$$= \int_A\int_{-\zeta}^{\zeta}a^{ik}(\overset{0}{\beta}\overset{1}{\overline{v}}_k + \overset{1}{\beta}\overset{0}{\overline{v}}_k)\delta\overline{v}_i d\xi dA. \tag{3.18}$$

Choosing $\delta\overline{v}_i = \delta\overline{v}_i(\theta^{\alpha})$, and integrating through the thickness, gives the following equations for the $\xi$ independent part of $\overset{1}{\overline{v}}_i$:

$$\int_A\left\{\left(\overset{0}{\overline{\mathcal{P}}}{}^{i\beta k\delta}\langle\overset{1}{\overline{v}}_{k,\delta}\rangle + \langle\overset{1}{\overline{\mathcal{P}}}{}^{i\beta k\delta}\rangle\overset{0}{\overline{v}}_{k,\delta} + \frac{1}{2}\overset{0}{\overline{L}}{}^{i\beta m3}(\overset{0}{\overline{L}}{}^{n3m3})^{-1}\overset{0}{\overline{K}}{}^{nk}\overset{0}{\overline{v}}_k\right)\delta\overline{v}_{i,\beta}\right.$$
$$\left. + \frac{1}{2\zeta}\left[(\overset{1}{\mu}\overset{0}{\overline{K}}{}^{ik} + \overset{1}{\overline{K}}{}^{ik})\overset{0}{\overline{v}}_k + \overset{0}{\overline{K}}{}^{in}\overset{1}{\overline{v}}_k\right]_{\xi=-\zeta}\delta\overline{v}_i\right\}dA$$
$$= \int_A a^{ik}(\overset{0}{\beta}\langle\overset{1}{\overline{v}}_k\rangle + \overset{1}{\beta}\overset{0}{\overline{v}}_k)\delta\overline{v}_i dA, \tag{3.19}$$

where the thickness average of a function $f$ is defined as

$$\langle f\rangle \equiv \frac{1}{2\zeta}\int_{-\zeta}^{\zeta}f d\xi. \tag{3.20}$$

With the solution of (3.19), the mode, including first-order bending terms, is determined up to order $\varepsilon$:

$$\overline{v}_i = \overset{0}{\overline{v}}_i + \varepsilon(\langle\overset{1}{\overline{v}}_i\rangle - \xi(\overset{0}{\overline{L}}{}^{n3i3})^{-1}\overset{0}{\overline{L}}{}^{n3k\delta}\overset{0}{\overline{v}}_{k,\delta}) + O(\varepsilon^2). \tag{3.21}$$

An expression for the stability functional, accurate to $O(\varepsilon^3)$, can be found using this form of the mode that includes bending terms.

*Stability Functional up to $O(\varepsilon^3)$.* Using the previous results, the stability functional is assembled with accuracy up to $O(\varepsilon^3)$. Starting with the following expression for $\mathcal{F}$,

$$\mathcal{F} = \frac{\varepsilon}{2}\int_A\int_{-\zeta}^{\zeta}\left\{\overline{L}^{i3k3}\frac{\partial\overline{v}_k}{\partial\xi}\frac{\partial\overline{v}_i}{\partial\xi} + \overline{L}^{i3k\gamma}q^{\delta}_{\gamma}\overline{v}_{k,\delta}\frac{\partial\overline{v}_i}{\partial\xi}\right.$$
$$\left. + \overline{L}^{i\alpha k3}q^{\beta}_{\alpha}\frac{\partial\overline{v}_k}{\partial\xi}\overline{v}_{i,\beta} + \overline{L}^{i\alpha k\gamma}q^{\delta}_{\gamma}q^{\beta}_{\alpha}\overline{v}_{k,\delta}\overline{v}_{i,\beta}\right\}\mu d\xi dA$$
$$+ \frac{\varepsilon}{2}\int_A[\overline{K}^{ik}\overline{v}_k\overline{v}_i\mu]_{\xi=-\zeta}dA, \tag{3.22}$$

making the substitutions of the expansions (3.1)–(3.4) into (3.22), using (3.8), (3.11) and choosing $\delta\overline{v}_i = \langle\overset{2}{\overline{v}}_i\rangle$ in (3.16) and invoking the mode uniqueness condition, we find the following simplified expression for $\mathcal{F}$:

$$\mathcal{F} = \frac{\varepsilon}{2}\int_A\left\{\int_{-\zeta}^{\zeta}[(\overset{0}{\overline{\mathcal{P}}}{}^{i\beta k\delta} + \varepsilon\overset{1}{\overline{\mathcal{P}}}{}^{i\beta k\delta} + \varepsilon^2\overset{2}{\overline{\mathcal{P}}}{}^{i\beta k\delta})(\overset{0}{\overline{v}}_{k,\delta} + \varepsilon\overset{1}{\overline{v}}_{k,\delta})\right.$$
$$\times(\overset{0}{\overline{v}}_{i,\beta} + \varepsilon\overset{1}{\overline{v}}_{i,\beta})(\overset{0}{\mu} + \varepsilon\overset{1}{\mu} + \varepsilon^2\overset{2}{\mu})]d\xi + [(\overset{0}{\overline{K}}{}^{ik} + \varepsilon\overset{1}{\overline{K}}{}^{ik} + \varepsilon^2\overset{2}{\overline{K}}{}^{ik})$$
$$\times(\overset{0}{\overline{v}}_k + \varepsilon\overset{1}{\overline{v}}_k)(\overset{0}{\overline{v}}_i + \varepsilon\overset{1}{\overline{v}}_i)(\overset{0}{\mu} + \varepsilon\overset{1}{\mu} + \varepsilon^2\overset{2}{\mu})]_{\xi=-\zeta}$$
$$+ \varepsilon^2\int_{-\zeta}^{\zeta}[2f(\xi)\overset{0}{\overline{K}}{}^{im}(\overset{0}{\overline{L}}{}^{n3m3})^{-1}[(\overset{1}{\overline{L}}{}^{n3k\delta} + \overset{0}{\overline{L}}{}^{n3k\gamma}\overset{1}{q}{}^{\delta}_{\gamma}$$
$$- \overset{1}{\overline{L}}{}^{n3p3}(\overset{0}{\overline{L}}{}^{q3p3})^{-1}\overset{0}{\overline{L}}{}^{q3k\delta})\overset{0}{\overline{v}}_{k,\delta} + \overset{0}{\overline{L}}{}^{n3k\delta}\overset{1}{\overline{v}}_{k,\delta}]\overset{0}{\overline{v}}_i$$
$$\left. + (f(\xi))^2\overset{0}{\overline{K}}{}^{im}(\overset{0}{\overline{L}}{}^{n3m3})^{-1}\overset{0}{\overline{K}}{}^{nk}\overset{0}{\overline{v}}_k\overset{0}{\overline{v}}_i]d\xi\right\}dA + O(\varepsilon^4). \tag{3.23}$$

Examining the $\varepsilon^3$ contact terms in (3.23) it is recalled that for adequately large values of the foundation stiffness, the mode $\overset{0}{\overline{v}}_i = O(\varepsilon)$ in the areas of contact. This makes the second integral though the thickness $O(\varepsilon^4)$. In metal forming applications, the tooling is assumed to be rigid, and thus the foundation stiffness will be very large, and an expression for the stability functional accurate to $O(\varepsilon^4)$ is found:

$$\mathcal{F} = \frac{1}{2}\int_A\left\{\int_{-h/2}^{h/2}[\overline{\mathcal{P}}^{i\beta k\delta}\hat{v}_{k,\delta}\hat{v}_{i,\beta}\mu]d\theta^3 + [\overline{K}^{ik}\hat{v}_k\hat{v}_i]_{\xi=-\zeta}\right\}dA \tag{3.24}$$

where the mode $\hat{v}_i$ is given by

$$\hat{v}_i(\theta^i)=\overset{0}{\overline{v}}_i(\theta^\alpha)+\langle\overset{1}{\overline{v}}_i\rangle(\theta^\alpha)$$
$$-\theta^3(\overline{L}^{m3i3})^{-1}\overline{L}^{m3k\delta}\left[\frac{\partial\overset{0}{\overline{v}}_k(\theta^\alpha)}{\partial\theta^\delta}-t^r_{k\delta}\overset{0}{\overline{v}}_r(\theta^\alpha)\right], \tag{3.25}$$

and the plane stress incremental moduli, $\overline{\mathcal{P}}^{i\beta k\delta}$, and the contact terms, $\overline{K}^{ik}$, can be found to the accuracy of the prestressed solution.

For puckering problems in metal forming operations, the mode of the instability is a bending mode. When $\beta=0$, and assuming $\langle\overset{1}{\overline{\mathcal{P}}}{}^{i\beta k\delta}\rangle\approx 0$, a reasonable assumption in metal forming problems, it can be shown that $\langle\overset{1}{\overline{v}}_i\rangle\approx 0$, i.e., the mode is a bending mode. When this is the case, the $\hat{v}_i$ only depends on $\overset{0}{\overline{v}}_i$ and the expression for the stability functional simplifies further.

In practice, the positive definiteness of (3.24) is determined numerically. It is assumed that the stress state, and therefore the incremental moduli for the material, are known at any point of the loading path. For the sheet metal forming applications of interest, the load parameter $\lambda$ is taken to be the punch displacement $H$. A stamping process is stable against puckering at a height $H$, if for all punch displacements between 0 and $H$ the functional (3.24) is positive definite. An application of the general theory for the case of the hemispherical cup test is discussed in detail in Part II.

## 4 Conclusions

The goal of the present paper is to present a general and consistent methodology to model puckering in sheet metal forming processes. The starting point is Hill's three-dimensional stability functional for rate-independent solids, appropriately modified for contact. Positive definiteness of this functional, i.e., a positive minimum eigenvalue $\beta$, ensures the stability of the corresponding prestressed elastoplastic solid, while the onset of buckling corresponds to a vanishing $\beta$. The slenderness of the solid, i.e., the dimensionless thickness parameter $\varepsilon\ll 1$, permits using a multiple scale asymptotic method, the construction of a two-dimensional stability functional which is accurate up to $O(\varepsilon^4)$ and which takes into account bending effects. The advantage here lies in the avoidance of any shell theory type approximation. The result is a consistently derived stability functional which is defined on the middle surface of the sheet.

The method is meant to be employed with the finite element discretization of the sheet forming problem of interest and has several advantages. It can use any equilibrium prestress to check for the stability of the corresponding deformed state. It is particularly useful when a membrane solution is available, as in the case of tearing calculations, where the present methodology also allows a check for puckering. The stability functional is defined independently of the algorithm used for the principal solution, and hence one can selectively refine the mesh in those areas prone to puckering. It also provides $\beta$ which is a measure of the stability of the sheet against puckering. Since the stability functional is symmetric, the criterion for $\beta$ is equivalent to the minimum diagonal entry of **D** in an **LDU** decomposition of the stiffness matrix **K** which results from the finite element discretization of the stability functional. The application to puckering experiments for the hemispherical cup test are presented in Part II.


## Acknowledgments

The partial support of ALCOA and NSF under grant G-CMS-9503956 are gratefully acknowledged. The encouragement and support of E. Chu (ALCOA) and R. Stevenson (GM) during the course of this work are sincerely appreciated.


## References


[1] Devons, J. D., 1941, *The Metallurgy of Deep Drawing and Pressing*, Chapman Hall, London.

[2] Taylor, L., Cao, J., Karafillis, A. P., and Boyce, M. C., 1995, ''Numerical Simulations of Sheet-Metal Forming,'' J. Mater. Process. Technol., **50**, pp. 168–179.

[3] Neale, K. W., and Tugcu, P., 1990, ''A Numerical Analysis of Wrinkle Formation Tendencies in Sheet Metals,'' Int. J. Numer. Methods Eng., **30**, pp. 1595–1608.

[4] Triantafyllids, N., and Kwon, Y. J., 1987, ''Thickness Effects on the Stability of Thin Walled Structures,'' J. Mech. Phys. Solids, **35**, pp. 643–674.

[5] Scherzinger, W. M., and Triantafyllids, N., 1998, ''Asymptotic Analysis of Stability for Prismatic Solids Under Axial Loads,'' J. Mech. Phys. Solids, **46**, pp. 955–1007.

[6] Hill, R., 1957, ''On Uniqueness and Stability in the Theory of Finite Elastic Strain,'' J. Mech. Phys. Solids, **5**, pp. 229–241.

[7] Hill, R., 1958, ''A General Theory of Uniqueness and Stability in Elastic-Plastic Solids,'' J. Mech. Phys. Solids, **6**, pp. 236–249.

[8] Destuynder, P., 1980, ''*Application de la Méthode des Développements Asymptotiques à la Théorie des Plaques et des Coques*,'' Doctoral Dissertation, Universityé Pierre et Marie Curie, Paris.

[9] Goetz, A., 1970, *Introduction to Differential Geometry*, Addison-Wesley, New York.

**W. Scherzinger**
Sandia National Laboratories,
P.O. Box 5800, MS 0847,
Albuquerque, NM 87185
Mem. ASME

**E. Chu**
ALCOA Inc.,
Alcoa Technical Center,
Alcoa Center, PA 15069-0001
Mem. ASME

**N. Triantafyllidis**
Department of Aerospace Engineering,
The University of Michigan,
Ann Arbor, MI 48109-2140
Fellow ASME

# Asymptotic Stability Analysis for Sheet Metal Forming—Part II: Application

*The general asymptotic method proposed in Part I is applied to the investigation of the puckering instability in the hemispherical cup test. Both a membrane and a shell stress state for the principal axisymmetric solution have been considered. Due to the possibility of strong deviations from proportional loading in some cases, a corner theory of plasticity is also employed, in addition to the standard $J_2$ deformation theory which is often used in plastic buckling calculations. Results are compared to a previous experimental investigation for brass specimens.* [S0021-8936(00)00904-1]

## 1 Introduction and Motivation

As an application of the general methodology of Part I, we study the puckering instability of the hemispherical cup test. The reasons for the choice of this problem are threefold: (a) relative analytical simplicity, (b) easy calculation of membrane and shell prebifurcation solutions, allowing asymptotic method to be applied to two different prestressed states, and (c) it is an interesting problem with reliable experimental data and a clear asymmetric bifurcation away from the axisymmetric principal solution.

Although our goal is not an exhaustive, in-depth study of puckering per se, a few comments are in order concerning the hemispherical cup test. In addition to its practical interest as a benchmark test for the appearance of wrinkles around corners in sheet metal forming (see Devons [1]), the problem possesses a very interesting theoretical aspect (see Triantafyllidis [2]): unlike most plastic buckling problems that are studied, which have a proportional loading path (beams, plates, cylinders), the hemispherical cup test has large deviations from proportional loading. This feature has implications concerning the constitutive model that is used. The goal here is to show the validity of the asymptotic methodology presented in Part I and to compare the results found using a membrane theory prestress and a shell theory prestress. In addition we investigate, in a systematic fashion, the influence of the constitutive theory and the friction boundary conditions.

The presentation of the work begins with the formulation of the problem in Section 2. The kinematics for the axisymmetric principal solution are given first, followed by the stability functional for the hemispherical cup test as deduced from the general theory of Part I. Finally, a brief discussion of numerical considerations concludes the problem formulation. In Section 3 numerical results based on the above derived stability functional are compared to the experimental data of Donoghue et al. [3] for the hemispherical cup test performed on circular brass disks. The effect of sheet thickness on the loading history is examined, and as a result, the effect of the choice of the constitutive theory on the critical height is also considered. Finally, the influence of the boundary conditions, arising from the friction between the tooling surfaces and the sheet is also investigated.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Sept. 24 1999; final revision, Jan. 30 2000. Associate Technical Editor: S. Kyriakides. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

## 2 Problem Formulation

The application of the general methodology of Part I to the hemispherical cup puckering test requires the solution of the axisymmetric sheet punching problem. For this solution, two cases are considered: a membrane theory and a shell theory prestress. The membrane theory is an approximate solution, while the shell theory (see Triantafyllidis and Samanta [4]) is a consistent finite rotation and finite strain shell theory with a more accurate prestress.

Following the presentation of the kinematics and of the bifurcation functional for the problem at hand, there is a brief discussion of numerical considerations for the finite element discretization of the problem and the calculation of the critical punch height.

As mentioned in the Introduction, the choice of constitutive response is very important for plastic buckling problems. Due to nonproportional loading of the principal solution, the present stability calculations use the corner theory of Christoffersen and Hutchinson [5]. This theory gives the flexibility of investigating a range of constitutive theories from deformation theory to flow theory.

**2.1 Kinematics.** The geometry of the hemispherical cup test is shown in Fig. 1 while the kinematics for the undeformed and deformed configurations are shown in Fig. 2. For an axisymmetric sheet the convected coordinates $\theta^i$ are used, where $\theta^1=s$, $\theta^2=\theta$, and $\theta^3=z$. The meridian for the surface of revolution that describes the midsurface of the sheet in the undeformed configuration is depicted by the line $\mathcal{C}$ in Fig. 2. A material point on the midsurface is given by its position vector $\mathbf{R}$, and an arbitrary material point, at distance $z$ from the midsurface, is given by $\mathbf{P}$

$$P(s,\theta,z)=\mathbf{R}(s,\theta)+z\mathbf{N}(s,\theta), \quad -\frac{h}{2}\leqslant z\leqslant\frac{h}{2}, \tag{2.1}$$

where $\mathbf{N}$ is the unit normal to the midsurface at $\mathbf{R}$. During the deformation of the initially flat sheet, a point on the midsurface initially at $\mathbf{R}$, undergoes a displacement $\mathbf{u}$ and its new position on the deformed midsurface, depicted by curve $\bar{\mathcal{C}}$ in Fig. 2, is given by $\mathbf{r}=\mathbf{R}+\mathbf{u}$. The position vector for an arbitrary material point, at distance $\bar{z}$ from the midsurface, is given by $\mathbf{p}$,

 Copyright © 2000 by ASME 

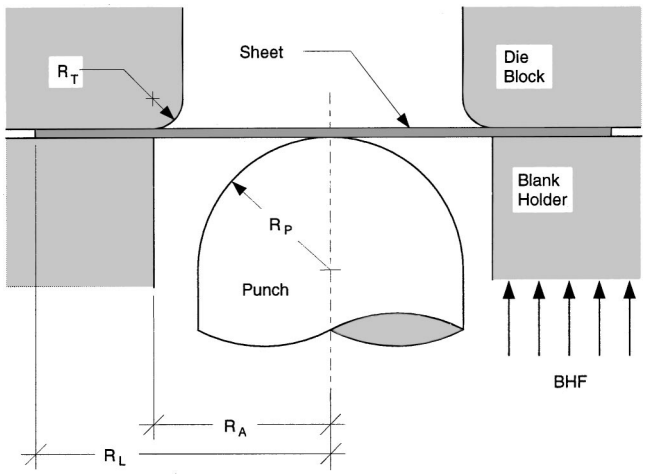



**Fig. 1 Schematic drawing for hemispherical cup test**

$$\mathbf{p}(s,\theta,z)=\mathbf{r}(s,\theta)+\bar{z}\mathbf{n}(s,\theta),\quad -\frac{\bar{h}(s)}{2}\leq\bar{z}\leq\frac{\bar{h}(s)}{2},\tag{2.2}$$

where $\mathbf{n}$ is the unit normal to the current midsurface at $\mathbf{r}$ and $\bar{h}(s)$ is the current sheet thickness. The displacement, $\mathbf{u}$, has components $v(s)$ and $w(s)$ along the radial and axial directions, respectively, i.e.,

$$\mathbf{u}(s,\theta)=v(s)\mathbf{e}_r(\theta)+w(s)\mathbf{e}_3;\quad \mathbf{e}_r(\theta)=\cos\theta\mathbf{e}_1+\sin\theta\mathbf{e}_2.\tag{2.3}$$

The basis vectors for the midsurface in the current configuration, $\mathbf{a}_i$, expressed in terms of the local orthonormal basis $\{\mathbf{t},\mathbf{n},\mathbf{b}\}$, are

$$\mathbf{a}_1=\frac{\partial\mathbf{r}}{\partial s}=\lambda_s\mathbf{t},\quad \mathbf{a}_2=\frac{\partial\mathbf{r}}{\partial\theta}=\rho\mathbf{b},\quad \mathbf{a}_3=\mathbf{n},\tag{2.4}$$

where $\mathbf{t}$ is the unit tangent to the curve $\bar{\mathcal{C}}$, $\lambda_s$ is the radial stretch ratio and $\rho$ is the current distance from the axis of symmetry, which are found with the help of (2.1)–(2.4) to be

$$\lambda_s=\frac{d\tilde{s}}{ds}=\left[\left(1+\frac{dv}{ds}\right)^2+\left(\frac{dw}{ds}\right)^2\right]^{1/2},\quad \rho=s+v(s).\tag{2.5}$$

From differential geometry (e.g., Goetz [6]), the nonzero components of the curvature tensor for the axisymmetric midsurface are

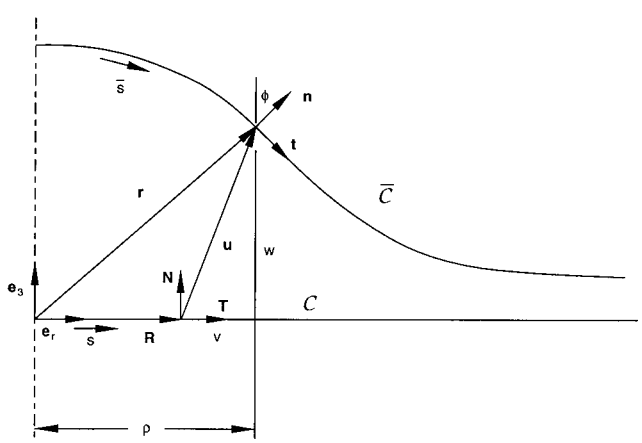



**Fig. 2 Kinematics for axisymmetric shell and membrane theories. $\mathcal{C}$ is the meridian for the sheet in the undeformed (reference) configuration, and $\bar{\mathcal{C}}$ is the meridian for the sheet in the deformed (current) configuration.**

$$b_1^1=-k_s=\frac{1}{\lambda_s^2}\left(\sin\phi\frac{d^2v}{ds^2}+\cos\phi\frac{d^2w}{ds^2}\right),\quad b_2^2=-k_\theta=-\frac{\sin\phi}{\rho}\tag{2.6}$$

where $\cos\phi=\mathbf{N}\cdot\mathbf{n}$. The nonzero Christoffel symbols for the current midsurface are found from Eq. (2.5) in Part I

$$t_{11}^1=\frac{1}{\lambda_s}\frac{d\lambda_s}{ds},\quad t_{12}^2=t_{21}^2=\frac{1}{\rho}\frac{d\rho}{ds},\quad t_{22}^1=-\frac{\rho}{\lambda_s^2}\frac{d\rho}{ds}.\tag{2.7}$$

With these kinematic quantities, the stability functional presented in Part I can be evaluated for the axisymmetric hemispherical cup test.

**2.2 Stability Functional.** Using the results from Part I, and representing tensor quantities with their physical components,[1] the stability functional for an axisymmetric shell is according to Eq. (3.24) in Part I

$$\mathcal{F}(\overset{0}{\bar{v}}_{\langle i\rangle};H)=\frac{1}{2}\int_A\left\{\int_{-h/2}^{h/2}[\bar{\mathcal{P}}^{\langle i\beta k\delta\rangle}\hat{v}_{\langle k,\delta\rangle}\hat{v}_{\langle i,\beta\rangle}\mu]dz+[\hat{K}^{\langle ik\rangle}\hat{v}_{\langle k\rangle}\hat{v}_{\langle i\rangle}]_{\theta^3=-h/2}\right\}dA,\tag{2.8}$$

where $\hat{v}_{\langle i\rangle}$ is deduced from Eq. (3.25) of Part I under the assumption $\langle\overset{1}{\bar{\mathcal{P}}}_{i\beta k\delta}\rangle\approx 0$,

$$\hat{v}_{\langle i\rangle}=\overset{0}{\bar{v}}_{\langle i\rangle}-\theta^3(L^{\langle m3i3\rangle})^{-1}L^{\langle m3k\delta\rangle}\left[\frac{\partial\overset{0}{\bar{v}}_{\langle k\rangle}}{\partial\theta^\delta}-t_{k\delta}^r\overset{0}{\bar{v}}_{\langle r\rangle}\right],\tag{2.9}$$

and the incremental moduli, appearing in (2.9) (and in the expression for the plane stress moduli $\mathcal{P}$), are given in Triantafyllidis [2]. For simplicity, we introduce the notation $\overset{0}{\bar{v}}_{\langle 1\rangle}\equiv\bar{u}$, $\overset{0}{\bar{v}}_{\langle 2\rangle}\equiv\bar{v}$ and $\overset{0}{\bar{v}}_{\langle 3\rangle}\equiv\bar{w}$. Using the kinematics from Section 2.1 in conjunction with (2.9), the plane stress assumption and the axisymmetry of the stress state, the strain rate terms $\hat{v}_{\langle i,j\rangle}$ in (2.8) are found to be

$$\hat{v}_{\langle 1,1\rangle}=\frac{1}{\lambda_s(1+k_s\bar{z})}\left\{\frac{\partial\bar{u}}{\partial s}-\lambda_s k_s\bar{w}-\bar{z}\left(\frac{\partial\alpha}{\partial s}\right)\right\}\tag{2.10}$$

$$\hat{v}_{\langle 1,2\rangle}=\frac{1}{\rho(1+k_\theta\bar{z})}\left\{\frac{\partial\bar{u}}{\partial\theta}-\bar{v}\cos\phi-\bar{z}\left(\frac{\partial\alpha}{\partial\theta}-\beta\cos\phi\right)\right\}\tag{2.11}$$

$$\hat{v}_{\langle 2,1\rangle}=\frac{1}{\lambda_s(1+k_s\bar{z})}\left\{\frac{\partial\bar{v}}{\partial s}-\bar{z}\left(\frac{\partial\beta}{\partial s}\right)\right\}\tag{2.12}$$

$$\hat{v}_{\langle 2,2\rangle}=\frac{1}{\rho(1+k_\theta\bar{z})}\left\{\frac{\partial\bar{v}}{\partial\theta}+\bar{u}\cos\phi+\bar{w}\sin\phi-\bar{z}\left(\frac{\partial\beta}{\partial\theta}+\alpha\cos\phi\right)\right\}\tag{2.13}$$

$$\hat{v}_{\langle 3,1\rangle}=\alpha;\quad \hat{v}_{\langle 3,2\rangle}=\beta\tag{2.14}$$

where for simplicity, the following auxiliary quantities $\alpha$ and $\beta$ have been introduced:

$$\alpha\equiv\frac{1}{\lambda_s}\frac{\partial\bar{w}}{\partial s}-k_s\bar{u},\quad \beta\equiv\frac{1}{\rho}\frac{\partial\bar{w}}{\partial\theta}-k_\theta\bar{v}.\tag{2.15}$$

Finally, it is noted that the stability functional in (2.8) can be used with a membrane or shell prestressed state. A membrane prestress state, one with no bending stiffness, gives a stress field that is constant through the thickness of the sheet, resulting in plane stress incremental moduli that depend solely on the midsurface coordinate $s$. On the other hand, a shell prestress state gives a

[1]The physical components of a tensor are denoted with the subscripts or superscripts of the tensor enclosed in brackets $\langle\cdot\rangle$.

varying stress state through the thickness of the sheet, in which case the plane stress incremental moduli are integrated through the sheet thickness.

**2.3 Numerical Considerations.** The methodology and algorithm for calculating the axisymmetric principal solutions is presented in detail in Triantafyllidis [2] for the membrane model and Triantafyllidis and Samanta [4] for the shell model, respectively. The same references also contain detailed derivations of the incremental moduli for the deformation and corner theory constitutive models employed here. The determination of the positive definiteness of the stability functional is done by employing the same finite element discretization used for the principal solution. The axisymmetry of the principal solution, suggests a Fourier decomposition for the eigenmode

$$\overline{u}(s,\theta)=\sum_{n=0}^{\infty}\,[u_n^*(s)\cos(n\theta)+u_n(s)\sin(n\theta)], \qquad (2.16)$$

$$\overline{v}(s,\theta)=\sum_{n=0}^{\infty}\,[v_n(s)\cos(n\theta)+v_n^*(s)\sin(n\theta)], \qquad (2.17)$$

$$\overline{w}(s,\theta)=\sum_{n=0}^{\infty}\,[w_n^*(s)\cos(n\theta)+w_n(s)\sin(n\theta)], \qquad (2.18)$$

where $n\geqslant 1$ is the eigenmode's wave number. Using (2.16)–(2.18) and exploiting the orthogonality of the trigonometric functions, the stability functional can be written as

$$\mathcal{F}(\overline{u},\overline{v},\overline{w};H)=\sum_{n=1}^{\infty}\,[\mathcal{F}_n(u_n(s),v_n(s),w_n(s)) + \mathcal{F}_n(u_n^*(s),v_n^*(s),w_n^*(s))], \qquad (2.19)$$

where $\mathcal{F}_n$ are quadratic functionals of the three $s$-dependent parts of the eigenmode $(u_n(s),v_n(s),w_n(s)$, or equivalently of $u_n^*(s),v_n^*(s),w_n^*(s))$ and the wave number $n$. By inspection of (2.19), it is seen that loss of positive definiteness of some $\mathcal{F}_n$, results in loss of positive definiteness of $\mathcal{F}$. Hence the stability analysis of the two-dimensional puckering problem is reduced to the determination of positive definiteness of a one-dimensional stability functional over a given range of wave numbers $n$, thus greatly simplifying the analysis.

For assembling the stability functional $\mathcal{F}_n$, we use the same mesh as for the principal axisymmetric solution but with a linear interpolation function for $\overline{u},\overline{v}$ and a Hermitian cubic interpolation for $\overline{w}$. A Cholesky decomposition of the resulting symmetric stiffness matrix $\mathbf{K}_n$ is used, i.e., $\mathbf{K}_n=\mathbf{L}_n\mathbf{D}_n\mathbf{L}_n^T$. For a given punch height, the stability of the structure is examined by looking at the sign of the minimum entry in diagonal matrix $\mathbf{D}_n$ over a large range of integers $n$ (typically from 1 to 25). The critical height $H_c$ is the lowest punch displacement at which this entry becomes negative for some integer $n$.

## 3 Results

The goal of this section is to compare the predictions of the theory derived in Section 2 with experimental results and also to examine the accuracy of the membrane prestress simplification, as opposed to the shell theory prestress, for the prediction of the critical punch height. The experimental results used in our comparisons were performed by Donoghue et al. [3] where circular brass blanks of four different thicknesses, $h$, of 0.508 mm, 0.635 mm, 0.813 mm, and 1.016 mm, were tested. In addition to varying the sheet thickness, the tooling size and draw ratio $(R_L/R_A)$ were also varied. Results will be presented here for two tooling sizes, $S_1$ and $S_2$, (shown in Table 1) and for two draw ratios $(R_L/R_A=1.7$ and $R_L/R_A=1.5)$.

**Table 1 Tooling sizes used in Donoghue et al. [3]**

| | $S_1$ | $S_2$ |
|---|---|---|
| $R_P$ | 50.00 mm | 33.33 mm |
| $R_A$ | 58.33 mm | 38.87 mm |
| $R_T$ | 10.71 mm | 6.94 mm |

Experimental data for the brass uniaxial stress-strain curves closely follow a bilinear hardening model. The initial elastic modulus is $E=110$ GPa. For the three thinest sheets the yield stress is $\sigma_y=195$ MPa and the tangent modulus is $E_t=812$ MPa, while for the thickest sheet, $h=1.016$ mm, the yield stress is $\sigma_y=120$ MPa, and the tangent modulus is $E_t=852$ MPa. For the brass experiments there was no measured plastic anisotropy in the sheets.

Considerable effort was devoted to removing the effects of friction in the tests reported by Donoghue et al. [3] by using teflon sheets between the blank and the tooling, and the friction coefficient between the tooling and the sheet, while not removed entirely, was reduced considerably. The value $\mu=0.04$ used in the simulations reported in Donoghue et al. [3], has also been used for all numerical calculations in this work.

To constrain wrinkling in the flange, a blank holder force (BHF) was applied to the sheet between the blank holder and the die block. Table 2 shows the blank holder forces for all of the experimental cases examined here.

For all comparisons between the experimental and theoretical results, $J_2$ deformation theory of plasticity is used, which gives a lower bound on the critical punch height. The influence of the constitutive response of the material is examined later. Note that all subsequent comparisons involve the critical height, which is very sensitive to the constitutive description employed. The critical wave number, being determined mainly by geometry, is always in good agreement with experimental results, as discussed by Donoghue et al. [3] and hence is of no further concern.

For the first set of comparisons, the largest tooling $(S_1)$ is used and the draw ratio is $R_L/R_A=1.7$. The results for this case are shown in Table 3. For thin sheets the theoretical predictions using both the shell and membrane theories, show good agreement with the experimental results. The percent error for the 0.508 mm sheet is 2.2 percent using the shell theory and 13.7 percent with the membrane theory. As the sheet thickness increases the theoretical predictions using the deformation theory grow progressively worse. For the 1.016-mm sheet, the shell theory result underpredicts the critical punch height by 36.5 percent, while for the membrane theory calculation the error is 48.4 percent. The main reason for this discrepancy lies in the large deviations from proportional loading that occur. For the thickest sheets the deviations from

**Table 2 Blankholder force (BHF) for cases examined in experimental study by Donoghue et al. [3]**

| Tooling | $R_L/R_A$ | Thickness (mm) | BHF (kN) | Tooling | $R_L/R_A$ | Thickness (mm) | BHF (kN) |
|---|---|---|---|---|---|---|---|
| S1 | 1.7 | 0.508 | 34.637 | S1 | 1.5 | 0.635 | 33.373 |
| S1 | 1.7 | 0.635 | 33.809 | S1 | 1.5 | 0.813 | 33.818 |
| S1 | 1.7 | 0.813 | 41.227 | S1 | 1.5 | 1.016 | 41.320 |
| S1 | 1.7 | 1.016 | 41.316 | S2 | 1.7 | 0.508 | 49.570 |
| S1 | 1.5 | 0.508 | 33.168 | S2 | 1.7 | 0.635 | 49.837 |

**Table 3 Experimental and theoretical results using deformation theory with $R_L/R_A=1.7$ and the $S_1$ tooling**

| Thickness | 0.508 mm | 0.635 mm | 0.813 mm | 1.016 mm |
|---|---|---|---|---|
| $H_c$ (experimental) | 22.86 mm | 29.46 mm | 36.38 mm | 54.36 mm |
| $H_c$ (shell theory) | 22.35 mm | 25.15 mm | 30.81 mm | 34.54 mm |
| $H_c$ (membrane theory) | 19.72 mm | 21.15 mm | 24.64 mm | 28.05 mm |

**Table 4 Experimental and theoretical results using deformation theory with $R_L/R_A=1.5$ and the $S_1$ tooling**

| Thickness | 0.508 mm | 0.635 mm | 0.813 mm | 1.016 mm |
|---|---|---|---|---|
| $H_c$ (experimental) | 13.46 mm | 18.03 mm | 24.38 mm | 41.91 mm |
| $H_c$ (shell theory) | 13.72 mm | 15.81 mm | 19.02 mm | 24.18 mm |
| $H_c$ (membrane theory) | 13.37 mm | 15.37 mm | 18.06 mm | 22.15 mm |

proportional loading are larger, especially for material points away from the midsurface of the sheet. As a result the deformation theory of plasticity will not be as accurate for the thicker sheets. This point will be addressed subsequently.

Next attention is turned to experiments run with the same tooling as before ($S_1$), but with a smaller draw ratio of $R_L/R_A=1.5$. The experimental and theoretical results for this case are shown in Table 4. It is seen, as in the previous case, that the predictions for thin sheets, using $J_2$ deformation theory, are very good, while those for thick sheets are not as accurate. For the 0.508-mm sheet the percent error, using both the shell and membrane theories, is under 2.0 percent. But as the sheet thickness increases, once again the error grows quite large. For the 1.016-mm sheet the percent error is 42.3 percent for the shell theory and 47.1 percent for the membrane theory.

Results are also examined for tests run with a smaller tooling, $S_2$, which gives higher curvatures, and a draw ratio of $R_L/R_A=1.7$. These results are shown in Table 5. Here it is seen, as in the previous cases, that the shell theory performs better than the membrane theory when estimating the critical punch height. The larger curvatures in this problem, especially near the die throat, necessitate a shell theory. (Experiments using the 0.813-mm and 1.016-mm sheets were not performed by Donoghue et al. [3].)

The role of the constitutive theory in plastic buckling problems is extremely important and is often the major reason for the discrepancy between theoretical buckling calculations and experimental observations (see review article by Hutchinson [7]). For most commonly studied problems, the prebuckling solution has proportional loading, and deformation theory of plasticity will predict the instability. However, in the problem examined here, large deviations from proportional loading occur, especially for the thickest sheets. Deformation theory is inadequate when large deviations from proportional loading occur, and as a result, it is not expected to accurately predict instabilities for thicker sheets.

To examine the effect of sheet thickness on the deviation from proportional loading, we follow the stress histories for material points in the unsupported region of the cup, which lie initially near the die throat. This choice is dictated by the fact that the stress distribution in the unsupported region, has the greatest effect on the stability calculations. Presented here are the stress histories for the thinest sheet, 0.508 mm, and the thickest sheet, 1.016 mm, using the largest tooling and a draw ratio of $R_L/R_A=1.7$. The stress history for material points in the 0.508-mm sheet, at an initial radius of 48.12 mm are shown in Fig. 3. This point in the sheet has a location in the unsupported region of the sheet when the punch reaches $H_c$ and the resultant hoop force at this point is compressive. The yield surface is plotted in $\sigma_{11}-\sigma_{22}$ space. Material points on the top, middle, and bottom surface load through the elastic range, reach the yield surface, and continue loading in the plastic range. For the middle surface there is not a strong deviation from proportional loading after yield. The top and bottom surfaces, however, do show strong deviations from proportional loading after yielding has occurred. Notice, however, that the corresponding stresses are just out of the initial yield surface.

These results are compared to the stress histories for the thickest sheet, $h=1.016$ mm, shown in Fig. 4. Again the stress histories are plotted for a material point at an initial radius of 48.12 mm. Here large deviations from proportional loading are seen throughout the sheet and the stresses are now deeply into the plastic range. The stresses in the sheet are redistributed after yielding in order to sustain equilibrium, and this stress redistribution is greater if the sheet is thicker, resulting in larger deviations from proportional loading. Nearly proportional loading is necessary for the accuracy of a deformation theory, and as a result, the plasticity theory chosen to model the behavior of the thinner sheets, i.e., deformation theory, is not necessarily valid for the thicker sheets. To improve on the deformation theory, a corner theory is employed to predict buckling.

The effect of sheet thickness on the critical punch height for

**Table 5 Experimental and theoretical results using deformation theory with $R_L/R_A=1.5$ and the $S_2$ tooling**

| Thickness | 0.508 mm | 0.635 mm |
|---|---|---|
| $H_c$ (experimental) | 34.80 mm | 38.86 mm |
| $H_c$ (shell theory) | 30.18 mm | 31.32 mm |
| $H_c$ (membrane theory) | 21.09 mm | 22.23 mm |

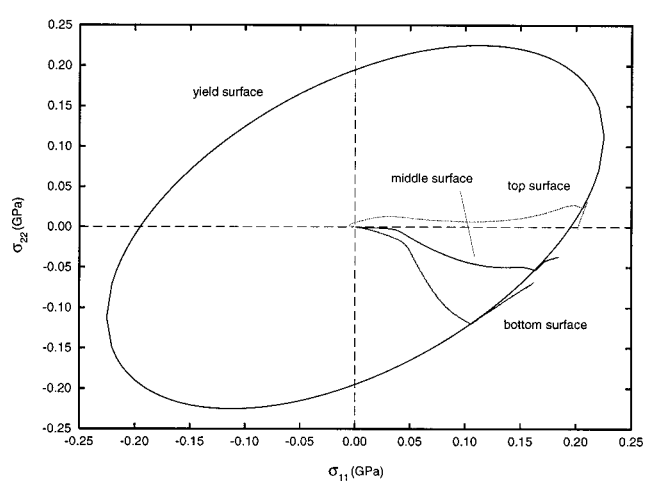



**Fig. 3 Stress history for material points initially at a radius of 48.12 mm for the 0.508 mm sheet. Stress history is shown for the bottom, middle, and top surfaces of the shell.**

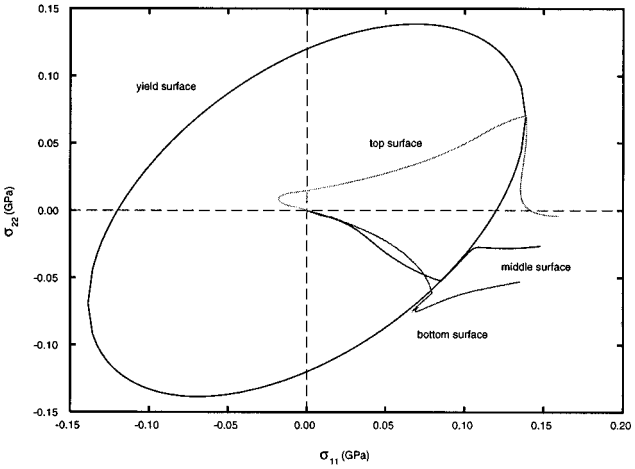



**Fig. 4 Stress history for material points initially at a radius of 48.12 mm for the 1.016 mm sheet. Stress history is shown for the bottom, middle and top surfaces of the shell.**

various corner theories can be seen in Fig. 5. The coefficient of friction for all the cases shown is 0.04 and the BHF is 41.316 kN. These are the same boundary conditions for the case of the 1.016-mm sheet with the $S_1$ tooling and the draw ratio of $R_L/R_A=1.7$. It is recalled that this case had the largest discrepancy between the experimental and theoretical results. The first two curves, (a) and (b), show the results found using the deformation theory of plasticity with a membrane and shell theory, respectively. Here it is seen that the bending terms and thickness distribution of the stress, both ignored in the membrane formulation, have the effect of raising the critical punch height. Altering the response of the material by changing the corner characteristics can further increase the critical punch height. Curves (c), (d), and (e) show the effects of different corner theories with no forward loading cone, $\theta_0=0$. As the unloading cone angle, $\theta_c$, approaches

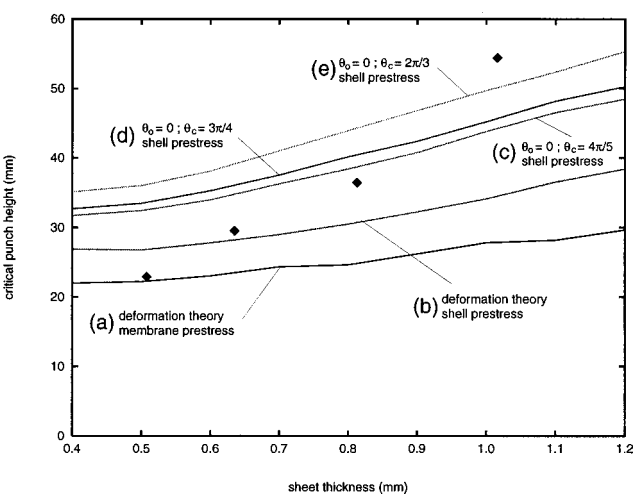



**Fig. 5 Effect of sheet thickness on the critical punch height for a brass sheet with $R_L/R_A=1.7$, the $S_1$ tooling, BHF $=41.316$ kN a coefficient of friction $\mu=0.04$. The membrane prestress underpredicts the shell theory results for the same constitutive relations. Using a corner theory, and varying its corner geometry so as to approach a flow theory, the critical punch height increases. It is noted that the BHF used for all the numerical calculations is the experimental value corresponding to the thickest blank. Experimental results are marked by (♦).**

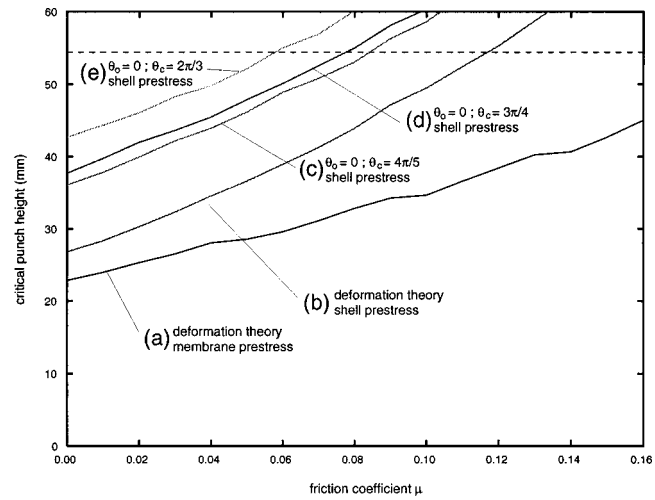



**Fig. 6 Effect of friction on the critical punch height for 1.016 mm brass sheet with $R_L/R_A=1.7$, the $S_1$ tooling and a BHF $=41.316$ kN. The trend in the results is the same as that seen in Fig. 5, and it shows the great dependence of the results on the friction boundary conditions, which are often difficult to measure. The experimental result for this case is shown by the dashed line.**

$\pi/2$, the critical punch height increases. Flow theory is obtained when $\theta_c=\pi/2$. This case is examined, but no puckering instability is found regardless of the sheet thickness.

Finally, the sensitivity of the results to the boundary conditions are examined by varying the coefficient of friction between the tooling surfaces and the sheet. The friction, combined with the BHF, supplies a restraining force on the sheet that prevents material from flowing into the die. The amount of material that flows into the die determines the compressive hoop stresses that cause puckering. Therefore, the coefficient of friction plays an important role in determining the critical punch height, as can be seen in Fig. 6. Once again the cases shown in Fig. 6 are variations of the case for the 1.016-mm sheet with the $S_1$ tooling and the draw ratio of $R_L/R_A=1.7$. The blank holder force for all the cases is 41.316 kN and the coefficient of friction is varied from 0.00 to 0.16. The effect of the coefficient of friction is greater for the shell theories (curves (b)–(e)) than for the membrane theory (curve (a)) as evidenced by the lower slope for the membrane theory curve. The effects for the different plasticity theories is the same for the shell theories since they have similar slopes. Looking at the region near the experimental results, $\mu=0.04$, it is seen that a slight change, say about 0.02, in the coefficient of friction can result in a change of up to 5 mm in the critical punch height. Therefore, it is determined that the boundary conditions (the coefficient of friction and BHF) have a very large effect on stability in the hemispherical cup test.

## 4 Conclusions

The general method proposed in Part I is applied to the modeling of the hemispherical cup test. The goal is to check how the method performs for different sheet thicknesses and to investigate the influence of simpler (membrane) prestress states versus more accurate (shell) prestress states.

As expected, better agreement with experimental results is obtained for thinner sheets using $J_2$ deformation theory of plasticity. As the sheet thickness increases, so does the discrepancy between the experimental results and the theoretical predictions. The main reason for this discrepancy is the presence in the principal solution of large deviations from proportional loading, which requires more sophisticated constitutive models (corner theories) to accurately predict the critical punch height.

In the application at hand, the minimum eigenvalue of the stability functional goes through zero at the onset of bifurcation, since a bifurcation type instability is guaranteed by the axisymmetry of the principal solution. For the general case of a nonaxisymmetric geometry, as for example for the punch test of noncircular blanks or blanks with some initial asymmetric imperfection, recent calculations done by the authors show that the minimum eigenvalue will not change sign at criticality. Consequently the onset of instability is not clearly defined, although it must still be connected to the minimum eigenvalue of the stability functional which measures the stiffness of the structure. Current work (unpublished) examines a nonaxisymmetric blank (a circular blank with a small initial imperfection) and a square blank deformed by a hemispherical punch and correlates the minimum eigenvalue with the initial appearance of surface wrinkles.

## Acknowledgments

The partial support of ALCOA and NSF under grant G-CMS-9503956 are gratefully acknowledged. The encouragement of R. Stevenson (GM) during the course of this work are sincerely appreciated.

## References

[1] Devons, J. D., 1941, *The Metallurgy of Deep Drawing and Pressing*, Chapman Hall, London.
[2] Triantafyllidis, N., 1985, ''Puckering Instability in the Hemispherical Cup Test,'' J. Mech. Phys. Solids, **33**, pp. 117–139.
[3] Donoghue, M., Stevenson, R., Kwon, Y. J., and Triantafyllidis, N., 1989, ''An Experimental Verification of the Hemispherical Cup Puckering Problem,'' ASME J. Eng. Mater. Technol., **111**, pp. 248–254.
[4] Triantafyllidis, N., and Samanta, S. K., 1986, ''Bending Effects on Flow Localization in Metallic Sheets,'' Proc. R. Soc. London, Ser. A, **406**, pp. 205–226.
[5] Christoffersen, J., and Hutchinson, J. W., 1979, ''A Class of Phenomenological Corner Theories of Plasticity,'' J. Mech. Phys. Solids, **27**, pp. 465–487.
[6] Goetz, A., 1970, *Introduction to Differential Geometry*, Addison-Wesley, New York.
[7] Hutchinson, J. W., 1974, ''Plastic Buckling,'' Adv. Appl. Mech., **14**, pp. 67–144.

# Modified Lagrange Method to Analyze Problems of Sliding and Rolling

**A. Rosen**
Faculty of Aerospace Engineering,
Technion–Israel Institute of Technology,
Haifa 32000, Israel
Mem. ASME

*A modified Lagrange method to analyze problems of sliding and rolling is presented. The method is based on modeling the friction as a process of collisions between the sliding and rolling body, and particles of the surface on which it slips and rolls. The process of collisions does not need to describe the exact friction process. Instead it can represent another equivalent mechanism of loss of energy of the body due to sliding and rolling. The function that describes the rate of increase of the kinetic energy of the particles, as a result of the collisions, plays a major roll in the modified Lagrange equations. Cases of isotropic and anisotropic friction can be modeled. Three examples of using the method are presented. It is shown that when an infinitely rough surface is assumed, the classical equations for nonholonomic constraints of rolling without sliding are obtained. Lagrange multipliers that appear in these equations obtain direct physical meaning and the mechanism behind the constraint becomes clear.* [S0021-8936(00)01404-5]

## 1 Introduction

Many kinds of motion include sliding and rolling of a body on a surface. There are many practical examples of such motions that are dealt with in many textbooks. Neimark and Fufaev [1] defined very accurately sliding and rolling and in what follows their definition, that is similar to other definitions in the literature, will be used.

Imagine that a moving body $B$ is constrained to touch a fixed surface $S$ at all times. At each instant of time $t$, a point $O_B$ on body $B$ touches a point $O_S$ of surface $S$. Actually there is a small area of contact between the body and the surface, but when a rigid body and a rigid surface are considered, an ideal point contact is given.

The velocity of point $O_B$ at the instant of time at which it is the point of contact, lies in the common tangent plane to the moving body $B$ and surface $S$, at their point of contact. The velocity $\mathbf{v}_s$ of the point of contact $O_B$ relative to point $O_S$ (see Fig. 1), is called the velocity of sliding of body $B$ with respect to surface $S$. Apart from translational motion of point $O_B$, body $B$ may also rotate about point $O_B$ with an angular velocity $\overline{\omega}$ (see Fig. 1). In the case of a rigid body, the vector $\overline{\omega}$ is the angular velocity of the body with respect to the surface. It is common to express $\overline{\omega}$ by two components: normal to the tangent plane $\overline{\omega}_p$ and a component $\overline{\omega}_s$ parallel to the tangent plane ($\omega_p$ and $\omega_s$ denote the magnitude of these components). $\omega_p$ is called the pivoting velocity, while $\omega_s$ is the pure rolling velocity (thus the rotation of $B$ about the point $O_B$ is comprised of pure rolling and pivoting).

Of special interest is the case of no sliding, $\mathbf{v}_s = \mathbf{0}$. In the literature this case is referred to in many cases as rolling, although as indicated above rolling can also be combined with sliding. In what follows we will refer to this case as a motion without sliding (or zero sliding).

Many cases of motion without sliding of a body on a surface have been investigated in the literature. The demand of zero sliding velocity at the point of contact is expressed mathematically by introducing nonholonomic constraints, namely constraints that include velocities. Solutions of dynamic problems that include nonholonomic constraints of zero sliding, using Lagrange method, appear in many books on analytical mechanics. Neimark and Fufaev [1] have published the most detailed collection of such problems and their solutions.

The author is not aware of examples in the literature of using Lagrange method to analyze problems that include nonzero sliding. The purpose of this paper is to present a method of addressing problems of combined sliding and rolling using the Lagrange approach. This method is based on a physical description of a procedure of energy transfer due to sliding and pivoting. It is assumed that the sliding body collides with particles of the surface, and as a result of these collisions the kinetic energy of these particles is increased. Lagrange equations are extended to include the effect of these collisions and the transfer of energy to these particles. This process is described using a newly defined function that can be looked upon as a special dissipation function for the case of sliding and rolling. The regular Lagrange equations are extended by considering not only body $B$, but also the particles that absorb its energy. Two new kinds of dissipation function are considered that, as shown in the text, represent dry friction and viscous friction. The derivation includes cases of isotropic friction (where the friction force is independent of the sliding direction), as well as cases of anisotropic friction (where the frictional characteristics change while moving in different directions).

It should be emphasized that the collisions mechanism proposed to describe the friction phenomenon is not by any means a description of the actual mechanisms involved in the friction process. The actual mechanisms are very complicated and very difficult to simulate. Instead the present mechanism should be considered as a simplified equivalent mechanism of the actual one. The model of collisions is a conceptual model that in the end, as a

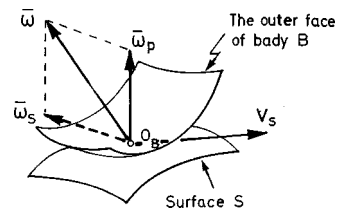



**Fig. 1 Definition of sliding, pivoting, and rolling**

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Nov. 29, 1999; final revision, Mar. 12, 2000. Associate Technical Editor: N. C. Perkins. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

 **Copyright © 2000 by ASME** 

result of using linear or quadratic dissipation functions, reproduces results of dry or viscous friction, respectively. By using other functions, other kinds of friction forces can be modeled.

Three examples will be solved, explained, and discussed. These cases include isotropic and anisotropic friction. Both linear and quadratic dissipation functions will be used. It will be shown that the zero sliding case is obtained from the more general solutions of combined sliding and rolling, if it is assumed that the surface (on which the body moves) exhibits infinite roughness. In such cases the results of the present method converge to those that appear in the literature, that has been obtained using the Lagrange multiplier approach for nonholonomic constraints.

## 2 The Physical Model

The main effect of sliding is the loss of energy of the body $B$, due to friction between it and the surface $S$. Rolling is also associated with a loss of energy, but this loss of energy is usually very small compared to the loss of energy due to sliding. In most of the cases the loss of energy due to pivoting is larger than the loss of energy due to pure rolling. Thus in what follows the loss of energy due to pivoting will be included and the one due to pure rolling will be neglected. If necessary, the small energy loss due to pure rolling can be included following the method that will be presented for the two other mechanisms.

For ease of physical-mechanical understanding, it is assumed that there is a region of surface $S$ that is responsible for the process of friction. This region includes an infinite number of very small identical particles that each has a mass $m_p$ and an initial velocity $v_l$. The friction force that acts on $B$ is a result of collisions between these particles and the contact region of body $B$. Due to a collision of a certain particle, the velocity of this particle is increased to $(v_l+\rho v_s+\zeta a\omega_p)$ where $\rho$ and $\zeta$ are always positive and greater than zero. $a$ is a typical length of body $B$. It is assumed that at a certain moment, $N_p$ particles collide with body $B$ per unit time. Thus the change of the kinetic energy of the particles per unit time, at the moment $t$, $K_f(t)$, is

$$K_f(t)=\frac{1}{2}N_p m_p(v_l+\rho v_s+\zeta a\omega_p)^2-\frac{1}{2}N_p m_p v_l^2$$
$$=N_p m_p v_l\left[\rho v_s+\zeta a\omega_p+\frac{\rho\zeta a v_s\omega_p}{v_l}+\frac{1}{2}\frac{\rho^2 v_s^2}{v_l}+\frac{1}{2}\frac{\zeta^2 a^2\omega_p^2}{v_l}\right]. \tag{1}$$

In the last equation $v_s$ and $\omega_p$ stand for $v_s(t)$ and $\omega_p(t)$, respectively (these are the absolute values that are always positive). In general, $N_p$ is also a function of $v_s$ and $\omega_p$. Thus we can conclude that the increase in the kinetic energy of the particles, per unit time, is

$$K_f(t)=K_f(v_s(t),\omega_p(t)). \tag{2}$$

$K_f$ is a positive definite function of $v_s(t)$ and $\omega_p(t)$.

In what follows two cases will be considered:

(a) $K_f$ is a linear function of $v_s$ and $\omega_p$. For convenience, ease of physical interpretation and comparison with known results from the literature, the linear function is written as follows:

$$K_{f1}(t)=m_B g[\kappa v_s(t)+a\chi\omega_p(t)]. \tag{3}$$

$m_B$ is the mass of body $B$ while $g$ is the gravity acceleration. $\kappa$ and $\chi$ are nondimensional coefficients.

(b) The function $K_f$ is a quadratic function of the sliding and pivoting velocities:

$$K_{f2}(t)=\tfrac{1}{2}[\mu v_s(t)^2+\nu\omega_p(t)^2]. \tag{4}$$

The factor 1/2 was introduced for convenience in the coming derivations. $\mu$ and $\nu$ are dimensional positive coefficients: (N·sec/m) and (N·sec·m), respectively.

There are cases where the friction is not isotropic, namely it varies in different directions. There may be two kinds of anisotropy:

(a) Anisotropy that depends on body $B$. Consider, for example, that body $B$ touches surface $S$ by a small sharp wheel. The friction in a direction that lies in the plane of the wheel is much smaller than the friction in a direction normal to that plane. In this case the behavior depends on the orientation of body $B$, namely the wheel (see the example in Section 4).

(b) Anisotropy that depends on surface $S$. Surface $S$ may have different friction characteristics in different directions. In this case, anisotropy is described relative to surface $S$ (see the example in Section 5).

In this paper, orthotropic cases of anisotropy will be considered where the anisotropy is described by the characteristics in two orthogonal directions. These directions are defined by the unit vectors $\mathbf{e}_\alpha$ and $\mathbf{e}_\beta$. These unit vectors are always perpendicular to each other, they lie in the tangent plane, but their orientation may change with time and it may also be a function of the position or orientation of body $B$ (see examples in what follows). In this case it is convenient to express $\mathbf{v}_s$ by its components in the directions $\mathbf{e}_\alpha$ and $\mathbf{e}_\beta$:

$$\mathbf{v}_s=v_{s\alpha}\mathbf{e}_\alpha+v_{s\beta}\mathbf{e}_\beta. \tag{5}$$

When orthotropy is considered, Eq. (3) is modified as follows:

$$K_{f1}=m_B g(\kappa_\alpha|v_{s\alpha}|+\kappa_\beta|v_{s\beta}|+a\chi|\omega_p|). \tag{6}$$

The absolute sign is necessary in order to make the function $K_{f1}$ positive definite. Equation (4) becomes

$$K_{f2}=\tfrac{1}{2}(\mu_\alpha v_{s\alpha}^2+\mu_\beta v_{s\beta}^2+\nu\omega_p^2). \tag{7}$$

The coefficients $\kappa_\alpha$, $\kappa_\beta$, $\mu_\alpha$, and $\mu_\beta$ describe the orthotropy.

Now let us turn to Lagrange method and the procedure of obtaining the equations of motion. The system that is considered here includes body $B$ and all the particles that produce the friction mechanism. All these particles have only a kinetic energy (they do not have any potential), $T_f$:

$$T_f(t)=\int_0^t K_f(\tau)d\tau+\frac{1}{2}m_f v_l^2. \tag{8}$$

$m_f$ is the total mass of the particles that participated in producing the friction. As it will be shown later on, $m_f$ and $v_l$ do not affect the equations directly.

The Lagrangian integrand of the entire system, $L$, is

$$L=L_B+T_f \tag{9a}$$

$$L_B=T_B-V_B. \tag{9b}$$

$T_B$ and $V_B$ are the kinetic energy and potential, respectively, of body $B$. If necessary, holonomic constraints can be included by adding them to $L$, using the Lagrange multiplier approach.

Lagrange's equations are

$$\frac{d}{dt}\left(\frac{\partial L}{\partial\dot{q}_i}\right)-\frac{\partial L}{\partial q_i}=Q_i \quad (i=1,2,\ldots N). \tag{10}$$

$q_i$ is the $i$th generalized coordinate while $Q_i$ is the $i$th generalized applied force. $N$ is the number of degrees-of-freedom of the system. The collision forces are internal impulse forces of the system.

Since there are collisions between body $B$ and the small particles, Lagrange's equation will be integrated, as is commonly done in the case of impulsive motion. The integration will be carried over the time interval $t_1\leqslant t\leqslant t_2$, letting $t_2\to t_1$:

$$\lim_{t_2\to t_1}\int_{t_1}^{t_2}\left[\frac{d}{dt}\left(\frac{\partial L}{\partial\dot{q}_i}\right)-\frac{\partial L}{\partial q_i}\right]dt=\lim_{t_2\to t_1}\int_{t_1}^{t_2}Q_i dt. \tag{11}$$

There is a discontinuous change in the momentum of the particles that collide with body $B$ during the interval $t_1\leqslant t\leqslant t_2$. On the other hand, since body $B$ has a finite mass while the particles

are very small, the change in the momentum of body $B$ during this time interval is assumed to be continuous. Thus if Eq. (9$a$) is substituted into Eq. (11) and integration of the left side of the equation is carried out, we obtain

$$\lim_{t_2\to t_1}\int_{t_1}^{t_2}\left[\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right)-\frac{\partial L}{\partial q_i}\right]dt$$
$$=\left[\frac{d}{dt}\left(\frac{\partial T_B}{\partial \dot{q}_i}\right)-\frac{\partial(T_B-V_B)}{\partial q_i}\right]_{t=t_1}\Delta t$$
$$+\lim_{t_2\to t_1}\int_{t_1}^{t_2}\left[\frac{d}{dt}\left(\frac{\partial T_f}{\partial \dot{q}_i}\right)-\frac{\partial T_f}{\partial q_i}\right]dt \qquad (12a)$$

$$\Delta t=t_2-t_1. \qquad (12b)$$

The integral on the right side of Eq. (12$a$) includes the contributions of all the particles that collide with body $B$ during the time interval $t_1<t<t_2$. Let us examine the contribution to this integral of a single particle that collides with body $B$ at time $\tau$, where $t_1\leqslant\tau\leqslant t_2$. $T_f$ of that particle is (see Eq. (1)):

$$T_p(t)=\frac{1}{N_p}K_f(\tau)\cdot h(t-\tau)+\frac{1}{2}m_p v_l^2. \qquad (13)$$

$h(t-\tau)$ is the Heaviside (step) function.

Thus the generalized momentum of the particle, $\partial T_p/\partial \dot{q}_i$, is

$$\frac{\partial T_p}{\partial \dot{q}_i}=\frac{1}{N_p}\left.\frac{\partial K_f}{\partial \dot{q}_i}\right|_{t=\tau}h(t-\tau) \qquad (14)$$

Similarly,

$$\frac{\partial T_p}{\partial q_i}=\frac{1}{N_p}\left.\frac{\partial K_f}{\partial q_i}\right|_{t=\tau}h(t-\tau). \qquad (15)$$

Substitution of Eqs. (14) and (15) into the integral on the right side of Eq. (12$a$) for a single particle results in

$$\lim_{t_2\to t_1}\int_{t_1}^{t_2}\left[\frac{d}{dt}\left(\frac{\partial T_p}{\partial \dot{q}_i}\right)-\frac{\partial T_p}{\partial q_i}\right]dt$$
$$=\frac{1}{N_p}\int_{t_1}^{t_2}\left[\left.\frac{\partial K_f}{\partial \dot{q}_i}\right|_{t=\tau}\delta(t-\tau)-\left.\frac{\partial K_f}{\partial q_i}\right|_{t=\tau}h(t-\tau)\right]dt \qquad (16)$$

where $\delta(t-\tau)$ is the Dirac (impulse) function.

If $K_f$ is continuous in the interval, and since $t_2\to t_1$, Eq. (16) becomes

$$\lim_{t_2\to t_1}\int_{t_1}^{t_2}\left[\frac{d}{dt}\left(\frac{\partial T_p}{\partial \dot{q}_i}\right)-\frac{\partial T_p}{\partial q_i}\right]dt=\frac{1}{N_p}\left[\left.\frac{\partial K_f}{\partial \dot{q}_i}\right|_{t=t_1}-\left.\frac{\partial K_f}{\partial q_i}\right|_{t=t_1}\cdot\xi\cdot\Delta t\right]. \qquad (17)$$

$\xi$ takes values between zero and one $(0<\xi\leqslant 1)$. During the time interval $\Delta t$ the number of particles that collide with body $B$ is $(N_p\cdot\Delta t)$. Therefore

$$\lim_{t_2\to t_1}\int_{t_1}^{t_2}\left[\frac{d}{dt}\left(\frac{\partial T_f}{\partial \dot{q}_i}\right)-\frac{\partial T_f}{\partial q_i}\right]dt$$
$$=(N_p\cdot\Delta t)\lim_{t_2\to t_1}\int_{t_1}^{t_2}\left[\frac{d}{dt}\left(\frac{\partial T_p}{\partial \dot{q}_i}\right)-\frac{\partial T_p}{\partial q_i}\right]dt$$
$$=\left.\frac{\partial K_f}{\partial \dot{q}_i}\right|_{t=t_1}\Delta t-\left.\frac{\partial K_f}{\partial q_i}\right|_{t=t_1}\xi\Delta t^2. \qquad (18)$$

It will be assumed that the applied (external) generalized forces are continuous and do not include impulse forces. Thus the right side of Eq. (11) becomes

$$\lim_{t_2\to t_2}\int_{t_1}^{t_2}Q_i dt=Q_i|_{t=t_1}\Delta t. \qquad (19)$$

Substitution of Eq. (18) into Eq. (12$a$), and substitution of the result and Eq. (19) into Eq. (11), lead to the following equation:

$$\left[\frac{d}{dt}\left(\frac{\partial L_B}{\partial \dot{q}_i}\right)-\frac{\partial L_B}{\partial q_i}+\frac{\partial K_f}{\partial \dot{q}_i}-\xi\frac{\partial K_f}{\partial q_i}\Delta t-Q_i\right]_{t=t_1}\Delta t=0. \qquad (20)$$

Since $t_2\to t_1$ and $\Delta t\to 0$, we obtain the following Lagrange equation for the case of sliding and rolling:

$$\frac{d}{dt}\left(\frac{\partial L_B}{\partial \dot{q}_i}\right)-\frac{\partial L_B}{\partial q_i}+\frac{\partial K_f}{\partial \dot{q}_i}=Q_i \quad 1\leqslant i\leqslant N. \qquad (21)$$

$K_f$ plays the role of a dissipation function, but this is a special dissipation function that depends on the physical variables $v_s$ and $\omega_p$ and it has a physical background based on the collision mechanism that has been assumed.

The present derivation was based on a mechanism where the energy loss of body $B$ is due to collisions with particles of surface $S$. Yet, as was emphasized in the Introduction, this mechanism is not a description of the actual mechanism involved in the friction process. Instead the present mechanism should be considered as a simplified equivalent mechanism of the actual one. The model of collisions is a conceptual model that reproduces results of various friction mechanisms.

In what follows, Eq. (21) will be used to investigate the behavior of various systems that include sliding and rolling.

## 3 Rolling and Sliding of a Circular Disk on a Rough Surface

The two-dimensional case of a circular disk that slides and rolls on a horizontal rough surface is considered (see Fig. 2). The disk has a radius $R$, its mass is $m_B$ and the moment of inertia about its center is $m_B k^2$. A linear coordinate $x$ and an angular coordinate $\theta$ are used to describe the disk motion. $x$ is parallel to the surface and passes through the disk center. In this two-dimensional problem there is no pivoting. Except for the gravity force and the friction between the disk and the surface, there are no other forces that act on it. At $t=0$ the disk coordinates are $x_0$ and $\theta_0$, while the initial linear and angular velocities are $\dot{x}_0$ and $\dot{\theta}_0$, respectively.

The absolute value of the sliding velocity, $v_s$, is

$$v_s=|\dot{x}-R\dot{\theta}|. \qquad (22)$$

The kinetic energy of the disk is

$$T_B=\tfrac{1}{2}m_B\dot{x}^2+\tfrac{1}{2}m_B k^2\dot{\theta}^2. \qquad (23)$$

Since the motion is on a horizontal surface the gravity potential is constant and chosen equal to zero.

In what follows two cases of linear and quadratic $K_f$ functions will be considered.

**Linear $K_f$ Function.** The function $K_{f1}$ according to Eqs. (3) and (22) is

$$K_{f1}=m_B g\kappa|\dot{x}-R\dot{\theta}|. \qquad (24)$$

The last equation can be written as

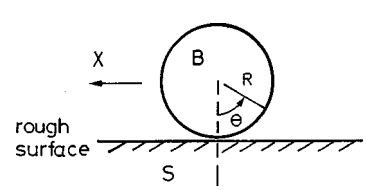



**Fig. 2 A circular disk that rolls and slides on a rough surface**

$$K_{f1}=m_B g\kappa(\dot{x}-R\dot{\theta})\mathrm{sign}(\dot{x}_0-R\dot{\theta}_0) \tag{25}$$

where the function sign $(z)$ indicates the sign of variable $z$ and takes values 1 or $-1$, accordingly. If in the course of the solution $(\dot{x}-R\dot{\theta})$ changes its sign, new initial conditions can be defined, and the solution is continued (in what follows this will be shown as unnecessary).

The two general coordinates in the present example are $x$ and $\theta$. Substitution of Eqs. (23) and (25) into Eq. (21) results in the following two equations of motion:

$$\ddot{x}+g\kappa\,\mathrm{sign}(\dot{x}_0-R\dot{\theta}_0)=0 \tag{26a}$$

$$k^2\ddot{\theta}-g\kappa R\,\mathrm{sign}(\dot{x}_0-R\dot{\theta}_0)=0. \tag{26b}$$

These equations are identical to the equations that are obtained while using Newton's method, when dry friction is assumed. $\kappa$ is then the coefficient of dry friction.

The solution of the last two equations is

$$x=-\frac{1}{2}g\kappa\,\mathrm{sign}(\dot{x}_0-R\dot{\theta}_0)t^2+\dot{x}_0t+x_0 \tag{27a}$$

$$\theta=\frac{g\kappa R}{2k^2}\mathrm{sign}(\dot{x}_0-R\dot{\theta}_0)t^2+\dot{\theta}_0t+\theta_0. \tag{27b}$$

The sliding velocity is (see Eq. (22))

$$v_s=\left|(\dot{x}_0-R\dot{\theta}_0)-g\kappa\left(1+\frac{R^2}{k^2}\right)\mathrm{sign}(\dot{x}_0-R\dot{\theta}_0)t\right|. \tag{28}$$

It turns out that the sliding velocity decreases linearly with time, until it becomes zero at time $t_R$:

$$t_R=\frac{1}{g\kappa\left(1+\frac{R^2}{k^2}\right)}|\dot{x}_0-R\dot{\theta}_0|. \tag{29}$$

From this moment on the disk moves without sliding, having constant linear and angular velocities, $\dot{x}_R$ and $\dot{\theta}_R$, respectively, that are equal to the values at $t=t_R$:

$$\dot{x}_R=\dot{x}_0-\frac{1}{\left(1+\frac{R^2}{k^2}\right)}(\dot{x}_0-R\dot{\theta}_0) \tag{30a}$$

$$\dot{\theta}_R=\dot{\theta}_0+\frac{R}{k^2\left(1+\frac{R^2}{k^2}\right)}(\dot{x}_0-R\dot{\theta}_0). \tag{30b}$$

Thus the complete solution to the problem is comprised of two parts:

(a) During the interval $0\leqslant t\leqslant t_R$ the motion is described by Eqs. (27*a*,27*b*). At the end of this time interval $(t=t_R)$ the coordinates of the disk are $x_R$ and $\theta_R$.

(b) For $t_R<t$ a motion with constant velocities $\dot{x}_R$ and $\dot{\theta}_R$ is obtained (see Eqs. (30*a*,30*b*)).

Thus in this region

$$x=x_R+\dot{x}_R(t-t_R) \tag{31a}$$

$$\theta=\theta_R+\dot{\theta}_R(t-t_R). \tag{31b}$$

It is interesting to note that the final velocities $\dot{x}_R$ and $\dot{\theta}_R$ are not dependent on the friction force, but are only dependent on the initial conditions, $R$ and $k$.

**Quadratic $K_f$ Function.** The function $K_f$, according to Eq. (4), is

$$K_{f2}=\tfrac{1}{2}\mu(\dot{x}-R\dot{\theta})^2. \tag{32}$$

Substitution of Eqs. (23) and (32) into Eq. (21), leads to the following equations of motion:

$$m_B\ddot{x}+\mu(\dot{x}-R\dot{\theta})=0 \tag{33a}$$

$$m_Bk^2\ddot{\theta}-\mu R(\dot{x}-R\dot{\theta})=0. \tag{33b}$$

It is interesting to note that Eqs. (33*a*,33*b*) are identical to those that are obtained after using Newton's method, when viscous friction is assumed.

The solution to the last two equations is

$$x=-\frac{m_B(\dot{x}_0-R\dot{\theta}_0)}{\mu\left(1+\frac{R^2}{k^2}\right)^2}e^{-\mu(1+R^2/k^2)t/m_B}+\left[\dot{x}_0-\frac{(\dot{x}_0-R\dot{\theta}_0)}{\left(1+\frac{R^2}{k^2}\right)}\right]t+x_0+\frac{m_B(\dot{x}_0-R\dot{\theta}_0)}{\mu\left(1+\frac{R^2}{k^2}\right)^2} \tag{34a}$$

$$\theta=-\frac{m_BR(\dot{x}_0-R\dot{\theta}_0)}{\mu k^2\left(1+\frac{R^2}{k^2}\right)^2}e^{-\mu(1+R^2/k^2)t/m_B}+\left[\dot{\theta}_0+\frac{R(\dot{x}_0-R\dot{\theta}_0)}{k^2\left(1+\frac{R^2}{k^2}\right)}\right]t+\theta_0-\frac{m_BR(\dot{x}_0-R\dot{\theta}_0)}{\mu k^2\left(1+\frac{R^2}{k^2}\right)}. \tag{34b}$$

The velocities $\dot{x}$ and $\dot{\theta}$ are obtained by differentiating the last equations. The sliding velocity becomes

$$v_s=|\dot{x}-R\dot{\theta}_0|=|\dot{x}_0-R\dot{\theta}_0|e^{-\mu(1+R^2/k^2)t/m_B}. \tag{35}$$

It should be noted that $v_s$ approaches zero asymptotically, as $\dot{x}$ and $\dot{\theta}$ approach their asymptotic values. The asymptotic values of $\dot{x}$ and $\dot{\theta}$ are identical to $\dot{x}_R$ and $\dot{\theta}_R$ (see Eqs. (30*a*,30*b*)) for the case of a linear $K_f$.

**Comment.** According to Eqs. (23) and (26*a*,26*b*) the rate of loss of energy of body $B$, $P=|dT_B/dt|$, in the case of dry friction, during the period $0<t<t_R$, is

$$P_1=m_Bg\kappa|\dot{x}-R\dot{\theta}|. \tag{36}$$

Comparison of Eq. (36) with Eq. (24) indicates that all the energy lost by body $B$ is absorbed by the particles that are responsible for the friction procedure.

In the case of viscous friction, the rate of loss of energy according to Eqs. (33*a*,33*b*) is

$$P_2=\mu(\dot{x}-R\dot{\theta})^2. \tag{37}$$

Comparison of the last equation with Eq. (32) indicates that in this case only half of the energy lost by body $B$, is absorbed by the particles.

## 4 Caplygin's Sleigh—A Special Case

The free motion of a solid body $B$ on a horizontal rough plane is considered. The body touches the plane at three points. The first two points of contact can slide freely on the plane in all directions. The projection of the center of mass of $B$ coincides with the third point of contact of a massless rigid sharp wheel and the plane (see Fig. 3). This is a special case of Caplygin's sleigh that was considered by various researchers (e.g., [1,2]).

The kinetic energy of body $B$ is

$$T_B=\tfrac{1}{2}m_B(\dot{x}^2+\dot{y}^2)+\tfrac{1}{2}m_Bk^2\dot{\varphi}^2 \tag{38}$$

where $x$ and $y$ are the planar Cartesian coordinates of the center of mass. $\varphi$ is the angle between the plane of the wheel and the $x$-axis (see Fig. 3). $k$ is the radius of gyration and $m_B$ the body mass.

The potential of body $B$ is constant and thus chosen equal to zero.

This is a case of orthotropic friction. Since the wheel has a sharp rigid rim, the friction along the plane of the wheel is much smaller than the friction normal to this plane.

As defined in Section 2, $\alpha$ and $\beta$ indicate the orientation of the orthotropy, $\alpha$ is the direction along the plane of the wheel, and $\beta$ is the direction normal to it. If $\mathbf{e}_x$ and $\mathbf{e}_y$ are unit vectors in the directions of the axes $x$ and $y$, respectively, then (Fig. 3)

$$\mathbf{e}_\alpha = \cos\varphi \mathbf{e}_x + \sin\varphi \mathbf{e}_y \tag{39a}$$

$$\mathbf{e}_\beta = -\sin\varphi \mathbf{e}_x + \cos\varphi \mathbf{e}_y . \tag{39b}$$

The sliding velocity in this example is the velocity of the body center of mass, thus

$$\mathbf{v}_s = \dot{x}\mathbf{e}_x + \dot{y}\mathbf{e}_y . \tag{40}$$

The components of $\mathbf{v}_s$ in the directions $\mathbf{e}_\alpha$ and $\mathbf{e}_\beta$ are

$$v_{s\alpha} = \dot{x}\cos\varphi + \dot{y}\sin\varphi \tag{41a}$$

$$v_{s\beta} = -\dot{x}\sin\varphi + \dot{y}\cos\varphi. \tag{41b}$$

The pivoting velocity in the present case is

$$\omega_p = \dot{\varphi}. \tag{42}$$

**Linear $K_f$ Function.** According to Eq. (6), for the present case

$$K_{f1} = m_B g(\kappa_\alpha|\dot{x}\cos\varphi + \dot{y}\sin\varphi| + \kappa_\beta|-\dot{x}\sin\varphi + \dot{y}\cos\varphi| + a\chi|\dot{\varphi}|)$$

where

$$\kappa_\beta \gg \kappa_\alpha, a\chi. \tag{43}$$

For the sake of comparison with results from the literature the solution is restricted to the cases where $\kappa_\beta$ is much larger than $\kappa_\alpha$ or $a\chi$, so the last two are neglected. Then we obtain

$$K_{f1} = m_B g\kappa_\beta(-\dot{x}\sin\varphi + \dot{y}\cos\varphi)\mathrm{sign}(-\dot{x}\sin\varphi + \dot{y}\cos\varphi). \tag{44}$$

There are three generalized coordinates in this case $(x,y,\varphi)$. Substitution of Eq. (38) into Eq. (9$b$), and then substitution of the result into Eq. (21), lead to the following three equations:

$$\ddot{x} - g\kappa_\beta \sin\varphi\, \mathrm{sign}(-\dot{x}\sin\varphi + \dot{y}\cos\varphi) = 0 \tag{45a}$$

$$\ddot{y} + g\kappa_\beta \cos\varphi\, \mathrm{sign}(-\dot{x}\sin\varphi + \dot{y}\cos\varphi) = 0 \tag{45b}$$

$$\ddot{\varphi} = 0. \tag{45c}$$

Again, as in the previous section, it is shown that a linear $K_f$ function leads to a dry friction force. The solution of the last equations is

$$x = -\frac{g\kappa_\beta}{\dot{\varphi}_0^2}[(\sin\varphi - \sin\varphi_0) - \dot{\varphi}_0\cos\varphi_0 t] \times \mathrm{sign}(-\dot{x}_0\sin\varphi_0 + \dot{y}_0\cos\varphi_0) + \dot{x}_0 t + x_0 \tag{46a}$$

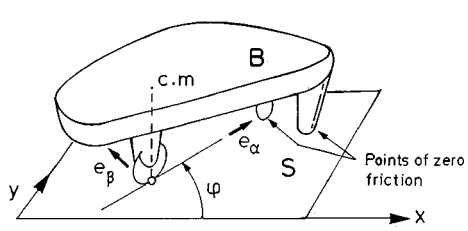



**Fig. 3 A special case of Caplygin's sleigh**

$$y = \frac{g\kappa_\beta}{\dot{\varphi}_0^2}[(\cos\varphi - \cos\varphi_0) + \dot{\varphi}_0\sin\varphi_0 t] \times \mathrm{sign}(-\dot{x}_0\sin\varphi_0 + \dot{y}_0\cos\varphi_0) + \dot{y}_0 t + y_0 \tag{46b}$$

$$\varphi = \dot{\varphi}_0 t + \varphi_0 . \tag{46c}$$

$x_0$, $y_0$, and $\varphi_0$ are the initial values of the generalized coordinates while $\dot{x}_0$, $\dot{y}_0$, and $\dot{\varphi}_0$ are the values of the initial generalized velocities.

The sliding velocity component normal to the plane of the wheel, $v_{S\beta}$, is

$$v_{S\beta} = (-\dot{x}\sin\varphi + \dot{y}\cos\varphi) = -\dot{x}_0\sin\varphi + \dot{y}_0\cos\varphi - \frac{g\kappa_\beta}{\dot{\varphi}_0}\sin(\varphi - \varphi_0)\mathrm{sign}(-\dot{x}_0\sin\varphi_0 + \dot{y}_0\cos\varphi_0). \tag{47}$$

The solution (46$a$–46$c$) is correct as long as $v_{S\beta}$ does not become zero. At the moment when $v_{S\beta}$ becomes zero, it is not clear what is the magnitude of the component of the friction force normal to the plane of the wheel, nor what is the direction of this force. Yet one can examine the situation just after $v_{S\beta}$ becomes zero. Differentiating Eq. (41$b$) and using Eqs. (41$a$) and (46$c$) imply

$$\dot{v}_{S\beta} = -v_{S\alpha}\dot{\varphi}_0 + (-\ddot{x}\sin\varphi + \ddot{y}\cos\varphi) = -v_{S\alpha}\dot{\varphi}_0 + \frac{F_{f\beta}}{m}. \tag{48}$$

$F_{f\beta}$ is the component of the friction force in the direction $\mathbf{e}_\beta$.

If at the moment when $v_{S\beta} = 0$ also $F_{f\beta} = 0$, then the sign of $v_{S\beta}$ an instant later is defined by the sign of $v_{S\alpha}$ and $\dot{\varphi}_0$. Knowing the sign of $v_{S\beta}$ one can consider the conditions at that moment as initial conditions for a solution, which is identical in its form to Eqs. (46$a$,46$b$).

Let us consider a ''more sophisticated mechanism,'' similar to static friction, where the force $F_{f\beta}$ (at the moment when $v_{S\beta}$ becomes zero) is determined such that $v_{S\beta}$ will be zero in what follows, namely (see Eq. (48))

$$F_{f\beta} = m v_t \dot{\varphi}_0 . \tag{49}$$

$v_t$ in the last equation is the magnitude of the component $v_{S\alpha}$ (the component in the plane of the wheel) at the moment where $v_{S\beta}$ becomes zero. The equations of motion become

$$\ddot{x} = -v_t\dot{\varphi}_0\sin\varphi \tag{50a}$$

$$\ddot{y} = v_t\dot{\varphi}_0\cos\varphi. \tag{50b}$$

The solution of these equations is

$$x = \frac{v_t}{\dot{\varphi}_0}(\sin\varphi - \sin\varphi_t) + x_t \tag{51a}$$

$$y = -\frac{v_t}{\dot{\varphi}_0}(\cos\varphi - \cos\varphi_t) + y_t . \tag{51b}$$

$x_t$, $y_t$, and $\varphi_t$ are the coordinates at the moment when $v_{S\beta}$ becomes zero.

The solution (51$a$,51$b$) is identical to the one given in the literature ([1], pp. 94–95) for a motion without sliding normal to the wheel plane. This is a circular motion. In the literature the mechanism that results in a motion described by these equations is not of interest. Yet, from a physical point of view, the above sophisticated mechanism seems to be unrealistic. In what follows it will be shown that when viscous friction is considered the phenomenon becomes more feasible and logical.

**Quadratic $K_f$ Function.** Similar to the case of a linear $K_f$ function, it will be assumed that $\mu_\alpha$ and $\nu$ in Eq. (7) are very small compared to $\mu_\beta$, and therefore they will be neglected. Thus we obtain

$$K_{f2}=\tfrac{1}{2}\mu_\beta(-\dot{x}\sin\varphi+\dot{y}\cos\varphi)^2. \tag{52}$$

Using Eq. (21) the following three equations of motion are obtained:

$$\ddot{x}-\rho(-\dot{x}\sin\varphi+\dot{y}\cos\varphi)\sin\varphi=0 \tag{53a}$$

$$\ddot{y}+\rho(-\dot{x}\sin\varphi+\dot{y}\cos\varphi)\cos\varphi=0 \tag{53b}$$

$$\ddot{\varphi}=0 \tag{53c}$$

$$\rho=\mu_\beta/m. \tag{53d}$$

The solution of the above equations is (for the sake of simplicity the velocities $\dot{x}$ and $\dot{y}$ are given, if necessary they can be integrated to obtain $x$ and $y$)

$$\dot{x}=\frac{\dot{\varphi}_0}{\zeta_1-\zeta_2}\left[-\left(v_{t0}+\frac{\dot{\varphi}_0}{\zeta_2}v_{n0}\right)\left(\sin\varphi+\frac{\dot{\varphi}_0}{\zeta_1}\cos\varphi\right)e^{-\zeta_1 t}+\left(v_{t0}+\frac{\dot{\varphi}_0}{\zeta_1}v_{n0}\right)\left(\sin\varphi+\frac{\dot{\varphi}_0}{\zeta_2}\cos\varphi\right)e^{-\zeta_2 t}\right] \tag{54a}$$

$$\dot{y}=\frac{\dot{\varphi}_0}{\zeta_1-\zeta_2}\left[\left(v_{t0}+\frac{\dot{\varphi}_0}{\zeta_2}v_{n0}\right)\left(\cos\varphi-\frac{\dot{\varphi}_0}{\zeta_1}\sin\varphi\right)e^{-\zeta_1 t}-\left(v_{t0}+\frac{\dot{\varphi}_0}{\zeta_1}v_{n0}\right)\left(\cos\varphi-\frac{\dot{\varphi}_0}{\zeta_2}\sin\varphi\right)e^{-\zeta_2 t}\right] \tag{54b}$$

$$\varphi=\dot{\varphi}_0 t+\varphi_0 \tag{54c}$$

$$v_{t0}=\dot{x}_0\cos\varphi_0+\dot{y}_0\sin\varphi_0 \tag{54d}$$

$$v_{n0}=-\dot{x}_0\sin\varphi_0+\dot{y}_0\cos\varphi_0 \tag{54e}$$

$$\zeta_{1,2}=\frac{\rho}{2}\left[1\pm\sqrt{1-4\frac{\dot{\varphi}_0^2}{\rho^2}}\right]. \tag{54f}$$

$v_{t0}$ and $v_{n0}$ are the initial velocity components in the directions $\mathbf{e}_\alpha$ and $\mathbf{e}_\beta$ at $t=0$, respectively (the components along the plane of the wheel and normal to it).

Of special interest is the case of an infinitely rough surface, namely

$$\mu_\beta,\rho\to\infty. \tag{55}$$

According to Eq. (54*f*) we obtain

$$\zeta_1\cong\rho\to\infty \tag{56a}$$

$$\zeta_2\cong\frac{\dot{\varphi}_0^2}{\rho}\to 0. \tag{56b}$$

Based on the above equations

$$\frac{\dot{\varphi}_0}{\zeta_1-\zeta_2}\cong\frac{\dot{\varphi}_0}{\rho}\to 0 \tag{57a}$$

$$\frac{\dot{\varphi}_0^2}{(\zeta_1-\zeta_2)\zeta_1}\cong\frac{\dot{\varphi}_0^2}{\rho^2}\to 0 \tag{57b}$$

$$\frac{\dot{\varphi}_0^2}{(\zeta_1-\zeta_2)\zeta_2}\cong 1 \tag{57c}$$

$$\frac{\dot{\varphi}_0^3}{(\zeta_1-\zeta_2)\zeta_1\zeta_2}\cong\frac{\dot{\varphi}_0}{\rho}\to 0 \tag{57d}$$

$$e^{-\zeta_1 t}\to 0 \tag{57e}$$

$$e^{-\zeta_2 t}\to 1. \tag{57f}$$

Using Eqs. (57*a*–57*f*), it turns out that for an infinitely rough surface

$$\dot{x}\cong v_{t0}\cos\varphi \tag{58a}$$

$$\dot{y}\cong v_{t0}\sin\varphi \tag{58b}$$

$$\dot{\varphi}=\dot{\varphi}_0 t+\varphi_0. \tag{58c}$$

Equations (58*a*–58*c*) describe a circular motion at a constant angular speed, the same motion as obtained by Neimark and Fufaev for the nonholonomic constraint of zero sliding velocity normal to the plane of the wheel. Unlike the previous case of dry friction, here this motion is obtained as a result of a ''realistic'' physical mechanism: As $\rho$ in Eqs. (53*a*,53*b*) is increased, the sliding velocity during the motion, $(-\dot{x}\sin\varphi+\dot{y}\cos\varphi)$, is getting smaller. At the limit

$$\lim_{\rho\to\infty}(-\dot{x}\sin\varphi+\dot{y}\cos\varphi)=0 \tag{59a}$$

$$\lim_{\rho\to\infty}\rho(-\dot{x}\sin\varphi+\dot{y}\cos\varphi)=\lambda. \tag{59b}$$

$\lambda$ is the viscous friction force normal to the wheel plane. Thus for the limit case where $\rho\to\infty$, the equations of motion (53*a*–53*c*) become

$$\ddot{x}-\lambda\sin\varphi=0 \tag{60a}$$

$$\ddot{y}+\lambda\cos\varphi=0 \tag{60b}$$

$$\ddot{\varphi}=0 \tag{60c}$$

$$-\dot{x}\sin\varphi+\dot{y}\cos\varphi=0. \tag{60d}$$

These are exactly Eqs. (2.5) and (2.6) of Neimark and Fufaev, where $\lambda$ is the Lagrange multiplier.

## 5 Motion of a Sphere on a Rough Surface

Consider a homogeneous sphere $B$ of mass $m_B$, radius $a$ and a radius of gyration $k$. This sphere rolls on a rough horizontal flat surface $S$ that has an orthotropy of friction.

The motion of the center of the sphere is described relative to the fixed system of coordinates $x$, $y$, $z$. $x$ and $y$ are parallel to the surface $S$ and chosen such that they coincide with the directions of the orthotropy. The rotation of the sphere is described by three Eulerian angles $\varphi$, $\psi$, and $\theta$, which are defined in Fig. 4 (the sequence of the rotation is $\psi$, $\theta$, and $\varphi$). From Fig. 4 it is easy to see that the vector of angular velocity of the sphere, $\overline{\omega}$, is

$$\overline{\omega}=\omega_x\mathbf{e}_x+\omega_y\mathbf{e}_y+\omega_z\mathbf{e}_z \tag{61}$$

where

$$\omega_x=\dot{\theta}\cos\psi+\dot{\varphi}\sin\theta\sin\psi \tag{62a}$$

$$\omega_y=\dot{\theta}\sin\psi-\dot{\varphi}\sin\theta\cos\psi \tag{62b}$$

$$\omega_z=\dot{\varphi}\cos\theta+\dot{\psi}. \tag{62c}$$

$\mathbf{e}_x$, $\mathbf{e}_y$, $\mathbf{e}_z$ are unit vectors in the directions of the coordinate lines $x$, $y$, $z$, respectively.

The kinetic energy of the sphere is

$$T_\beta=\tfrac{1}{2}m_B(\dot{x}^2+\dot{y}^2)+\tfrac{1}{2}m_B k^2(\omega_x^2+\omega_y^2+\omega_z^2)=\tfrac{1}{2}m_B(\dot{x}^2+\dot{y}^2)+\tfrac{1}{2}m_B k^2(\dot{\theta}^2+\dot{\varphi}^2+\dot{\psi}^2+2\dot{\varphi}\dot{\psi}\cos\theta). \tag{63}$$

Since the surface $S$ is horizontal, $V_B$ is chosen equal to zero.

The sliding velocity of the sphere on surface $S$, $\mathbf{v}_S$, is

$$\mathbf{v}_S=\dot{x}\mathbf{e}_x+\dot{y}\mathbf{e}_y+\overline{\omega}x(-a\mathbf{e}_z)=(\dot{x}-a\omega_y)\mathbf{e}_x+(\dot{y}+a\omega_x)\mathbf{e}_y. \tag{64}$$

In the present example a quadratic $K_f$ function will be considered. It will be assumed that the friction due to pivoting is very small and thus it will be neglected. According to Eq. (7),

$$K_{f2}=\tfrac{1}{2}[\mu_x(\dot{x}-a\omega_y)^2+\mu_y(\dot{y}+a\omega_x)^2]. \tag{65}$$

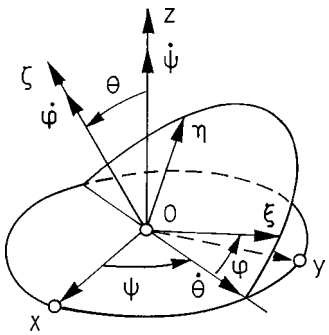



**Fig. 4 Description of Euler angles. $(x,y,z)$ and $(\xi,\eta,\zeta)$ represent a body system of coordinates before and after the rotation, respectively.**

Substitution of Eqs. (63) and (65) into Eq. (21), and using Eqs. (62*a*–62*c*), lead to the following five equations of motion (the generalized coordinates are: $x$, $y$, $\theta$, $\varphi$, and $\psi$):

$$m_B\ddot{x}+\mu_x(\dot{x}-a\omega_y)=0 \tag{66a}$$

$$m_B\ddot{y}+\mu_y(\dot{y}+a\omega_x)=0 \tag{66b}$$

$$m_Bk^2(\ddot{\theta}+\dot{\varphi}\dot{\psi}\sin\theta)-\mu_x a(\dot{x}-a\omega_y)\sin\psi + \mu_y a(\dot{y}+a\omega_x)\cos\psi=0 \tag{66c}$$

$$m_Bk^2(\ddot{\varphi}+\ddot{\psi}\cos\theta+\dot{\psi}\dot{\theta}\sin\theta)+\mu_x a(\dot{x}-a\omega_y)\sin\theta\cos\psi + \mu_y a(\dot{y}+a\omega_x)\sin\theta\sin\psi=0 \tag{66d}$$

$$m_Bk^2(\ddot{\psi}+\ddot{\varphi}\cos\theta-\dot{\varphi}\dot{\theta}\sin\theta)=0. \tag{66e}$$

The terms inside the parenthesis in Eq. (66*e*) are simply $\dot{\omega}_z$ (see Eq. (62*c*)). Equations (66*c*,66*d*) are replaced by another two equations that are obtained as follows:

(a) Equation (66*d*) is multiplied by $\cos\psi$ and subtracted from Eq. (66*c*) that is multiplied before that by $(\sin\theta\sin\psi)$. Use is made of Eq. (66*e*) to eliminate $\ddot{\psi}$.

(b) Equation (66*d*) is multiplied by $\sin\psi$, Eq. (66*c*) is multiplied by $(\sin\theta\cos\psi)$, both equations are added, and Eq. (66*e*) is used to eliminate $\ddot{\psi}$.

The new system of five equations is

$$m_B\ddot{x}+\mu_x(\dot{x}-a\omega_y)=0 \tag{67a}$$

$$m_B\ddot{y}+\mu_y(\dot{y}+a\omega_x)=0 \tag{67b}$$

$$m_Bk^2\dot{\omega}_y-\mu_x a(\dot{x}-a\omega_y)=0 \tag{67c}$$

$$m_Bk^2\dot{\omega}_x+\mu_y a(\dot{y}+a\omega_x)=0 \tag{67d}$$

$$\dot{\omega}_z=0. \tag{67e}$$

Consider now the case of an infinitely rough surface where

$$\mu_x,\mu_y\to\infty. \tag{68}$$

In this case, similar to what was shown in the previous section,

$$\lim_{\mu_x\to\infty}(\dot{x}-a\omega_y)\to 0 \tag{69a}$$

$$\lim_{\mu_y\to\infty}(\dot{y}+a\omega_x)\to 0 \tag{69b}$$

$$\lim_{\mu_x\to\infty}\mu_x(\dot{x}-a\omega_y)\to\lambda_x \tag{69c}$$

$$\lim_{\mu_y\to\infty}\mu_y(\dot{y}+a\omega_x)\to\lambda_y. \tag{69d}$$

$\lambda_x$ and $\lambda_y$ are newly defined variables that have finite values. The system of equations of motion for an infinitely rough surface becomes

$$m_B\ddot{x}+\lambda_x=0 \tag{70a}$$

$$m_B\ddot{y}+\lambda_y=0 \tag{70b}$$

$$m_Bk^2\dot{\omega}_y-a\lambda_x=0 \tag{70c}$$

$$m_Bk^2\dot{\omega}_x+a\lambda_y=0 \tag{70d}$$

$$\dot{\omega}_z=0 \tag{70e}$$

$$\dot{x}-a\omega_y=0 \tag{70f}$$

$$\dot{y}+a\omega_x=0. \tag{70g}$$

This is the same set that is obtained by the classical approach to nonholonomic constraints when one assumes rolling without sliding of the sphere (see [1]). $\lambda_x$ and $\lambda_y$ are Lagrange multipliers in this case.

## 6 Conclusions

An approach using Lagrange's method to derive the equations of motion of bodies which undergo sliding and rolling was presented. The method is based on modeling the friction between the body (that slides and rolls) and the surface, especially the loss of energy of the moving body, as a process of collisions between the body and particles of the surface. As a result of these collisions energy is transferred from the body to the particles, namely to the surface. The mechanism is defined by the function $K_f$ that describes the rate of increase of the kinetic energy of the particles that participate in the process. $K_f$ depends on the sliding velocity between the body and the surface and the pivoting speed. Two cases of the function $K_f$ have been investigated: linear and quadratic functions. It has been shown that linear functions represent dry friction while quadratic functions represent viscous friction.

The dynamic system includes the body $B$ and the particles of $S$ that participate in the process. Since collisions inside the system (between body $B$ and the particles) are considered, the Lagrange equations of the entire system are integrated with respect to time, similar to what is done in the case of an impulse motion. Modified Lagrange equations are obtained that are capable of describing motions that include simultaneous sliding and rolling. In these equations the function $K_f$ plays a role similar to a dissipation function.

It should be clear that the collisions mechanism proposed to describe the friction phenomenon is not a description of the actual mechanism involved in the friction process. Instead, the present mechanism should be considered as a simplified equivalent mechanism of the actual one.

Three examples of using the method have been presented, including isotropic friction and orthotropic friction. It has been shown that the case of zero sliding is obtained as a special case of the general solution, when infinite roughness of the surface is assumed. In the case of infinite roughness, quadratic $K_f$ functions, namely viscous friction, give equations of motion that are identical to those that are obtained when nonholonomic constraints of zero sliding are applied. Lagrange multipliers that appear in these equations, that emerge from the introduction of nonholonomic constraints, obtain a ''realistic physical meaning,'' and the mechanism behind the phenomenon becomes clear.

The newly presented method is very general in its nature. It can accommodate other forms (not only linear or quadratic) of the $K_f$ functions that will represent different forms of friction. Examples of other models of friction were presented by Oden and Martins [3], Banerjee and Kane [4], Bauchau [5], and others. The examples that were presented here dealt with a moving body $B$ and a stationary surface $S$. Yet, the application of the method to sliding and rolling between moving bodies, including more than two

bodies (multibodies) is straightforward. In the above investigation the (usually) very small effect of friction due to rolling has been neglected. Its addition, if necessary, is also straightforward using the technique presented above.

## Acknowledgment

The author would like to thank Mrs. A. Goodman for typing the manuscript and Mrs. D. Rosen for the preparation of the figures. This research was supported by the Fund for the Promotion of Research in the Technion. The author would also like to thank the reviewers of this paper for their useful comments that helped improve it.

## References

[1] Neimark, I., Jr., and Fufaev, N. A., 1972, *Dynamics of Nonholonomic Systems*, The American Mathematical Society, Providence, RI (translation of the Russian book from 1967).

[2] Hamel, G., 1949, ''Theoretische Mechanik,'' *Die Grundlehren der Mathematischen Wissenschaften in Ein Zeldarstellungen*, Band LVII, Springer, Berlin, pp. 464–473.

[3] Oden, J. T., and Martins, J. A. C., 1985, ''Models and Computational Methods for Dynamic Friction Phenomena,'' Comput. Methods Appl. Mech. Eng., **52**, pp. 527–634.

[4] Banerjee, A. K., and Kane, T. R., 1994, ''Modeling and Simulation of Rotor Bearing Friction,'' AIAA J. Guidance and Control, **17**, No. 5, pp. 1137–1139.

[5] Bauchau, O. A., 1999, ''On the Modeling of Friction and Rolling in Flexible Multi-Body Systems,'' Multibody Syst. Dyn., **3**, pp. 209–239.

# On the General Solution for Piezothermoelasticity for Transverse Isotropy With Application

**W. Q. Chen**
Associate Professor,
Department of Civil Engineering,
Zhejiang University,
Hangzhou 310027, P. R. China
e-mail: caijb@ccea.zju.edu.cn

*This paper derives a general solution of the three-dimensional equations of transversely isotropic piezothermoelastic materials (crystal class, 6 mm). Two displacement functions are first introduced to simplify the basic equations and a general solution is then derived using the operator theory. For the static case, the proposed general solution is very simple in form and can be used easily in certain boundary value problems. An illustrative example is given in the paper by considering the symmetric crack problem of an arbitrary temperature applied over the faces of a flat crack in an infinite space. The governing integro-differential equations of the problem are derived. It is found that exact expressions for the piezothermoelastic field for a penny-shaped crack subject to a uniform temperature can be obtained in terms of elementary functions.* [S0021-8936(00)01704-9]

## 1 Introduction

Piezoelectric ceramics and composites have been extensively used in many engineering applications such as sensors, actuators, intelligent structures, etc. ([1]). The mechanics of these so-called smart materials thus have attracted considerable academic attention ([2–6]). Owing to the pyroelectric effect occupied by most piezoelectric materials ([7]), the response characteristics may change considerably when piezoelectric structures work in environments where temperature varies notably. In order for the practical use of piezoelectric-based intelligent structures, a wide and thorough understanding of piezothermoelastic behaviors is required.

Most of the earlier investigations in the area of piezothermoelasticity have considered the development of constitutive relations as well as basic principles ([8–12]). Since 1990, the piezothermoelastic behaviors of piezoelectric plates and shells have been of intense research effort (see [13–18], to name a few). A potential function method was proposed by Ashida et al. [19–22] to effectively analyze various problems in piezothermoelasticity [23–26]. Kapuria et al. also applied the potential function method to investigate piezothermoelastic responses of a finite transversely isotropic cylindrical shell ([27]) and a clamped circular plate ([28]). However, as pointed out in Ding et al. [29], the derivation of the general solution presented by Ashida et al. is somehow tedious. Ding et al. [29] employed the operator theory and obtained a general solution in simpler form for static piezothermoelastic problems of crystal class 6 mm solids. The general solution for dynamic piezothermoelastic problems was also obtained recently in Ding et al. [30]. It is noted here that both the general solutions suggested by Ashida et al. [21,22] and Ding et al. [29,30] include a so-called piezothermoelastic potential function that satisfies a nonhomogeneous sixth-order differential equation with the nonhomogeneous term related to the temperature field. The general solution in such a form will however bring out certain inconvenience for dealing with boundary value problems associated with cracks and punches ([31]), while great achievement has been developed in the potential theory ([32]).

This paper presents a general solution for dynamic problems of transversely isotropic piezothermoelastic materials. In contrast to the one obtained by Ding et al. [30], this solution does not include a particular part that satisfies a nonhomogeneous differential equation. In particular, for static problems, the general solution is finally expressed in terms of five harmonic functions (more precisely, weighted or quasi-harmonic functions). Such a form allows one to use the solution feasibly to solve boundary value problems associated with cracks and punches. In fact, utilizing this solution, the potential theory method proposed by Fabrikant [32] and generalized by Chen et al. [5,6,33] for piezoelasticity can be easily applied to piezothermoelasticity. As an illustrative example, the stationary problem of a flat crack with its surfaces being parallel to the plane of isotropy and subjected to a prescribed temperature is considered in the paper. The governing integro-differential equations are derived. In particular, explicit expressions for the elastic, electric, and thermal field can be obtained for the problem that uniform temperature is applied over the faces of a penny-shaped crack. The stress intensity factor as well as the electric displacement intensity factor are presented.

## 2 Basic Equations

In Cartesian coordinates $(x,y,z)$, with the plane $x-y$ parallel to the plane of isotropy, the constitutive relations of a transversely isotropic piezothermoelastic medium (crystal class 6 mm) are

$$\begin{aligned}
\sigma_x &= c_{11}\frac{\partial u}{\partial x} + c_{12}\frac{\partial v}{\partial y} + c_{13}\frac{\partial w}{\partial z} + e_{31}\frac{\partial \Phi}{\partial z} - \lambda_{11}\theta,\\
\sigma_y &= c_{12}\frac{\partial u}{\partial x} + c_{11}\frac{\partial v}{\partial y} + c_{13}\frac{\partial w}{\partial z} + e_{31}\frac{\partial \Phi}{\partial z} - \lambda_{11}\theta,\\
\sigma_z &= c_{13}\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) + c_{33}\frac{\partial w}{\partial z} + e_{33}\frac{\partial \Phi}{\partial z} - \lambda_{33}\theta,\\
\tau_{yz} &= c_{44}\left(\frac{\partial v}{\partial z} + \frac{\partial w}{\partial y}\right) + e_{15}\frac{\partial \Phi}{\partial y},
\end{aligned} \tag{1}$$

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Jan. 3, 2000; final revision, May 1, 2000. Associate Technical Editor: J. R. Barber. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

 Copyright © 2000 by ASME 

$$\tau_{zx}=c_{44}\left(\frac{\partial u}{\partial z}+\frac{\partial w}{\partial x}\right)+e_{15}\frac{\partial \Phi}{\partial x},$$

$$\tau_{xy}=c_{66}\left(\frac{\partial u}{\partial y}+\frac{\partial v}{\partial x}\right),$$

$$D_x=e_{15}\left(\frac{\partial u}{\partial z}+\frac{\partial w}{\partial x}\right)-\epsilon_{11}\frac{\partial \Phi}{\partial x},$$

$$D_y=e_{15}\left(\frac{\partial v}{\partial z}+\frac{\partial w}{\partial y}\right)-\epsilon_{11}\frac{\partial \Phi}{\partial y}, \tag{2}$$

$$D_z=e_{31}\left(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}\right)+e_{33}\frac{\partial w}{\partial z}-\epsilon_{33}\frac{\partial \Phi}{\partial z}+p_3\theta,$$

where $u(v,w)$, $D_i$, and $\sigma_i(\tau_{ij})$ are components of the mechanical displacement, electric displacement, and normal (shear) stress, respectively; $\Phi$ and $\theta$ are the electric potential and temperature increment, respectively; $c_{ij}$, $e_{ij}$, $\epsilon_{ij}$, and $p_3$ are the elastic, piezoelectric, dielectric, and pyroelectric constants, respectively; and $\lambda_{ij}$ are the stress-temperature coefficients. The relation $c_{11}=c_{12}+2c_{66}$ holds for materials with transverse isotropy.

The dynamic equations of equilibrium and the electrostatic equation can be found in Tiersten [34] for example and are not repeated here. The temperature is assumed to satisfy the Fourier heat conduction equation ([26])

$$\left(k_{11}\Delta+k_{33}\frac{\partial^2}{\partial z^2}\right)\theta=\frac{k_{11}}{\kappa}\frac{\partial \theta}{\partial t}, \tag{3}$$

where $\Delta=\partial^2/\partial x^2+\partial^2/\partial y^2$ is the planar Laplacian, $k_{ij}$ are the coefficients of heat conduction, and $\kappa$ denotes the thermal diffusivity.

By virtue of Eqs. (1) and (2), the dynamic equations of equilibrium and the electrostatic equation for transversely isotropic piezothermoelasticity can be written in terms of the mechanical displacement, the electric potential, and the temperature increment as follows (in absence of body force and free charge density):

$$\begin{aligned}
&\left(c_{11}\frac{\partial^2}{\partial x^2}+c_{66}\frac{\partial^2}{\partial y^2}+c_{44}\frac{\partial^2}{\partial z^2}-\rho\frac{\partial^2}{\partial t^2}\right)u+(c_{12}+c_{66})\frac{\partial^2 v}{\partial x\partial y}\\
&\quad+(c_{13}+c_{44})\frac{\partial^2 w}{\partial x\partial z}+(e_{15}+e_{31})\frac{\partial^2\Phi}{\partial x\partial z}-\lambda_{11}\frac{\partial\theta}{\partial x}=0,\\
&(c_{12}+c_{66})\frac{\partial^2 u}{\partial x\partial y}+\left(c_{66}\frac{\partial^2}{\partial x^2}+c_{11}\frac{\partial^2}{\partial y^2}+c_{44}\frac{\partial^2}{\partial z^2}-\rho\frac{\partial^2}{\partial t^2}\right)v\\
&\quad+(c_{13}+c_{44})\frac{\partial^2 w}{\partial y\partial z}+(e_{15}+e_{31})\frac{\partial^2\Phi}{\partial y\partial z}-\lambda_{11}\frac{\partial\theta}{\partial y}=0,\\
&(c_{13}+c_{44})\frac{\partial}{\partial z}\left(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}\right)+\left(c_{44}\Delta+c_{33}\frac{\partial^2}{\partial z^2}-\rho\frac{\partial^2}{\partial t^2}\right)w\\
&\quad+\left(e_{15}\Delta+e_{33}\frac{\partial^2}{\partial z^2}\right)\Phi-\lambda_{33}\frac{\partial\theta}{\partial z}=0,\\
&(e_{15}+e_{31})\frac{\partial}{\partial z}\left(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}\right)+\left(e_{15}\Delta+e_{33}\frac{\partial^2}{\partial z^2}\right)w\\
&\quad-\left(\epsilon_{11}\Delta+\epsilon_{33}\frac{\partial^2}{\partial z^2}\right)\Phi+p_3\frac{\partial\theta}{\partial z}=0,
\end{aligned} \tag{4}$$

where $\rho$ is the mass density.

## 3 Dynamic General Solution

Two displacement functions $\psi$ and $G$ are introduced to represent the displacement components $u$ and $v$ as follows:

$$u=\frac{\partial\psi}{\partial y}-\frac{\partial G}{\partial x},\quad v=-\frac{\partial\psi}{\partial x}-\frac{\partial G}{\partial y}. \tag{5}$$

By virtue of Eq. (5), in a similar manner to that of Ding et al. [35], Eqs. (3) and (4) can be transformed into the following form:

$$\left(c_{66}\Delta+c_{44}\frac{\partial^2}{\partial z^2}-\rho\frac{\partial^2}{\partial t^2}\right)\psi=0, \tag{6}$$

$$D\begin{Bmatrix}G\\ w\\ \Phi\\ \theta\end{Bmatrix}=\begin{Bmatrix}0\\0\\0\\0\end{Bmatrix}, \tag{7}$$

where $D$ is a differential operator matrix

$$D=\begin{bmatrix}
c_{11}\Delta+c_{44}\frac{\partial^2}{\partial z^2}-\rho\frac{\partial^2}{\partial t^2} & -(c_{13}+c_{44})\frac{\partial}{\partial z} & -(e_{15}+e_{31})\frac{\partial}{\partial z} & \lambda_{11}\\
-(c_{13}+c_{44})\Delta\frac{\partial}{\partial z} & c_{44}\Delta+c_{33}\frac{\partial^2}{\partial z^2}-\rho\frac{\partial^2}{\partial t^2} & e_{15}\Delta+e_{33}\frac{\partial^2}{\partial z^2} & -\lambda_{33}\frac{\partial}{\partial z}\\
(e_{15}+e_{31})\Delta\frac{\partial}{\partial z} & -\left(e_{15}\Delta+e_{33}\frac{\partial^2}{\partial z^2}\right) & \epsilon_{11}\Delta+\epsilon_{33}\frac{\partial^2}{\partial z^2} & -p_3\frac{\partial}{\partial z}\\
0 & 0 & 0 & k_{11}\Delta+k_{33}\frac{\partial^2}{\partial z^2}-\frac{k_{11}}{\kappa}\frac{\partial}{\partial t}
\end{bmatrix}. \tag{8}$$

It is seen that the displacement function $\psi$ can be solved from Eq. (6) separately. Equation (7) is a homogeneous set of differential equations in $G$, $w$, $\Phi$, and $\theta$. The general solution can be obtained routinely by the operator theory as follows:

$$G=A_{i1}F,\quad w=A_{i2}F,\quad \Phi=A_{i3}F,\quad \theta=A_{i4}F,\quad (i=1,2,3,4), \tag{9}$$

where $A_{ij}$ are the algebraic cominors of the matrix $D$, of which the determinant is

$$\begin{aligned}
|D|=&\left(a\frac{\partial^6}{\partial z^6}+b\Delta\frac{\partial^4}{\partial z^4}+c\Delta^2\frac{\partial^2}{\partial z^2}+d\Delta^3+g\Delta\frac{\partial^4}{\partial t^4}\right.\\
&\left.+h\Delta^2\frac{\partial^2}{\partial t^2}+k\Delta\frac{\partial^2}{\partial z^2}\frac{\partial^2}{\partial t^2}+l\frac{\partial^4}{\partial z^4}\frac{\partial^2}{\partial t^2}+n\frac{\partial^2}{\partial z^2}\frac{\partial^4}{\partial t^4}\right)\\
&\times\left(k_{11}\Delta+k_{33}\frac{\partial^2}{\partial z^2}-\frac{k_{11}}{\kappa}\frac{\partial}{\partial t}\right),
\end{aligned} \tag{10}$$

where $a$, $b$, $c$, $d$, $g$, $h$, $k$, $l$, and $n$ are combinations of material constants that can be found in Ding et al. [35]. The function $F$ in Eq. (9) should satisfy the following homogeneous equation:

$$|D|F=0. \tag{11}$$

It can be seen immediately that if we take $i=1,2,3$ in Eq. (9), we will get three general solutions in which $\theta=0$. These solutions are in fact identical to the ones of piezoelasticity as reported in Ding et al. [35]. Thus, $i=4$ should be taken in Eq. (9) for coupled piezothermoelastic problems so that the following general solution of dynamic problems can be obtained:

$$\begin{gathered} u=\frac{\partial\psi}{\partial y}-A_{41}\frac{\partial F}{\partial x}, \quad v=-\frac{\partial\psi}{\partial x}-A_{41}\frac{\partial F}{\partial y}, \\ w=A_{42}F, \quad \Phi=A_{43}F, \quad \theta=A_{44}F, \end{gathered} \tag{12}$$

where $\psi$ and $F$ are general solutions of Eqs. (6) and (11), respectively.

Substituting the expressions of $A_{4j}$ into Eq. (12), the general solution can be written in the following form:

$$\begin{aligned} u&=\frac{\partial\psi}{\partial y}-\left[a_1\Delta^2+b_1\Delta\frac{\partial^2}{\partial z^2}+c_1\frac{\partial^4}{\partial z^4}+\left(d_1\Delta+f_1\frac{\partial^2}{\partial z^2}\right)\rho\frac{\partial^2}{\partial t^2}\right]\frac{\partial F}{\partial x}, \\ v&=-\frac{\partial\psi}{\partial x}-\left[a_1\Delta^2+b_1\Delta\frac{\partial^2}{\partial z^2}+c_1\frac{\partial^4}{\partial z^4}+\left(d_1\Delta+f_1\frac{\partial^2}{\partial z^2}\right)\rho\frac{\partial^2}{\partial t^2}\right]\frac{\partial F}{\partial y}, \\ w&=\left[a_2\Delta^2+b_2\Delta\frac{\partial^2}{\partial z^2}+c_2\frac{\partial^4}{\partial z^4}+\left(d_2\Delta+f_2\frac{\partial^2}{\partial z^2}\right)\rho\frac{\partial^2}{\partial t^2}\right]\frac{\partial F}{\partial z}, \\ \Phi&=\left[a_3\Delta^2+b_3\Delta\frac{\partial^2}{\partial z^2}+c_3\frac{\partial^4}{\partial z^4}\right. \\ &\quad\left.+\left(d_3\Delta+f_3\frac{\partial^2}{\partial z^2}+p_3\rho\frac{\partial^2}{\partial t^2}\right)\rho\frac{\partial^2}{\partial t^2}\right]\frac{\partial F}{\partial z}, \\ \theta&=\left[a\frac{\partial^6}{\partial z^6}+b\Delta\frac{\partial^4}{\partial z^4}+c\Delta^2\frac{\partial^2}{\partial z^2}+d\Delta^3+g\Delta\frac{\partial^4}{\partial t^4}+h\Delta^2\frac{\partial^2}{\partial t^2}\right. \\ &\quad\left.+k\Delta\frac{\partial^2}{\partial z^2}\frac{\partial^2}{\partial t^2}+l\frac{\partial^4}{\partial z^4}\frac{\partial^2}{\partial t^2}+n\frac{\partial^2}{\partial z^2}\frac{\partial^4}{\partial t^4}\right]F, \end{aligned} \tag{13}$$

where

$$\begin{aligned} a_1&=-\lambda_{11}(c_{44}\epsilon_{11}+e_{15}^2), \\ b_1&=(c_{13}+c_{44})(\lambda_{33}\epsilon_{11}-p_3e_{15})+(e_{15}+e_{31})(\lambda_{33}e_{15}+p_3c_{44}) \\ &\quad-\lambda_{11}(c_{33}\epsilon_{11}+c_{44}\epsilon_{33}+2e_{15}e_{33}), \\ c_1&=(c_{13}+c_{44})(\lambda_{33}\epsilon_{33}-p_3e_{33})+(e_{15}+e_{31})(\lambda_{33}e_{33}+p_3c_{33}) \\ &\quad-\lambda_{11}(e_{33}^2+c_{33}\epsilon_{33}), \\ d_1&=\epsilon_{11}\lambda_{11}, \quad f_1=\lambda_{11}\epsilon_{33}-p_3(e_{15}+e_{31}), \\ a_2&=c_{11}(\lambda_{33}\epsilon_{11}-p_3e_{15})-\lambda_{11}[\epsilon_{11}(c_{13}+c_{44})+e_{15}(e_{15}+e_{31})], \\ b_2&=c_{44}(\lambda_{33}\epsilon_{11}-p_3e_{15})+c_{11}(\lambda_{33}\epsilon_{33}-p_3e_{33})-\lambda_{11}\epsilon_{33}(c_{13}+c_{44}) \\ &\quad+(e_{15}+e_{31})[p_3(c_{13}+c_{44})+\lambda_{33}(e_{15}+e_{31})-\lambda_{11}e_{33}], \\ c_2&=c_{44}(\lambda_{33}\epsilon_{33}-p_3e_{33}), \\ d_2&=e_{15}p_3-\epsilon_{11}\lambda_{33}, \quad f_2=p_3e_{33}-\lambda_{33}\epsilon_{33}, \\ a_3&=c_{11}(\lambda_{33}e_{15}+p_3c_{44})+\lambda_{11}(e_{31}c_{44}-e_{15}c_{13}), \\ b_3&=p_3(c_{11}c_{33}-c_{13}^2-2c_{13}c_{44})+\lambda_{11}[c_{33}(e_{15}+e_{31}) \\ &\quad-e_{33}(c_{13}+c_{44})]-\lambda_{33}[c_{13}(e_{15}+e_{31})+c_{44}e_{31}-c_{11}e_{33}], \\ c_3&=c_{44}(c_{33}p_3+\lambda_{33}e_{33}), \\ d_3&=-p_3(c_{11}+c_{44})-\lambda_{11}(e_{15}+e_{31})-e_{15}\lambda_{33}, \\ f_3&=-p_3(c_{33}+c_{44})-e_{33}\lambda_{33}. \end{aligned} \tag{14}$$

The dynamic general solution in cylindrical coordinates $(r,\phi,z)$ can be readily obtained from Eq. (13). In fact, the expressions for $w$, $\Phi$, and $\theta$ will keep unaltered while $u$ and $v$ should be replaced by $u_r$ and $u_\phi$ as

$$u_r=\frac{\partial\psi}{r\partial\phi}-\left[a_1\Delta^2+b_1\Delta\frac{\partial^2}{\partial z^2}+c_1\frac{\partial^4}{\partial z^4}+\left(d_1\Delta+f_1\frac{\partial^2}{\partial z^2}\right)\rho\frac{\partial^2}{\partial t^2}\right]\frac{\partial F}{\partial r}, \tag{15}$$

$$\begin{aligned} u_\phi&=-\frac{\partial\psi}{\partial r}-\left[a_1\Delta^2+b_1\Delta\frac{\partial^2}{\partial z^2}+c_1\frac{\partial^4}{\partial z^4}\right. \\ &\quad\left.+\left(d_1\Delta+f_1\frac{\partial^2}{\partial z^2}\right)\rho\frac{\partial^2}{\partial t^2}\right]\frac{\partial F}{r\partial\phi}, \end{aligned}$$

in which $\Delta=\partial^2/\partial r^2+\partial/r\partial r+\partial^2/r^2\partial\phi^2$ is the planar Laplacian in cylindrical coordinates.

It can be seen that the dynamic general solution does not include a particular part that satisfies a nonhomogeneous differential equation as the one reported in Ding et al. [30]. This will bring out certain convenience for solving practical problems. Especially in the static case the general solution will take a form more conveniently to be used in potential theory as shown in the following.

## 4 Static General Solution

For static problems, the terms associated with the time in Eqs. (6) and (11) should vanish. In fact, functions $\psi$ and $F$ satisfy the following two equations, respectively,

$$\left(\Delta+\frac{\partial^2}{\partial z_0^2}\right)\psi=0, \tag{16}$$

$$\left(\Delta+\frac{\partial^2}{\partial z_1^2}\right)\left(\Delta+\frac{\partial^2}{\partial z_2^2}\right)\left(\Delta+\frac{\partial^2}{\partial z_3^2}\right)\left(\Delta+\frac{\partial^2}{\partial z_4^2}\right)F=0, \tag{17}$$

where $z_i=s_iz$, $s_0=\sqrt{c_{66}/c_{44}}$, $s_4=\sqrt{k_{11}/k_{33}}$ and $s_i\,(i=1,2,3)$ are three roots of the following algebraic equation with positive real part:

$$as^6-bs^4+cs^2-d=0. \tag{18}$$

Then by virtue of the generalized Almansi' theorem verified by Ding et al. [35], the function $F$ can be expressed in terms of four harmonic functions (more precisely, weighted or quasi-harmonic equations)

$$F=\begin{cases} F_1+F_2+F_3+F_4, & \text{for } s_i\,(i=1,2,3,4) \text{ are distinct}, \\ F_1+F_2+F_3+zF_4, & \text{for } s_1\neq s_2\neq s_3=s_4\neq s_1, \\ F_1+F_2+zF_3+z^2F_4, & \text{for } s_1\neq s_2=s_3=s_4, \\ F_1+zF_2+z^2F_3+z^3F_4, & \text{for } s_1=s_2=s_3=s_4, \end{cases} \tag{19}$$

where $F_i\,(i=1,2,3,4)$ satisfy

$$\left(\Delta+\frac{\partial^2}{\partial z_i^2}\right)F_i=0, \quad (i=1,2,3,4). \tag{20}$$

For the sake of simplicity, only in Eq. (19) in this paper we do not specify $s_4$ to the value of $\sqrt{k_{11}/k_{33}}$; it is merely one of the equal eigenvalues when the later three cases are considered.

We now proceed to consider the case of distinct eigenvalues $s_i\,(i=1,2,3,4)$; when equal eigenvalues appear, the derivation is similar. The general solution as shown in Eq. (13) can be thus written in the following form for static problems:

$$u=\frac{\partial\psi}{\partial y}-\sum_{i=1}^{4}\beta_{i1}\frac{\partial^5F_i}{\partial x\partial z_i^4}, \quad v=-\frac{\partial\psi}{\partial x}-\sum_{i=1}^{4}\beta_{i1}\frac{\partial^5F_i}{\partial y\partial z_i^4}, \tag{21}$$

$$w=\sum_{i=1}^{4}\beta_{i2}s_i\frac{\partial^5 F_i}{\partial z_i^5},\quad \Phi=\sum_{i=1}^{4}\beta_{i3}s_i\frac{\partial^5 F_i}{\partial z_i^5},\quad \theta=\beta_{44}\frac{\partial^6 F_4}{\partial z_4^6},$$

where

$$\beta_{ij}=a_j-b_js_i^2+c_js_i^4,\quad (j=1,2,3),\quad \beta_{44}=as_4^6-bs_4^4+cs_4^2-d. \tag{22}$$

Equation (21) can be further simplified if setting

$$\beta_{i1}\frac{\partial^4 F_i}{\partial z_i^4}=\psi_i,\quad (i=1,2,3,4). \tag{23}$$

Utilizing Eq. (23), and writing $\psi$ as $\psi_0$, the static general solution will take the following simple form:

$$u=\frac{\partial\psi_0}{\partial y}-\sum_{i=1}^{4}\frac{\partial\psi_i}{\partial x},\quad v=-\frac{\partial\psi_0}{\partial x}-\sum_{i=1}^{4}\frac{\partial\psi_i}{\partial y},$$
$$w=\sum_{i=1}^{4}\alpha_{i1}\frac{\partial\psi_i}{\partial z_i},\quad \Phi=\sum_{i=1}^{4}\alpha_{i2}\frac{\partial\psi_i}{\partial z_i},\quad \theta=\sum_{i=1}^{4}\alpha_{i3}\frac{\partial^2\psi_i}{\partial z_i^2}, \tag{24}$$

where $\alpha_{i1}=\beta_{i2}s_i/\beta_{i1}$, $\alpha_{i2}=\beta_{i3}s_i/\beta_{i1}$, $\alpha_{13}=\alpha_{23}=\alpha_{33}=0$, $\alpha_{43}=\beta_{44}/\beta_{41}$ and

$$\left(\Delta+\frac{\partial^2}{\partial z_i^2}\right)\psi_i=0,\quad (i=0,1,\ldots,4). \tag{25}$$

In cylindrical coordinates $(r,\phi,z)$, the radial and circumferential components of the mechanical displacement $u_r$ and $u_\phi$ in the static general solution will read

$$u_r=\frac{\partial\psi_0}{r\partial\phi}-\sum_{i=1}^{4}\frac{\partial\psi_i}{\partial r},\quad u_\phi=-\frac{\partial\psi_0}{\partial r}-\sum_{i=1}^{4}\frac{\partial\psi_i}{r\partial\phi} \tag{26}$$

and the axial component of the mechanical displacement $w$, the electric potential $\Phi$, as well as the temperature increment $\theta$, still take the same form as given in Eq. (24).

It is interesting to compare the present static general solution (Eq. (24)) with those obtained by Ashida et al. [21] and Ding et al. [29]. Whether the problem is static or quasi-static, totally seven potential functions (the electric potential is considered here as a natural variable) are involved in the general solution of Ashida et al. [21]. Basically, the static general solution of Ding et al. [29] is identical with that of Ashida et al. although rigorous mathematical derivation was developed. Ding et al. [29] also expressed the general solution in a more concise form in which only four harmonic functions and one particular function, which satisfies a nonhomogeneous sixth-order partial differential equation, are involved. It is obvious that the present static general solution expressed by five harmonic functions, as given in Eq. (24), is simpler than both of Ashida et al. [21] and Ding et al. [29]. More importantly, the general solution in the present form bridges the gap between some boundary value problems usually met in problems associated with cracks and punches ([36]) and the potential theory ([32]). The problem of a flat crack opened out by the application of a prescribed temperature to its faces in an infinite piezothermoelastic space is to be considered immediately to show the superiority of the present general solution.

For the convenience of the followed analysis, the following complex quantities are introduced:

$$U=u+iv,\quad \sigma_1=\sigma_x+\sigma_y,\quad \sigma_2=\sigma_x-\sigma_y+2i\tau_{xy},$$
$$\tau_z=\tau_{xz}+i\tau_{yz},\quad D=D_x+iD_y.$$

Utilizing Eqs. (24), (1), and (2), it is obtained that

$$U=-\Lambda\left(\sum_{i=1}^{4}\psi_i+i\psi_0\right),\quad \sigma_z=\sum_{i=1}^{4}\gamma_{i1}\frac{\partial^2\psi_i}{\partial z_i^2},$$
$$D_z=\sum_{i=1}^{4}\gamma_{i2}\frac{\partial^2\psi_i}{\partial z_i^2},$$
$$\sigma_1=\sum_{i=1}^{4}\gamma_{i3}\frac{\partial^2\psi_i}{\partial z_i^2},\quad \sigma_2=-2c_{66}\Lambda^2\left(\sum_{i=1}^{4}\psi_i+i\psi_0\right), \tag{27}$$
$$\tau_z=\Lambda\left(\sum_{i=1}^{4}\omega_{i1}\frac{\partial\psi_i}{\partial z_i}-is_0c_{44}\frac{\partial\psi_0}{\partial z_0}\right),$$
$$D=\Lambda\left(\sum_{i=1}^{4}\omega_{i2}\frac{\partial\psi_i}{\partial z_i}-is_0e_{15}\frac{\partial\psi_0}{\partial z_0}\right),$$

where

$$\gamma_{i1}=c_{13}+c_{33}s_i\alpha_{i1}+e_{33}s_i\alpha_{i2}-\lambda_{33}\alpha_{i3},$$
$$\gamma_{2i}=e_{31}+e_{33}s_i\alpha_{i1}-\epsilon_{33}s_i\alpha_{i2}+p_3\alpha_{i3}, \tag{28}$$
$$\gamma_{i3}=2[(c_{11}-c_{66})+c_{13}s_i\alpha_{i1}+e_{31}s_i\alpha_{i2}-\lambda_{11}\alpha_{i3}],$$
$$\omega_{i1}=-c_{44}s_i+c_{44}\alpha_{i1}+e_{15}\alpha_{i2},\quad \omega_{i2}=-e_{15}s_i+e_{15}\alpha_{i1}-\epsilon_{11}\alpha_{i2}.$$

Also noted are the following two identities:

$$\gamma_{i1}s_i=\omega_{i1},\quad \gamma_{i2}s_i=\omega_{i2}. \tag{29}$$

## 5 Application: Symmetric Crack Problem

Consider a transversely isotropic piezothermoelastic space weakened by a flat crack $S$ located in the plane $z=0$, with arbitrary temperature $\Theta(x,y)$ applied over the upper and lower crack faces. By virtue of the symmetric condition with respect to the crack surface ([5,37]) the problem is equivalent to a mixed boundary value problem for the half-space $z\geqslant 0$, with the following mixed boundary conditions on the plane $z=0$:

$$\sigma_z=0,\quad D_z=0,\quad \theta=\Theta(x,y),\quad \text{for } (x,y)\in S;$$
$$w=\Phi=\partial\theta/\partial z=0,\quad \text{for } (x,y)\notin S; \tag{30}$$
$$\tau_z=0,\quad \text{for } -\infty<(x,y)<\infty.$$

It is noted here that, in Shang et al. [31] where the problem of a penny-shaped crack with uniform temperature surface was solved by means of potential functions and Fourier-Hankel transformations, the thermal condition outside the crack was taken to be $\theta=0$, which turns to be not a natural condition ([38]).

By extending the potential theory for elasticity ([32]) and for piezoelasticity ([33]), it is assumed that,

$$\psi_0=0;\quad \psi_i(z)=h_{i1}H_1(z_i)+h_{i2}H_2(z_i)+h_{i3}H_3(z_i),$$
$$(i=1,2,3,4), \tag{31}$$

where $h_{ij}$ are undetermined constants, and

$$H_1(r,\phi,z)=\int\int_S\frac{\omega(N)}{R(M,N)}dS,$$
$$H_2(r,\phi,z)=\int\int_S\frac{\varphi(N)}{R(M,N)}dS, \tag{32}$$
$$H_3(r,\phi,z)=\int\int_S\vartheta(N)\{z\ln[R(M,N)+z]-R(M,N)\}dS,$$

in which $\omega$, $\varphi$, and $\vartheta$ stand for the crack-face displacement $w(x,y,0)$, electric potential $\Phi(x,y,0)$, and temperature gradient $\partial\theta(x,y,z)/\partial z|_{z=0}$, respectively, $R(M,N)$ is the distance between the points $M(r,\phi,z)$ and $N(r_0,\phi_0,0)$ [hereafter, cylindrical coordinates $(r,\phi,z)$ are alternatively used for the sake of convenience], and the integration is taken over the crack surface $S$. In

contrast to piezoelasticity ([5]), a new potential $H_3$ has been introduced here to account for the pyroelectric effect. To satisfy the zero-shear stress condition in Eq. (30), one can assume

$$\sum_{i=1}^{4} \omega_{i1} h_{ij} = 0, \quad (j=1,2,3). \tag{33}$$

It is shown that the property of the potential of a simple layer immediately leads to the satisfaction of the vanishing of $w$, $\Phi$, and $\partial\theta/\partial z$ outside the crack, while inside the crack one has

$$\left.\frac{\partial H_1}{\partial z}\right|_{z=0} = -2\pi\omega = -2\pi w(x,y,0),$$
$$\left.\frac{\partial H_2}{\partial z}\right|_{z=0} = -2\pi\varphi = -2\pi\Phi(x,y,0), \tag{34}$$
$$\left.\frac{\partial^3 H_3}{\partial z^3}\right|_{z=0} = -2\pi\vartheta = -2\pi\left.\frac{\partial\theta(x,y,z)}{\partial z}\right|_{z=0}.$$

The following relations then are obtained from Eqs. (24), (31), and (34):

$$\sum_{i=1}^{4} \alpha_{i1} h_{i1} = -\frac{1}{2\pi}, \quad \sum_{i=1}^{4} \alpha_{i1} h_{i2} = 0, \quad \sum_{i=1}^{4} \alpha_{i1} h_{i3} = 0,$$
$$\sum_{i=1}^{4} \alpha_{i2} h_{i1} = 0, \quad \sum_{i=1}^{4} \alpha_{i2} h_{i2} = -\frac{1}{2\pi}, \quad \sum_{i=1}^{4} \alpha_{i2} h_{i3} = 0, \tag{35}$$
$$\sum_{i=1}^{4} \alpha_{i3} h_{i1} = 0, \quad \sum_{i=1}^{4} \alpha_{i3} h_{i2} = 0, \quad \sum_{i=1}^{4} \alpha_{i3} h_{i3} = -\frac{1}{2\pi}.$$

Now $h_{ij}$ are then solved from Eqs. (33) and (35) as follows:

$$\begin{Bmatrix} h_{1j} \\ h_{2j} \\ h_{3j} \\ h_{4j} \end{Bmatrix} = -\frac{1}{2\pi} \begin{bmatrix} \omega_{11} & \omega_{21} & \omega_{31} & \omega_{41} \\ \alpha_{11} & \alpha_{21} & \alpha_{31} & \alpha_{41} \\ \alpha_{12} & \alpha_{22} & \alpha_{32} & \alpha_{42} \\ \alpha_{13} & \alpha_{23} & \alpha_{33} & \alpha_{43} \end{bmatrix}^{-1} \begin{Bmatrix} 0 \\ \delta_{1j} \\ \delta_{2j} \\ \delta_{3j} \end{Bmatrix},$$
$$(j=1,2,3), \tag{36}$$

where $\delta_{ij}$ is the Kronecker delta.

The satisfaction of the first condition in Eq. (30) gives the following integro-differential equations:

$$g_{11}\Delta \int\int_S \frac{\omega(N)}{R(N_0,N)} dS + g_{12}\Delta \int\int_S \frac{\varphi(N)}{R(N_0,N)} dS$$
$$- g_{13} \int\int_S \frac{\vartheta(N)}{R(N_0,N)} dS = 0,$$
$$g_{21}\Delta \int\int_S \frac{\omega(N)}{R(N_0,N)} dS + g_{22}\Delta \int\int_S \frac{\varphi(N)}{R(N_0,N)} dS \tag{37}$$
$$- g_{23} \int\int_S \frac{\vartheta(N)}{R(N_0,N)} dS = 0,$$

$$\int\int_S \frac{\vartheta(N)}{R(N_0,N)} dS = -2\pi\Theta(N_0), \tag{38}$$

where $g_{kj} = \Sigma_{i=1}^{4} \gamma_{ik} h_{ij}$ $(k=1,2;\ j=1,2,3)$, $R(N_0,N)$ represents the distance between two points $N_0$ and $N$, and both $N_0$, $N \in S$.

Making use of Eq. (38), Eq. (37) can be written as

$$(g_{13}g_{22} - g_{23}g_{12})\Theta(N_0) = -\frac{1}{4\pi^2 A}\Delta \int\int_S \frac{\omega(N)}{R(N_0,N)} dS, \tag{39}$$

$$(g_{11}g_{23} - g_{21}g_{13})\Theta(N_0) = -\frac{1}{4\pi^2 A}\Delta \int\int_S \frac{\varphi(N)}{R(N_0,N)} dS, \tag{40}$$

where $A = 1/2\pi(g_{11}g_{22} - g_{12}g_{21})$. For a flat crack with an arbitrary shape, the integral Eq. (38) and the integro-differential Eqs. (39) and (40) generally can be solved by numerical methods. After $\omega$, $\varphi$, and $\vartheta$ are known, the functions $H_i$ can be obtained from Eq. (32) by integration and then the whole piezothermoelastic field can be determined.

For a penny-shaped crack, with the radius being $a_0$, exact solutions to Eqs. (38), (39), and (40) can be obtained by using Fabrikant's results ([32]). In the case when the temperature is uniformly distributed, i.e., $\Theta = \Theta_0 = \text{const.}$, the solutions are

$$\vartheta(r,\phi) = -\frac{2\Theta_0}{\pi\sqrt{a_0^2 - r^2}}, \tag{41}$$

$$\omega(r,\phi) = 4A(g_{13}g_{22} - g_{23}g_{12})\Theta_0\sqrt{a_0^2 - r^2}, \tag{42}$$

$$\varphi(r,\phi) = 4A(g_{11}g_{23} - g_{21}g_{13})\Theta_0\sqrt{a_0^2 - r^2}. \tag{43}$$

Substituting Eqs. (41)–(43) into Eq. (32) yields

$$H_1(r,\phi,z) = 2\pi A(g_{13}g_{22} - g_{23}g_{12})\Theta_0 \left[ (2a_0^2 + 2z^2 - r^2) \times \sin^{-1}\left(\frac{a_0}{l_2}\right) - \frac{2a_0^2 - 3l_1^2}{a_0}\sqrt{l_2^2 - a_0^2} \right],$$

$$H_2(r,\phi,z) = 2\pi A(g_{11}g_{23} - g_{21}g_{13})\Theta_0 \left[ (2a_0^2 + 2z^2 - r^2) \times \sin^{-1}\left(\frac{a_0}{l_2}\right) - \frac{2a_0^2 - 3l_1^2}{a_0}\sqrt{l_2^2 - a_0^2} \right], \tag{44}$$

$$H_3(r,\phi,z) = -2\Theta_0 \left\{ \left(z^2 - a_0^2 - \frac{r^2}{2}\right)\sin^{-1}\left(\frac{a_0}{l_2}\right) - \frac{3(2a_0^2 - l_1^2)}{2a_0}\sqrt{l_2^2 - a_0^2} + 2a_0 z \ln[l_2 + \sqrt{l_2^2 - r^2}] \right\},$$

where

$$l_1 = \frac{1}{2}\{\sqrt{(r+a_0)^2 + z^2} - \sqrt{(r-a_0)^2 + z^2}\},$$

$$l_2 = \frac{1}{2}\{\sqrt{(r+a_0)^2 + z^2} + \sqrt{(r-a_0)^2 + z^2}\}.$$

Having the explicit expressions for the harmonic functions $H_i$ $(i=1,2,3)$ as shown in Eq. (44), the piezothermoelastic field can be obtained simply by differentiation. It is obvious that all the resulting expressions will be in terms of elementary functions. Since the derivation is straightforward, it is omitted here to save the space of the paper. In order to investigate the singular behavior of the piezothermoelastic field near the crack tip, however, the expressions for $\sigma_z$ and $D_z$ are given as follows:

$$\sigma_z = 8\pi A\Theta_0 \sum_{i=1}^{4} \gamma_{i1}[h_{i1}(g_{13}g_{22}-g_{23}g_{12})+h_{i2}(g_{11}g_{23}-g_{21}g_{13})] \times\left[\sin^{-1}\left(\frac{a_0}{l_{2i}}\right)-\frac{a_0\sqrt{l_{2i}^2-a_0^2}}{l_{2i}^2-l_{1i}^2}\right]-4\Theta_0\sum_{i=1}^{4}\gamma_{i1}h_{i3}\sin^{-1}\left(\frac{a_0}{l_{2i}}\right), \tag{45}$$

$$D_z = 8\pi A\Theta_0 \sum_{i=1}^{4} \gamma_{i2}[h_{i1}(g_{13}g_{22}-g_{23}g_{12})+h_{i2}(g_{11}g_{23}-g_{21}g_{13})] \times\left[\sin^{-1}\left(\frac{a_0}{l_{2i}}\right)-\frac{a_0\sqrt{l_{2i}^2-a_0^2}}{l_{2i}^2-l_{1i}^2}\right]-4\Theta_0\sum_{i=1}^{4}\gamma_{i2}h_{i3}\sin^{-1}\left(\frac{a_0}{l_{2i}}\right),$$

where

$$l_{1i}=\frac{1}{2}\{\sqrt{(r+a_0)^2+z_i^2}-\sqrt{(r-a_0)^2+z_i^2}\},$$

$$l_{2i}=\frac{1}{2}\{\sqrt{(r+a_0)^2+z_i^2}+\sqrt{(r-a_0)^2+z_i^2}\}.$$

If the stress and electric displacement intensity factors are defined here in the following way,

$$k_\sigma=\lim_{r\to a_0}\{(r-a_0)^{1/2}\sigma_z|_{z=0}\},\quad k_D=\lim_{r\to a_0}\{(r-a_0)^{1/2}D_z|_{z=0}\}, \tag{46}$$

then from Eq. (45), we obtain

$$k_\sigma=-2\sqrt{2a_0}g_{13}\Theta_0,\quad k_D=-2\sqrt{2a_0}g_{23}\Theta_0. \tag{47}$$

It is then concluded that both the stress and electric displacement intensity factors depend on material constants for thermal loading, just as in the elastic case ([37]).

## 6 Conclusions

A dynamic general solution is first presented for problems of transversely isotropic piezothermoelastic media by using the operator theory. For static problems the solution is further simplified and expressed in terms of five quasi-harmonic functions. Comparison with that obtained by Ding et al. [29] shows that the present general solution is simpler and can be used more conveniently. More importantly, the present solution allows one to use feasibly the existent results in potential theory to analyze boundary value problems in piezothermoelasticity. It is also noted here that the general solution of Ding et al. [29] could not account for the case of equal eigenvalues involving $s_4$.

The symmetric crack problem of a prescribed temperature applied over the faces of a flat crack in an infinite piezothermoelastic solid is considered as an illustrative example. The problem is first transformed into a boundary value problem of a half-space by using the symmetric property. The five quasi-harmonic functions in the general solution are in turn expressed in terms of three potentials. The governing integro-differential equations of the problem are established from boundary conditions on the surface. Exact solutions to the resulting integro-differential equations are given, for instance, for a penny-shaped crack subjected to a uniform temperature to its faces. Explicit expressions for the stress and electric displacement intensity factors are then presented. It is found that both the factors are dependent on the material constants.

Using the present general solution, the static as well as dynamic behaviors of piezothermoelastic plates and shells can also be investigated. Further developments in this respect will be reported in other papers.


## Acknowledgment

The author is thankful for the project sponsored by the Scientific Research Foundation for the Returned Overseas Chinese Scholars, State Education Ministry, and is also indebted to Prof. Haojiang Ding for his helpful comments on the topic.


## References


[1] Tani, J., Takagi, T., and Qiu, J. H., 1998, ''Intelligent Material Systems: Application of Functional Materials,'' ASME Appl. Mech., **51**, pp. 505–521.

[2] Pak, Y. E., 1990, ''Crack Extension Force in a Piezoelectric Material,'' ASME J. Appl. Mech., **57**, pp. 647–653.

[3] Bisegna, P., and Maceri, F., 1996, ''An Exact Three-Dimensional Solution for Simply Supported Rectangular Piezoelectric Plates,'' ASME J. Appl. Mech., **63**, pp. 628–638.

[4] Chen, W. Q., 1999, ''Problems of Radially Polarized Piezoelastic Bodies,'' Int. J. Solids Struct., **36**, pp. 4317–4332.

[5] Chen, W. Q., and Shioya, T., 1999, ''Fundamental Solution for a Penny-Shaped Crack in a Piezoelectric Medium,'' J. Mech. Phys. Solids, **47**, pp. 1459–1475.

[6] Chen, W. Q., Shioya, T., and Ding, H. J., 1999, ''The Elasto-Electric Field for a Rigid Conical Punch on a Transversely Isotropic Piezoelectric Half-Space,'' ASME J. Appl. Mech., **66**, pp. 764–771.

[7] Nowinski, J. L., 1978, *Theory of Thermoelasticity With Applications*, Sijthoff and Noordhoff, Alpen aan den Rijn, The Netherlands.

[8] Tiersten, H. F., 1971, ''On the Nonlinear Equations of Thermoelectroelasticity,'' Int. J. Eng. Sci., **9**, pp. 587–604.

[9] Nowacki, W., 1978, ''Some General Theorems of Thermopiezoelectricity,'' J. Therm. Stresses, **1**, pp. 171–182.

[10] Nowacki, J. P., 1982, ''Steady-State Problems of Thermopiezoelectricity,'' J. Therm. Stresses, **5**, pp. 183–194.

[11] Chandrasekharaiah, D. S., 1988, ''A Generalized Linear Thermoelasticity Theory for Piezoelectric Media,'' Acta Mech., **71**, pp. 39–49.

[12] Iesan, D., 1989, ''On Some Theorems in Thermopiezoelectricity,'' J. Therm. Stresses, **12**, pp. 209–222.

[13] Tauchert, T. R., 1992, ''Piezothermoelastic Behavior of a Laminated Plate,'' J. Therm. Stresses, **15**, pp. 25–37.

[14] Dube, G. P., Kapuria, S., and Dumir, P. C., 1996, ''Exact Piezothermoelastic Solution of Simply-Supported Orthotropic Circular Cylindrical Panel in Cylindrical Bending,'' Arch. Appl. Mech., **66**, pp. 537–554.

[15] Lee, H. J., and Saravanos, D. A., 1996, ''Coupled Layerwise Analysis of Thermopiezoelectric Composite Beams,'' AIAA J., **34**, pp. 1231–1237.

[16] Stam, M., and Carman, G., 1996, ''Electrothermoelastic Behavior of Piezoelectric Coupled Cylinders,'' AIAA J., **34**, pp. 1612–1618.

[17] Tang, Y. Y., Noor, A. K., and Xu, K., 1996, ''Assessment of Computational Models for Thermoelectroelastic Multilayered Plates,'' Comput. Struct., **61**, pp. 915–923.

[18] Xu, K., and Noor, A. K., 1996, ''Three-Dimensional Analytical Solutions for Coupled Thermoelectroelastic Response of Multilayered Cylindrical Shells,'' AIAA J., **34**, pp. 802–812.

[19] Ashida, F., Tauchert, T. R., and Noda, N., 1993, ''Response of a Piezothermoelastic Plate of Crystal Class 6 mm Subject to Axisymmetric Heating,'' Int. J. Eng. Sci., **31**, pp. 373–384.

[20] Ashida, F., Noda, N., and Tauchert, T. R., 1994, ''A Two-Dimensional Piezothermoelastic Problem in an Orthotropic Plate Exhibiting Crystal Class 2 mm,'' JSME Int. J., **37**, pp. 334–340.

[21] Ashida, F., Tauchert, T. R., and Noda, N., 1994, ''A General Solution Technique for Piezothermoelasticity of Hexagonal Solids of Class 6 mm in Cartesian Coordinates,'' Z. Angew. Math. Mech., **74**, pp. 87–95.

[22] Ashida, F., Tauchert, T. R., and Noda, N., 1994, ''Potential Function Method for Piezothermoelastic Problems of Solids of Crystal Class 6 mm in Cylindrical Coordinates,'' J. Therm. Stresses, **17**, pp. 361–375.

[23] Ashida, F., Noda, N., and Tauchert, T. R., 1994, ''Inverse Problem of Two-Dimensional Piezothermoelasticity in an Orthotropic Plate Exhibiting Crystal Class mm 2,'' JSME Int. J., **37**, pp. 341–346.

[24] Ashida, F., Choi, J., and Noda, N., 1997, ''Control of Elastic Displacement in Piezoelectric-Based Intelligent Plate Subjected to Thermal Load,'' Int. J. Eng. Sci., **35**, pp. 851–868.

[25] Ashida, F., and Tauchert, T. R., 1997, ''Temperature Determination for a Contacting Body Based on an Inverse Piezothermoelastic Problem,'' Int. J. Solids Struct., **34**, pp. 2549–2561.

[26] Ashida, F., and Tauchert, T. R., 1998, ''Transient Response of a Piezothermoelastic Circular Disk Under Axisymmetric Heating,'' Acta Mech., **128**, pp. 1–14.

[27] Kapuria, S., Dumir, P. C., and Sengupta, S., 1996, ''Exact Piezothermoelastic Axisymmetric Solution of a Finite Transversely Isotropic Cylindrical Shell,'' Comput. Struct., **61**, pp. 1085–1099.

[28] Kapuria, S., Dumir, P. C., and Segupta, S., 1998, ''Three-Dimensional Axisymmetric Piezothermoelastic Solution of a Transversely Isotropic Piezoelectric Clamped Circular Plate,'' ASME J. Appl. Mech., **65**, pp. 178–183.

[29] Ding, H. J., Guo, F. L., Hou, P. F., and Chi, Y. W., 1999, ''A General Solution for Static Piezothermoelastic Problems of Crystal Class 6 mm Solids,'' Acta Mech. Solida Sinica, **12**, pp. 189–197.

[30] Ding, H. J., Guo, F. L., and Hou, P. F., 2000, ''A General Solution for Piezo-

thermoelasticity of Transversely Isotropic Piezoelectric Materials and Its Applications,'' Int. J. Eng. Sci., **38**, pp. 1415–1440.
[31] Shang, F. L., Wang, Z. K., and Li, Z. H., 1996, ''Thermal Stresses Analysis of a Three-Dimensional Crack in a Thermopiezoelectric Solid,'' Eng. Fract. Mech., **55**, pp. 737–750.
[32] Fabrikant, V. I., 1989, *Applications of Potential Theory in Mechanics: A Selection of New Results*, Kluwer Academic Publishers, The Netherlands.
[33] Chen, W. Q., 1999, ''On the Application of Potential Theory in Piezoelasticity,'' ASME J. Appl. Mech., **66**, pp. 808–811.
[34] Tiersten, H. F., 1969, *Linear Piezoelectric Plate Vibrations*, Plenum Press, New York.
[35] Ding, H. J., Chen, B., and Liang, J., 1996, ''General Solutions for Coupled Equations for Piezoelectric Media,'' Int. J. Solids Struct., **33**, pp. 2283–2298.
[36] Sneddon, I. N., 1966, *Mixed Boundary Value Problems in Potential Theory*, North-Holland, Amsterdam.
[37] Tsai, Y. M., 1983, ''Thermal Stress in a Transversely Isotropic Medium Containing a Penny-Shaped Crack,'' ASME J. Appl. Mech., **50**, pp. 24–28.
[38] Sneddon, I. N., and Lowengrub, M., 1969, *Crack Problems in the Classical Theory of Elasticity*, John Wiley and Sons, New York.

# Dynamic Wetting: Displacement of the Contact Line Induced by Air Pressure Oscillations in Its Close Vicinity

**P. Woehl**[1]
e-mail: woehlp@corning.com

**P. Bourgin**[2]
Mem. ASME
e-mail: bourgin@esp-oyonnax.fr

Université Louis Pasteur,
Institut de Mécanique des Fluides,
2, rue Boussingault,
67000 Strasbourg, France

*In most studies devoted to dynamic wetting, the viscosity of the ambient air is neglected compared with that of the liquid, which equivalently means that the surrounding air is considered to be static. However, the presence of air flows underneath the sheet leads to an increase of the pressure in the wedge formed by the liquid sheet and the solid substrate. The pressure deforms the interface and tends to separate the liquid sheet from the solid substrate. This is a coupled problem, the two unknown functions being the meniscus shape and the pressure generated in the air wedge. A simple model is proposed here to investigate the possible influence of slight pressure fluctuations in the air on the displacement of the contact line. This model incorporates the liquid wetting properties as well as parameters characteristic of the process and of the air pressure oscillations. The main result is that small amplitude pressure oscillations in the air close to the contact line may generate large lateral displacements of the contact line. At the same time, the air gap is reduced as compared to its value in the static case.* [S0021-8936(00)01904-8]

## 1 Introduction

Numerous experimental studies concerned with the mechanics of air entrainment in dynamic wetting have been reported by several authors since the pioneering work of Deryagin and Levi in 1964. It is now well established that one of the limiting factors in high-velocity coating flows is air entrainment into the wedge between the moving solid substrate and the liquid layer. Burley and Jolly [1] proposed several correlations between the process parameters and the onset of air entrainment and extensively described the complex mechanisms involved ([2,3]). After precursor phenomena, such as unsteady and serrated meniscus, the apparent contact angle reaches 180 deg at the onset of air entrainment, as observed by Gutoff and Kendrick [4] or Miyamoto and Scriven [5]. The visualization of coating flows is very difficult due to small dimensions in the vicinity of the three-phase junction: for example, Mues et al. [6] found by using laser Doppler velocimetry that when the capillary number is greater than 0.25, the observed contact line becomes apparent and a thin air layer is entrained between the coated liquid and the moving surface. Qualitative observations by means of a high-speed camera confirmed this tendency: see, for instance, Veverka and Aidun [7,8] and Miyamoto [9]. These studies clearly show that the dynamic effects of the ambient air in the vicinity of the contact line should not be neglected, even in the case of no air entrainment.

The theoretical counterpart of these studies is far from being so rich. In other words, there are only few theoretical studies dealing with the microfluid mechanics of air entrainment in the vicinity of the dynamic wetting line since the first analysis of Huh and Scriven [10] considering Stokes flow in the vicinity of the triple line. The key idea to consider here is that the pressure forces generated in the air wedge are large enough to lift up the meniscus in the close vicinity of the contact line. The amount of air entrained can be predicted by projecting the contact line to infinity (Emonot [11], Emonot et al. [12], and Bourgin and Tahiri [13]). In such conditions, an analytical solution can be achieved through asymptotic expansions, leading to a good agreement with numerical predictions (Tahiri [14], Tahiri et al. [15], or Bourgin and Saintlos [16]).

Here, a more realistic model is proposed to consider the role of air pressure by linking the free surface to the moving surface. Numerical predictions show the influence of slight pressure fluctuations in the air in the vicinity of the contact line on the motion of the triple line.

## 2 Flow Configuration and Governing Equations

Neglecting side-leakage effects, the basic configuration is that of a two-dimensional liquid sheet flowing onto a moving solid substrate, as shown in Fig. 1. The flow rate at the outlet of the ejector of opening $e_0$ is prescribed, as well as the translation velocity of the solid ($U$). Since the average ejection velocity is lower than $U$, the liquid sheet accelerates and is stretched so that its current thickness ($e$) decreases from the die opening down to the layer thickness. The challenge is to guarantee a uniform value for the coated layer, by controlling the process key parameters.

This is a free-surface problem, the shape of the liquid sheet (and consequently the final thickness of the coated layer) being unknown as well as the velocity distribution along and throughout the sheet. Far upstream of the three-phase contact line, complex recirculating flows take place, but without any significant pressure increase: The ambient air does not affect the shape of the liquid sheet surface and the classical free-surface boundary conditions can be applied at the liquid/air interface. Close to the three-phase junction, some air is entrapped in the wedge between the solid substrate and the liquid sheet and, as a result, the pressure increases, which deforms the interface and tends to separate the liquid sheet from the solid substrate. This is a coupled problem, the two unknown functions being the meniscus shape (or equivalently the air gap thickness) and the pressure generated in the air gap. The effects of the pressure fluctuations at the inlet of the

[1]Present address: Corning S.A., Centre Européen de Recherche de Fontainebleau, 7bis avenue de Valvins, 77210 Avon, France.

[2]Present address: Ecole Supérieure de Plasturgie, BP 807 01108 Oyonnax, France.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, December 21, 1998; final revision, May 22, 2000. Associate Technical Editor: D. A. Siginer. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

 **Copyright © 2000 by ASME** 

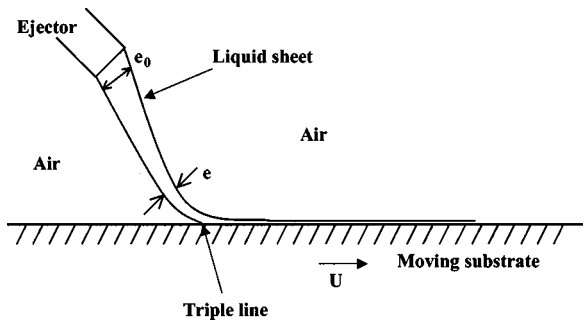



**Fig. 1 Flow configuration**

contacting zone are studied in this paper. Figure 2 represents a magnification of the contacting zone, the interface represented by a parabola. The air gap $h(x,t)$ which corresponds to the interface shape is given by the following relationship, the radius of curvature $R$ being the curvature of the liquid sheet:

$$h(x,t)=h_{\min}(t)+\frac{x^2}{2R}. \tag{1}$$

The radius $R$ is here supposed to be constant as observed in Levarlet and Finnicum [17] where the authors have performed a perturbation in the air upstream the triple line in a bead coating configuration and visualized no deformation of this radius. This meniscus portion is connected to the solid wall through an arc of a circle of radius $\rho_c$ and the following relationship holds:

$$\rho_c=\frac{h_{\min}}{2\cos\theta}. \tag{2}$$

The pressure is governed by the time-dependent Reynolds equation:

$$\frac{\partial}{\partial x}\left(\frac{h^3(x,t)}{12\eta}\frac{\partial P}{\partial x}-Uh(x,t)\right)=\frac{\partial h(x,t)}{\partial t}. \tag{3}$$

The contact zone considered here ranges between abscissas $x_e$ (inlet) and $x_s$ (outlet). The location of $x_e$ is a difficult question. Several points of view can be proposed: The inlet location can be considered either as the first point where the pressure in the air starts to rise and subsequently affects the meniscus shape, or as the last point of the upstream zone. Whatever the choice may be, a complete study of the upstream zone is necessary, which is beyond the scope of the present paper. Therefore, $x_e$ will be considered here as a given parameter. The outlet location $x_s$ corresponds to the three-phase junction, i.e., the attachment point of the meniscus to the substrate.

Clearly, the value of this abscissa $x_s$ is expected to be affected by pressure fluctuations at the inlet. However, it is assumed as a first step that the meniscus deformations reduce to vertical translations of a rigid solid which means that the unknowns are now the pressure field and the distance between the meniscus and the solid substrate. Therefore, the outlet location will be prescribed and taken equal to zero without loss of generality, which corresponds to assuming that the parabola has a horizontal tangent for $x=x_s$ (i.e., $h_{\min}=h_s$). A second step will consist of considering the outlet location as an unknown and then studying the lateral motion of the triple line due to pressure oscillations at the inlet.

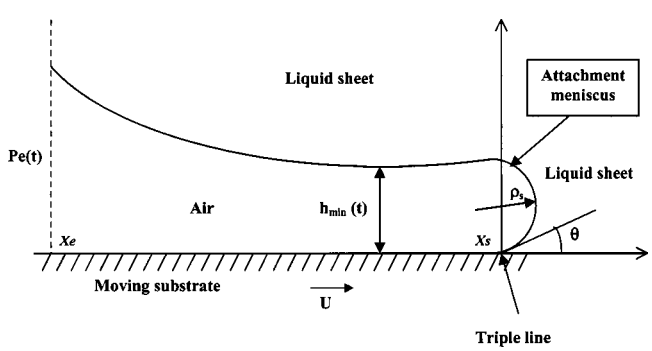



**Fig. 2 Magnification of the contacting zone**

## 3 First Step: Vertical Motion of the Meniscus Over the Triple Line

**3.1 Boundary Conditions.** The boundary conditions are as follows:

$$P(x=x_e)=P_m\sin\omega t,\quad P(x=0)=\frac{\sigma}{\rho_c}. \tag{4}$$

The first equation expresses that the inlet pressure oscillates (amplitude $P_M$ and frequency $\nu=\omega/2\pi$ are given) and the second one is the Laplace equation. The initial condition for $h_s(t)$ is the thickness value for the steady case (with the steady Reynolds equation):

$$h_s(t=0)=h_0. \tag{5}$$

The steady case problem has a closed-form solution which allows $h_0$ to be expressed as a function of the governing parameters by balancing the resulting of the pressure forces with the external ''pinning'' force ($W$):

$$\frac{W}{\sigma}=\frac{8\mathrm{Ca}\,R}{h_0}\left(\frac{1}{2}+\frac{4\cos\theta}{3\pi\,\mathrm{Ca}\sqrt{2R}}\sqrt{h_0}\right) \tag{6}$$

where Ca is a capillary number built on the viscosity of air ($\mathrm{Ca}=\eta U/\sigma$) and $W/\sigma$ represents a dimensionless number (ratio between external—gravity, . . . —and surface tension forces).

**3.2 Solution.** The problem defined by Eqs. (2)–(5) leads to a straightforward analytical solution for $P(x,t)$ as well as for its integrated value (loading force $W$). Assuming that $W$ is maintained constant in time and introducing the change of variables (Woehl [18]),

$$H(t)=\frac{h_s(t)}{h_0}, \tag{7}$$

the governing ordinary differential equation for $H$ reads:[3]

$$A\frac{1}{H^{3/2}}\frac{dH}{dt}=\frac{B}{H}+C(t)\sqrt{H}+\frac{D}{\sqrt{H}}+E(t) \tag{8}$$

where $A, B, \ldots, E$ are known functions given by

$$A=\frac{6\eta R\sqrt{2R}}{\sqrt{h_0}}\left(\frac{\pi}{4}-\frac{4}{3\pi}\right) \tag{9}$$

$$B=\frac{4\eta UR}{h_0} \tag{10}$$

$$C(t)=-\frac{4\sqrt{2R}}{3\pi}P_m\sin(\omega t)\sqrt{h_0} \tag{11}$$

$$D=\frac{16\sqrt{2R}}{3\pi\sqrt{h_0}}\sigma\cos\theta \tag{12}$$

$$E(t)=-W-x_eP_m\sin(\omega t). \tag{13}$$

[3]As Eq. (1) shows that the height function $h(x,t)$ is the sum of a function of time and a function of space, the unsteady Reynolds equation leads to an equation with the unique space variable $x$ which is easily solved.

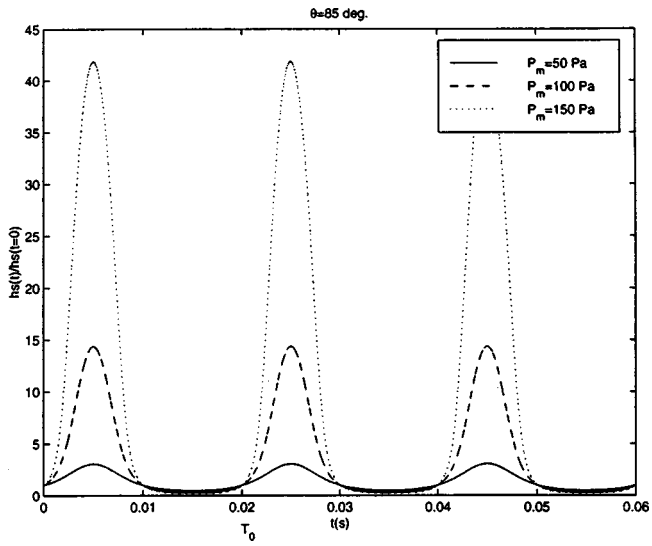



**Fig. 3 Influence of pressure amplitude on relative vertical displacement**

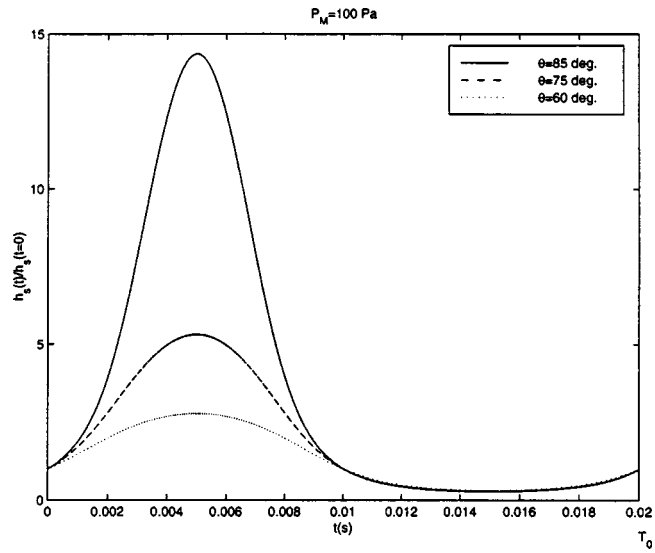



**Fig. 5 Influence of wetting angle on relative vertical displacement**

Equation (8) has been solved by a Runge-Kutta method of the fourth order, for the following set of data: $x_e = -1$ mm; $R = 1$ mm; $\eta = 18.10^{-6}$ Pa.s. The overall dimensions of the contacting zone are prescribed. Provided that $x_e$ be sufficiently large compared to the air gap thickness, it does not affect the results.

**3.3 Results.** Figure 3 shows the effect of the amplitude of the pressure oscillations at the inlet on the meniscus vertical displacement, the velocity being equal to 1 m/s, the frequency to 50 Hz, the wetting angle to 85 deg, the loading force to 0.1 N/m, and the surface tension to $50.10^{-3}$ N/m.

The response to a pressure sine signal is not sinusoidal but amplifies the gap between the liquid sheet and the solid wall, which tends to separate the meniscus from the solid substrate. This trend drastically develops as the pressure amplitude increases, although the absolute value of the latter remains quite small (150 Pa). It can be seen in Fig. 4 that when the nominal substrate velocity increases, the previous effects are still enhanced. A way to improve the quality of coating by reducing the fluctuations at the triple line is to force the liquid layer into con-

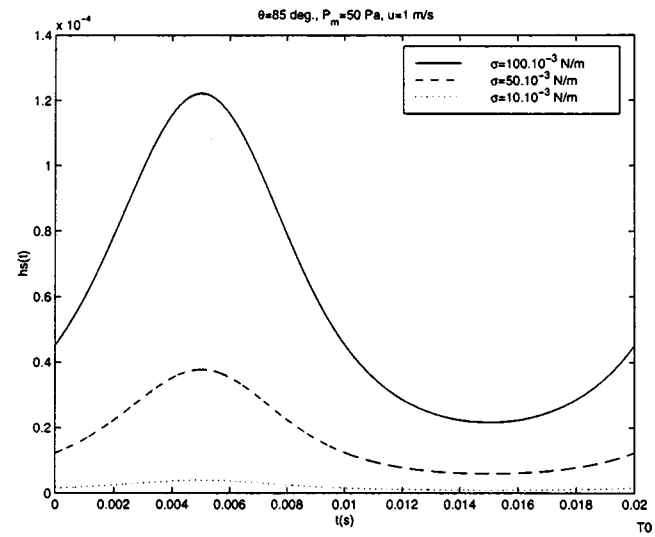



**Fig. 6 Influence of surface tension on vertical displacement**

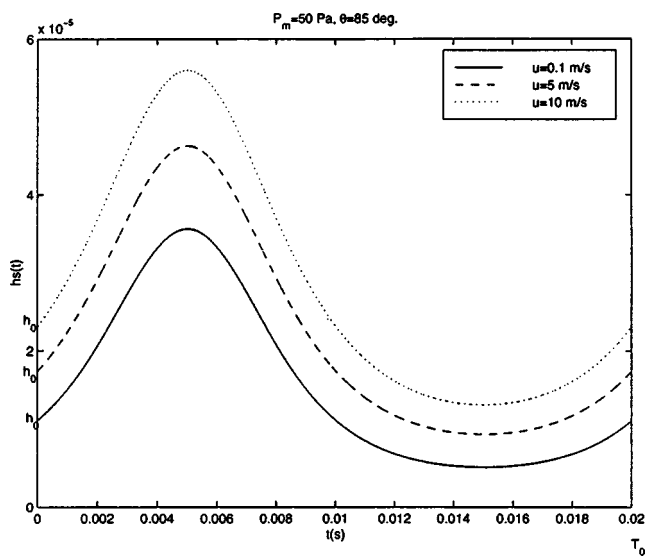



**Fig. 4 Influence of velocity on vertical displacement**

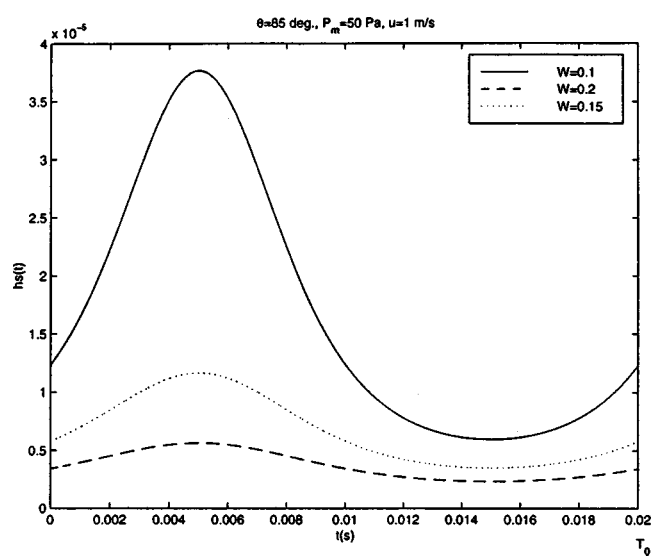



**Fig. 7 Influence of pinning force on vertical displacement**

tact with the solid substrate and to adhere to it either by modifying the fluid wetting properties (reducing the wetting angle and surface tension) or by using so-called pinning systems (for instance, by means of an electrostatic or a magnetic field). The influence of the wetting angle (Fig. 5) confirms the previous remarks: When the wetting angle is small, the liquid better sticks on the solid, which damps down the vertical fluctuations of the meniscus. In the same way, Fig. 6 illustrates the beneficial influence of low surface tension. Figure 7 confirms that increasing the pinning force is efficient indeed.

## 4 Second Step: Lateral Displacement of the Triple Line

As pointed out in Section 3, large pressure fluctuations occur in the vicinity of the triple line, which should lead to displacements of the three-phase junction when maintained at prescribed location. The variation in the position of the triple line (given by the unknown $x_s$) will now be addressed as a function of time $x_s(t)$.

**4.1 Additional Equation.** The boundary condition at the outlet $x_s$ remains given by the Laplace equation as in Eq. (4*b*):

$$P(x=x_s)=\frac{\sigma}{\rho_c}. \tag{14}$$

The three-phase junction is now free to move along the substrate. It means that its location $x_s(t)$ becomes the third unknown along with the vertical distance $h_s(t)$ and the pressure field $P(x,t)$. Abscissa $x_s$ is controlled via an additional condition at the inlet: The flow rate $Q_e$ at the entrance $(x_e)$ is prescribed; its value is evaluated by a study in the upstream zone in the spirit of Bourgin and Tahiri [19]—its order of magnitude being equivalent to a one micrometer height of air.

The initial conditions for $h_s(t)$ and $x_s(t)$ are given again by the result of the steady case:

$$h_s(t=0)=h_0 \tag{15}$$

$$x_s(t=0)=x_{s0}. \tag{16}$$

**4.2 Solution.** The problem defined by Eqs. (2)–(5) (Eq. (4*b*) being replaced by Eq. (14)) and the flow rate condition leads to a system of two coupled nonlinear differential equations of the first order (both similar to Eq. (8)) involving only $h_s(t)$ and $x_s(t)$. The same set of data as in Section 3 has been selected: $x_e=-1$ mm; $R=1$ mm; $\eta=18.10^{-6}$ Pa.s. In addition, only one value of the substrate velocity was chosen, as an illustrative example: $u=1$ m/s. The problem was solved by a Runge-Kutta method of the fourth order. As quoted before, the values of the inlet abscissa $(x_e)$ does not affect the results, provided it be sufficiently large far upstream of $x_s(t)$.

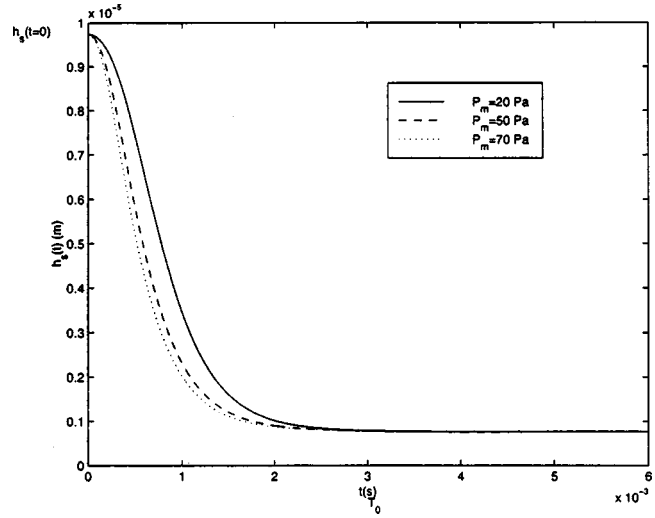



**Fig. 9 Influence of the amplitude of a small perturbation on vertical displacement**

### 4.3 Results

(i) For a given set of process parameters, the displacements of the triple line are very sensitive to any oscillations of the inlet pressure $P(x_e,t)$, as seen in Fig. 8. However, the response of $h_s(t)$ to inlet pressure oscillations are very different (Fig. 9): A transient regime takes place until the air gap distance reaches a constant value which is significantly lower than that in the steady case. There is no clear explanation why the oscillations are mainly transmitted to the triple line lateral displacements.

(ii) Keeping the same set of process parameters, the effects of the flow rate on $x_s(t)$ and $h_s(t)$ for prescribed pressure oscillations are respectively shown in Figs. 10 and 11. After the transient regime, which typically extends over one period ($\simeq 3.10^{-3}$ $s=T_0$) the inlet flow rate does not affect much the triple line

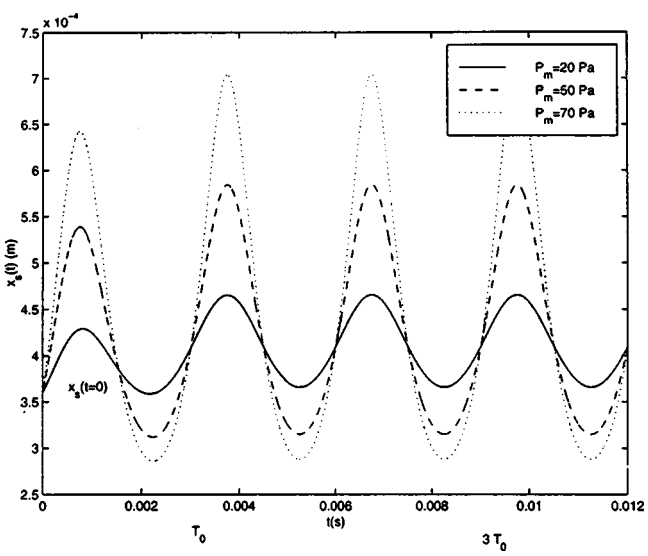



**Fig. 8 Influence of the amplitude of a small perturbation on lateral motion of triple line**

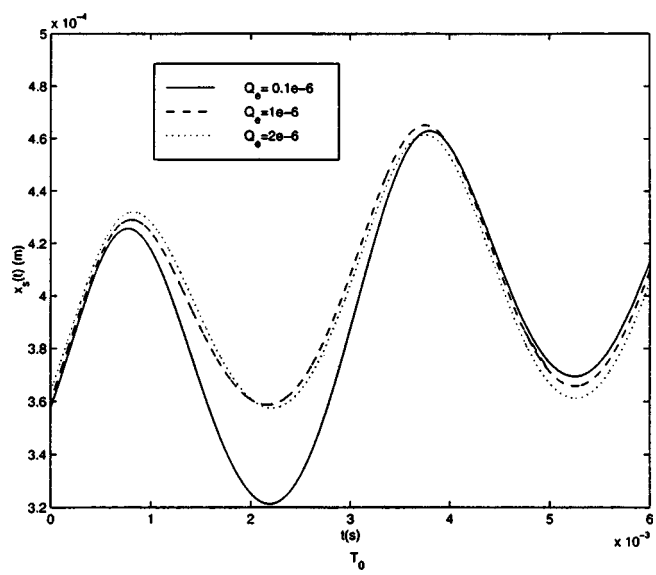



**Fig. 10 Influence of inlet flow rate on lateral motion of triple line**

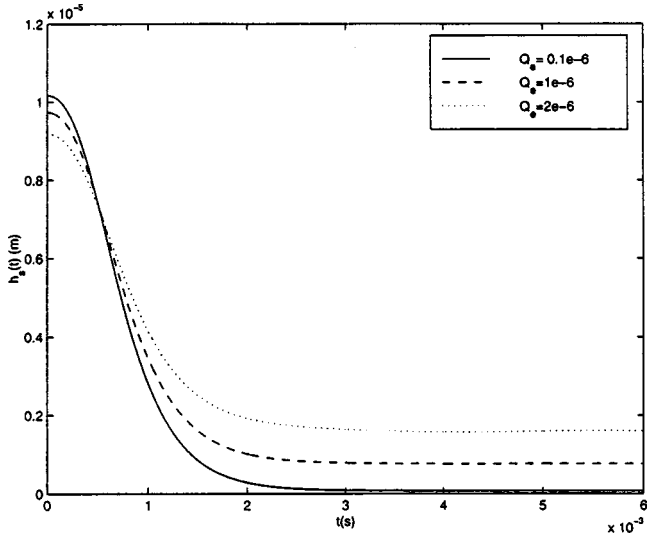



**Fig. 11 Influence of inlet flow rate on vertical displacement**

motion (Fig. 10). The asymptotic values of $h_s(t)$ reached after the transient regime ($\simeq 3.10^{-3} s = T_0$) are increasing functions of the inlet flow rate (Fig. 11).

(iii) The dependence of the lateral displacement upon the liquid wetting properties as well as the process parameters is the same as quoted in Section 3 for the vertical displacements: Improving the coating quality of the liquid reduces the effect of any fluctuation at the inlet.

## 5 Conclusion

Considering the configuration of high-velocity curtain coating flows, a simple model is proposed here to predict the influence of pressure fluctuations in the air on the motion of the three-phase junction in both vertical and lateral motion directions. The main result is that very slight pressure oscillations close to the triple line may induce large lateral displacements of the triple line together with a reduction of the air gap thickness compared to the static case. Therefore, the thickness regularity of the final coated product is not expected to suffer from the liquid sheet vertical shifts directly, but from the lateral displacements of the triple line. These phenomena can lead either to irregularities or eventually to a complete destruction of the junction with catastrophic air entrainment. The respective roles of the liquid wetting properties (wetting angle, surface tension) as well as the process parameters (translating velocity, pinning force) and parameters characteristic of the pressure oscillations (amplitude) are investigated: The theoretical predictions are in good qualitative agreement with well-known observations.

## Acknowledgments

The authors wish to acknowledge the French Office of Research and Education for financial funding (grant for P. Woehl).

## References

[1] Burley, R., and Jolly, R. P. S., 1984, ''Entrainment of Air Into Liquids by a High Speed Continuous Solid Surface,'' Chem. Eng. Sci., **39**, No. 89, pp. 1357–1372.
[2] Burley, R., and Kennedy, B. S., 1976, ''A Study of the Dynamic Wetting Behavior of Polyester Tapes,'' Br. Polym. J., **8**, pp. 140–143.
[3] Burley, R., 1992, ''Air Entrainment and the Limits of Coatability,'' JOCCA, **75**, No. 85, pp. 192–202.
[4] Gutoff, E. B., and Kendrick, C. E., 1982, ''Dynamic contact angles,'' AIChE J., **28**, pp. 459–466.
[5] Miyamoto, K., and Scriven, L. E., 1982, ''Breakdown of Air Film Entrained by Liquid Coated on a Web,'' AIChE Meeting, Los Angeles, CA, paper 101g.
[6] Mues, W., Hens, J., and Boiy, L., 1989, ''Observation of a Dynamic Wetting Process Using Laser-Doppler Velocimetry,'' AIChE J., **35**, No. 89, pp. 1521–1526.
[7] Veverka, P. J., and Aidun, C. K., 1991, ''Using High Speed Cinematography to Obtain Visual Evidence on Air Entrainment in Low-Speed Roll Coating,'' Tappi J., **75**, pp. 203–210.
[8] Veverka, P. J., and Aidun, C. K., 1996, ''An Investigation of Interfacial Instability During Catastrophic Air Entrainment,'' AIChE Spring National Meeting, New Orleans, LA, Paper 7d.
[9] Miyamoto, K., 1991, ''On the Mechanism of Air Entrainment,'' Ind. Coating Res., **1**, pp. 71–88.
[10] Huh, C., and Scriven, L. E., 1971, ''Hydrodynamic Model of Steady Movement of a Solid/Liquid/Fluid Contact Line,'' J. Colloid Interface Sci., **35**, No. 1, pp. 85–101.
[11] Emonot, J., 1994, ''Contribution à l'Étude du Contact Dynamique Entre une Nappe de Liquide et une Surface Solide,'' Ph.D. thesis, Université Claude Bernard de Lyon.
[12] Emonot, J., Bourgin, P., and Burley, R., 1995, ''Microfluid Mechanics of Air Entrainment in High Speed Coating Processes,'' *Int. Symp. on Two-Phase Flow Modelling and Experimentation*, Roma, Italy.
[13] Bourgin, P., and Tahiri, N., 1996, ''Triple Line : Myth or Reality?,'' AIChE Spring National Meeting, New Orleans, LA, Paper 7b.
[14] Tahiri, N., 1997, ''Etude de l'Écoulement d'un Fluide Très Visqueux sur un Substrat Solide se Déplaçant à Grande Vitesse : Entraînement d'Air,'' Ph.D. thesis, Université Louis Pasteur de Strasbourg.
[15] Tahiri, N., Bourgin, P., and Saintlos, S., 1997, ''High Velocity Dynamic Contact Between a Viscous Liquid and a Solid Substrate : Analytical and Numerical Study,'' *Proc. Euromech Symposium 367*, Strasbourg, France.
[16] Bourgin, P., and Saintlos, S., 1998, ''Dynamic Effects of Air Entrainment in High Velocity Coating Flows,'' *Rheology and Fluid Mechanics of Nonlinear Materials*, ASME FED-Vol. **246**, ASME, New York, pp. 95–100.
[17] Levarlet, C., and Finnicum, D. S., 1996, ''Unsteady State Bead Visualization,'' AIChE Meeting, New Orleans, LA.
[18] Woehl, P., 1999, ''Mouillage Dynamique à Grande Vitesse: Entraînement d'Air et Stabilité,'' Ph.D. thesis, Université Louis Pasteur de Strasbourg.
[19] Bourgin, P., and Tahiri, N., 1995, ''Generalized Jeffery-Hamel Flow : Application to High Velocity Coating,'' *Proc. First European Coating Symposium*, Leeds, UK.

# Stability Analysis of an Initially Inclined Ship in Following Sea

**S. K. Lee**
Department of Naval Architecture
and Ocean Engineering,
Pusan National University,
Pusan, Korea

*Generally speaking, the causes of ship capsizing in waves are due to the pure loss of stability, parametric oscillation, and broaching-to phenomenon. In this paper, an analytic expression for the GZ value of a ship, inclined initially by some reasons like damage in longitudinal waves, is derived. And, the roll motion is calculated through the computer simulation to give the necessary information for the safe regions of various parameters of a ship moving in a longitudinal wave.* [S0021-8936(00)00704-2]

## 1 Introduction

When a ship is moving in the following waves and if the encounter frequency becomes about one half of ship's natural roll frequency, low-cycle parametric oscillation occurs and sometimes develops into the capsizing of a ship.

Grim [1], Kerwin [2] Hamamoto [3], and Paulling [4] solved this problem by the ''Mathieu Type'' differential equation and by computer simulations.

In this paper, the authors developed the analytic expression for the *GZ* value for an initially inclined ship with an angle of $\phi_i$ in longitudinal waves. And various calculations are executed with the changes of *GM* and $T_e$, the encounter period of ship to waves. As a result, the stable and unstable region of these parameters are found.

## 2 Equation of Motion

The roll motion of a ship in a sinusoidal wave can be expressed by Eq. (1), using the coordinate system in Fig. 1.

$$(I_{xx}+J_{xx})\ddot{\phi}+N\dot{\phi}+W\cdot GZ(\xi_G,\phi)=0 \tag{1}$$

Here, $(I_{xx}+J_{xx})$ is the virtual moment of roll and $N$ is damping coefficient, and $\xi_G$ is the longitudinal position of C.G. of a ship relative to sinusoidal waves.

In Eq. (1), the wave exciting forces on the ship are excluded to simplify the problems. And the variation of *GZ* along the change of relative position $\xi_G$ to waves is assumed as a sinusoidal function of Eq. (2).

$$GZ=\frac{\phi_r}{\pi}GM\sin\left(\pi\frac{\phi}{\phi_r}\right) \tag{2}$$

Here, $\phi_r$ is the angle of vanishing stability. And *GZ* for the initially inclined ship with an angle of $\phi_i$ is also expressed as

$$GZ=\frac{\phi_r}{\pi}GM\sin\left(\pi\frac{\phi+\phi_t}{\phi_r}\right)-GZ_{\phi_i} \tag{3}$$

Here

$$GZ_{\phi_i}=\frac{\phi_r}{\pi}GM\sin\left(\pi\frac{\phi_i}{\phi_r}\right). \tag{4}$$

According to Kerwin [2], the variation of *GM* along the waves can be presented as follows:

$$GM_W=GM+\Delta GM\cos k(\xi_G-ct). \tag{5}$$

Here, $k$ is the wave number and $c$ is the phase velocity of the wave. $\phi_r$, the angle of vanishing stability, is assumed similarly.

$$\phi_r=\phi_R\left\{1+\frac{\Delta\phi}{\phi_R}\cos k(\xi_G-ct)\right\} \tag{6}$$

Here, $\phi_R$ is the angle of vanishing stability in still water, and $\Delta\phi$ is the small change from $\phi_R$. Meanwhile *GM* is related to *GZ* by the following equation:

$$GM=\left.\frac{dGZ}{d\phi}\right|_{\phi=0}. \tag{7}$$

Using the equations above, we can derive the following relation:

$$GZ=\frac{\phi_r}{\pi}GM\left\{1+\frac{\Delta GM}{GM}(\xi_G-ct)\right\}\left\{\sin\left(\pi\frac{\phi+\phi_i}{\phi_r}\right)-\sin\left(\pi\frac{\phi_t}{\phi_r}\right)\right\} \tag{8}$$

Substituting Eq. (8) to Eq. (1), we get

$$\ddot{\phi}+2\alpha_e\dot{\phi}+\frac{\phi_r}{\pi}\omega_\phi^2\left\{1+\frac{\Delta GM}{GM}\cos(k\xi_0-\omega_e t)\right\}\cdot\left\{\sin\left(\pi\frac{\phi+\phi_i}{\phi_r}\right)-\sin\left(\pi\frac{\phi_i}{\phi_r}\right)\right\}=0. \tag{9}$$

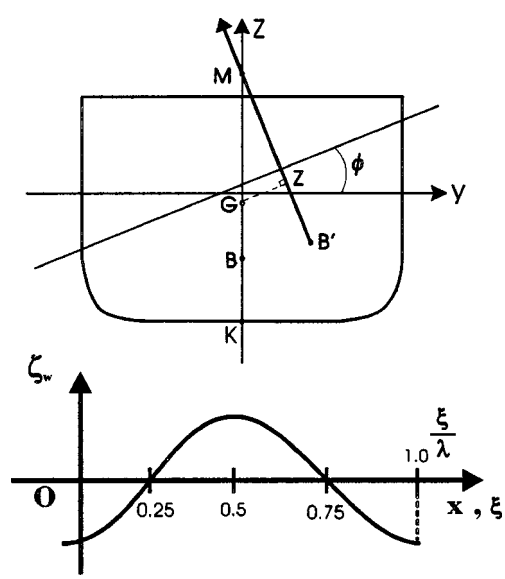



**Fig. 1 Coordinate system**

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, April 6, 1999; final revision, May 5, 2000. Associate Technical Editor: W. K. Liu. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS.

 **Copyright © 2000 by ASME** 

Here,

$$\phi_r = \phi_R \left\{ 1 + \frac{\Delta\phi}{\phi_R} \cos(k\xi_0 - \omega_e t) \right\} \tag{10}$$

$$\alpha_e = \frac{N}{2(I_{xx} + J_{xx})} \tag{11}$$

$$\omega_\phi^2 = \frac{W \cdot GM}{(I_{xx} + J_{xx})} \tag{12}$$

and

$$\xi_G = \xi_0 + Ut \tag{13}$$

$$\omega_e = k(c - U). \tag{14}$$

## 3 Numerical Calculation of Roll Motion

Nonlinear Eq. (9) can be linearized and converted to the Mathieu-type equation as follows:

$$\frac{d^2\Phi}{d\tau^2} + \left(\frac{T_e}{T}\right)^2 \left[ 1 - \left(\frac{a_e}{\pi}\right)^2 + \frac{\Delta GM}{GM} \cos\tau \right] \Phi = 0 \tag{15}$$

Here

$$\phi(t) = \Phi e^{-\alpha_e t}, \quad \tau = \omega_e t.$$

$a_e$ is the effective extinction coefficient,

$$a_e = \frac{T}{2} \cdot \alpha_e . \tag{16}$$

For a ship $(a_e/\pi)^2$ is a too small value and the outline of rolling motion will mainly be determined by two parameters $\Delta GM/GM$ and $T/T_e$.

Therefore, we can say the variation of $GM$ and the change of the encounter period $T_e$ are main parameters that affect the characteristics of roll motion in longitudinal waves.

We made two nondimensional parameters, $\Delta GM/GM$ and $T/T_e$ for convenience. Here $GM$ is the initial metacentric height, and $T$ is the natural period of a roll. We presumed $\phi_R = 40$ deg. $\Delta GM/GM = 0.5$. $\Delta\phi/\phi_R = 0.5$ and $\phi_i = 10$ deg and the calculation was done by the Runge-Kutta-Gill's method.

Figure 2 shows the variation of $GZ$ along the waves of an inclined ship, and Fig. 3 shows the stable, critical, and unstable region of $\Delta GM/GM$ and the $T/T_e$ parameters.

Extra calculation for $\Delta\phi/\phi_R = 0.3$ is carried out to see the effects of $\Delta\phi/\phi_R$ on rolling motion, and it was found that the stable

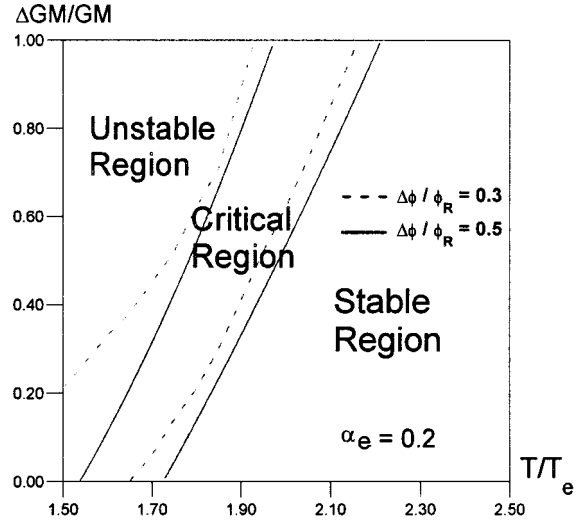



**Fig. 3 Stable and unstable regions for nonlinear equation of the ten-deg inclined ship**

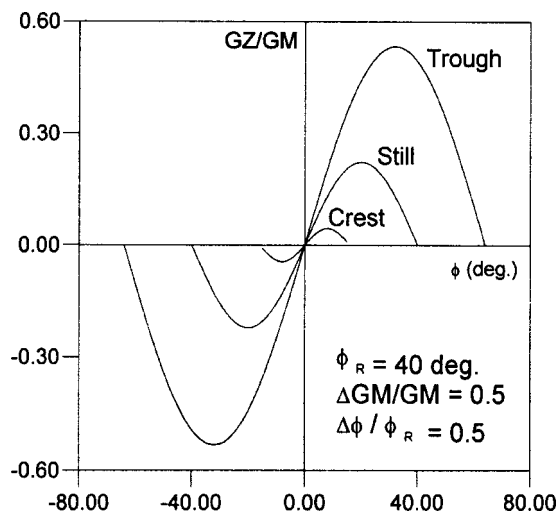



**Fig. 4 Righting arm for computation of roll motion of an upright ship**

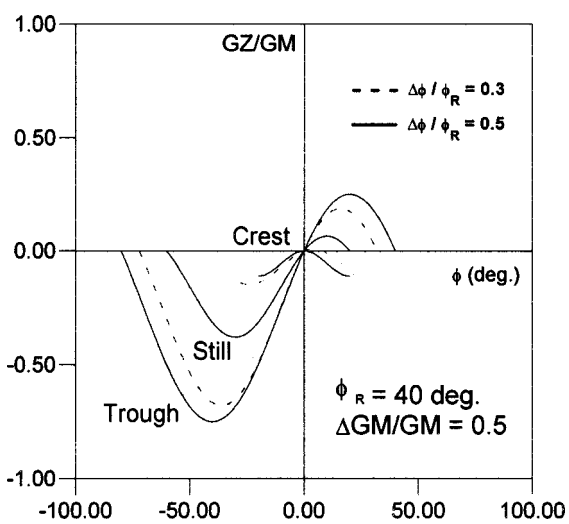



**Fig. 2 Righting arm for computation of roll motion of the ten-deg inclined ship**

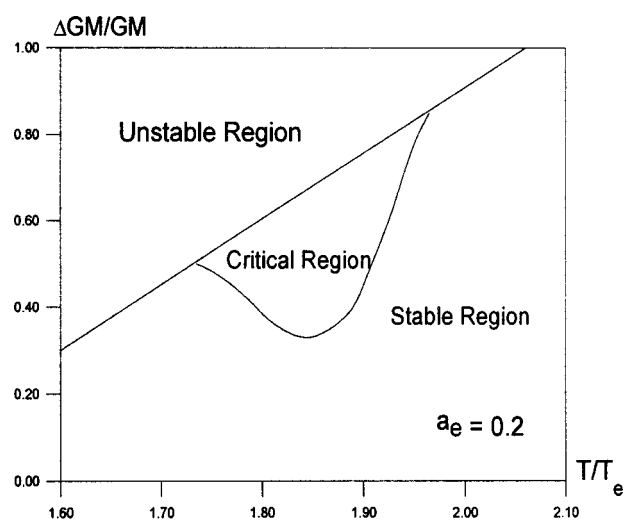



**Fig. 5 Stable and unstable regions for nonlinear equation of an upright ship**

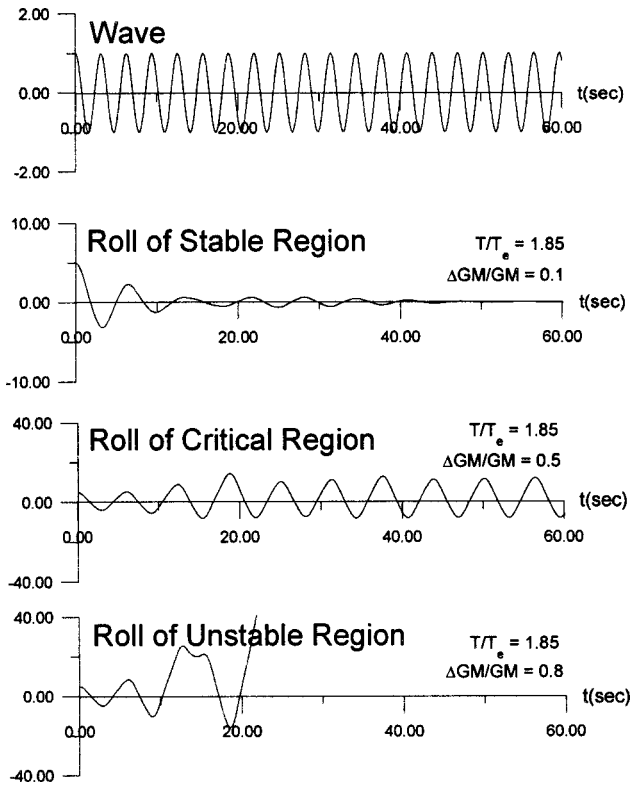



**Fig. 6 Time histories of stable and unstable regions for a nonlinear equation in the ten-deg inclined ship**

region became wider with a smaller $\Delta\phi/\phi_R$ value. Here, the critical zone is a region, where the roll motion becomes steady oscillatory.

And another calculation of the same condition, but initially in an upright case ($\phi_i = 0°$) is carried out, to examine the initial inclination effects. These results are indicated in Figs. 4 and 5.

From above results we can clearly see the difference of the safety region in an initially inclined and upright ship.

Finally, Fig. 6 shows the time histories of roll angles in a stable, critical, and unstable situation of an initially inclined ship.

From the figures above, we can see that the nonlinear behaviors of roll motion are successfully calculated.

## 4 Conclusion

The lateral stability of an initially inclined ship in longitudinal waves is studied, with use of the nonlinear roll equation. The following are the main results derived from this study.

1 Proposed nonlinear roll equation can simulate the roll motion of an initially inclined ship in longitudinal waves.
2 Using this method, the stable zone in longitudinal motion can be assessed.
3 This approach can be applied to calculate the safety margin of a ship in various conditions, including damage.

## Nomenclature

$a_e$ = effective extinction coefficient
$c$ = phase velocity of wave
$GM$ = metacentric height
$\Delta GM$ = small change from $GM$
$GZ$ = righting arm
$I_{xx}$ = mass moment of inertia of a ship
$J_{xx}$ = added mass moment of inertia of a ship
$k$ = wave number
$K_\phi$ = roll damping coefficient
$N$ = damping coefficient of a ship
$W$ = weight of a ship
$t$ = time
$T$ = natural period of a roll
$T_e$ = encounter period of wave
$\phi$ = roll angle
$\phi_R$ = angle of vanishing stability
$\Delta\phi$ = small change from $\phi_R$
$\phi_i$ = initial inclination angle
$\omega_\phi$ = natural circular frequency
$\omega_e$ = encounter wave frequency
$\xi_G$ = relative position of a ship to wave

## References

[1] Grim, O., 1952, ''Rollschwingungen, Stabilitat und Sicherheit in Seegang,'' *Schiffstechnik*, pp. 20–35.
[2] Kerwin, J. E., 1955, ''Notes on Roll in Longitudinal Waves,'' Int. Ship Prog., **16**, No. 2, pp. 597–614.
[3] Hamamoto M., and Fujino M., 1986, ''Capsizing of Ships in Following Seas,'' *Safety and stability of ships and offshore structures*, 3rd Marine Dynamic Symposium, pp. 125–158.
[4] Pauling, J. R., 1961, ''The Transverse Stability of a Ship in Longitudinal Seaway,'' J. Ship Res., **4**, No. 4, pp. 37–49.

# Stick-Slip, Imposition-Removal of Constraints and the Spinning Ball Problem

**S. Djerassi**
RAFAEL,
P.O. Box 2250,
Haifa, Israel

*It has been observed that balls pressed between elastic bodies spin when subjected to linear, cyclic motion. This paper proposes an explanation to this phenomenon, based upon the stick-slip theory. To this end, a modified, vectorial formulation of the stick-slip theory is presented. The formulation is applied to a model comprising a ball pressed between pairs of springs and dampers. A computer program based on this formulation is used to predict the resulting motion. Predictions are shown to agree with experimental results.* [S0021-8936(01)00701-2]

## 1 Introduction

Figure 1(*a*) shows a uniform steel ball $B$ of mass $m$ and radius $R$ pressed between two elastic plates $P_B$ and $P_H$, connected to a rigid housing $B'$. A simplified model of the system in Fig. 1(*a*) is shown in Fig. 1(*b*). There, $P_B$ and $P_H$ are replaced with two rigid, massless plates, two massless springs $S_B$ and $S_H$, each of rate $k$, and two massless dampers $D_B$ and $D_H$, each of constant $b$ which is significantly smaller than the critical damping. Suppose that $B'$ is uniformly displaced by an amount $y(\cos\alpha\mathbf{n}_1+\sin\alpha\mathbf{n}_2)$, where $y=a\sin 2\pi ft$. Furthermore, suppose that the static deformation of the springs is $x_0$. Then, for certain values of $\alpha$ and $a$, the ball spins essentially with a constant angular speed $\Omega$ either clockwise or counterclockwise, depending on whether $f<1/2\pi\sqrt{2k/m}$ or $f>1/2\pi\sqrt{2k/m}$. This behavior was the subject of an extensive research by Edelstein [1]. He subjected a system similar to that in Fig. 1(*a*) to 150 m/sec$^2$ $(=a(2\pi f)^2)$ uniform sine-sweep vibrations in the 500–1100 Hz frequency range (then $a$ varies in the range 15.2–3.14 micrometer), changing $f$ at a rate of 10/3 Hz/sec; and measured $\Omega$ as a function of $f$. For a system characterized by $f_n\triangleq 1/2\pi\sqrt{2k/m}=720$ Hz, $F\triangleq kx_0=336$ N (a preload 150 larger than the weight of the ball), $2R=0.0381$ m and $\alpha=37$ deg he obtained Fig. 2, showing $\Omega(f)$.

It is the purpose of the present work to show that this kind of behavior can be predicted with reasonable accuracy, as shown in Fig. 3(*a*), if use is made of a modified formulation of the stick-slip theory. The modified formulation will next be presented and discussed in light of existing formulations; and then applied to the problem of a particle moving on a vibrating slider, and to the problem of the spinning ball.

## 2 Stick-Slip Theory

**Equations of Motion.** A generic stick-slip theory can be described with the aid of $B'$ and $P$, a rigid body and a particle, respectively, that remain in contact with one another during motions of $S$, a system including $B'$ and $P$. According to this theory, two phases of motion may occur: slipping phase, when $P$ moves relative to $B'$; and sticking phase, when $P$ remains fixed in $B'$. Then stick-slip is said to take place at $P$.

This state of affairs can be dealt with analytically if $S$ is assumed to be a simple, nonholonomic system of $\nu$ particles $P_i(i=1,\ldots,\nu)$, possessing $\bar{n}$ generalized coordinates $q_1,\ldots,q_{\bar{n}}$ and $n$ generalized speeds $u_1,\ldots,u_n$ in $N$, a Newtonian reference frame, where $\bar{n}\geqslant n$. Suppose that the motion of $S$ in $N$ is *defined* as unconstrained in the slipping phase, and governed by $n$ dynamical equations, namely,

$$F_r+F_r^*=0 \quad (r=1,\ldots,n) \tag{1}$$

where $F_r$ and $F_r^*$ are the $r$th generalized active force and the $r$th generalized inertia force for $S$, respectively. Moreover, suppose that the sticking phase is associated with constraints imposed on the motion of $S$. These constraints can be represented by $m$ linear relationships between the generalized speeds, and formulated as follows:

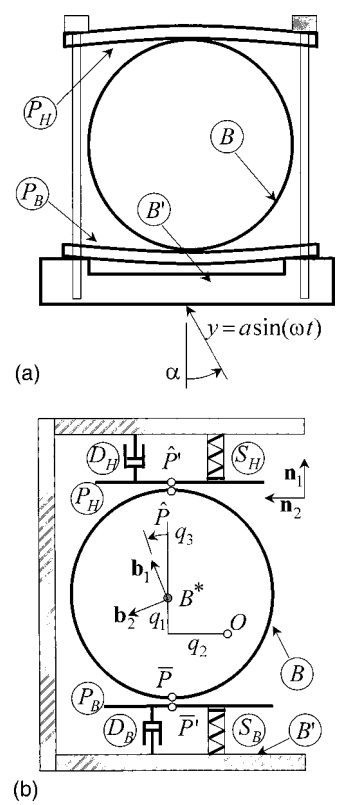



**Fig. 1 A ball pressed between two elastic plates**

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Feb. 10, 1999; final revision, July 18, 2000. Associate Technical Editor: R. C. Benson. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

 **Copyright © 2000 by ASME** 

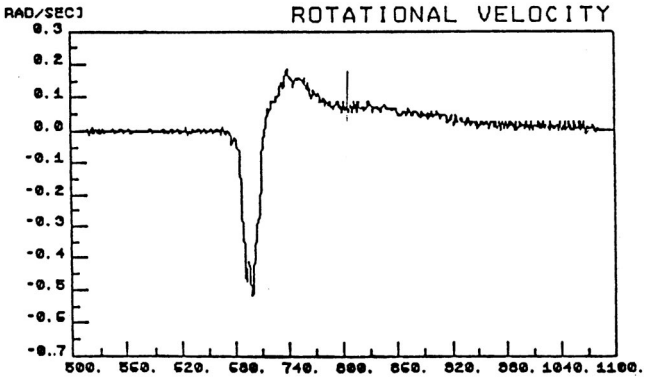



**Fig. 2 Ω as a function *f* for *α*=37 deg: experimental results**

$$u_k=\sum_{r=1}^{p} C_{kr}u_r+D_k \quad (k=p+1,\ldots,n, p=n-m), \tag{2}$$

where $C_{kr}$ and $D_k$ $(k=p+1,\ldots,n;\ r=1,\ldots,p)$ are functions of $q_1,\ldots,q_{\bar{n}}$ and time $t$. Then the motion of $S$ in the sticking phase is *defined* as constrained and governed by the following equations ([2]):

$$F_r+F_r^*+\sum_{k=p+1}^{u} C_{kr}(F_k+F_k^*)=0 \quad (r=1,\ldots,p). \tag{3}$$

Next, suppose that a transition from the slipping phase to the sticking phase occurs between $t_1$, the last instant of the slipping phase, and $t_2$, the first instant of the sticking phase; and that $t_1$ and $t_2$ are infinitely close to one another. Then

$$q_r(t_2)=q_r(t_1) \quad (r=1,\ldots,\bar{n}); \tag{4}$$

i.e., the configuration of $S$ in $N$ remains unaltered between $t_1$ and $t_2$. Moreover, no instantaneous change in the velocities of points accompanies the transition, an observation based on physical grounds. Synnestvedt [3] reports that, in accordance with experiments conducted at NASA Ames, this is not strictly true. Certain relative velocities become zero abruptly after acquiring a sufficiently small value. It is generally assumed, however, that the associated instantaneous changes in the indicated velocities can be disregarded. This means that precisely $n$ independent relationships of the form $\mathbf{v}^{P_i}\cdot\mathbf{c}(t_2)=\mathbf{v}^{P_i}\cdot\mathbf{c}(t_1)$ can be identified, where $\mathbf{v}^{P_i}$ is the velocity of $P_i$ in $N$, and $\mathbf{c}$ is a unit vector; and these relationships reduce to

$$u_r(t_2)=u_r(t_1) \quad (r=1,\ldots,n) \tag{5}$$

when written explicitly. Consequently, Eqs. (2) are satisfied both at $t_1$ and at $t_2$.

Equations (4) and (5) are also valid in connection with transitions from the sticking phase to the slipping phase. Then $t_1$ and $t_2$ are the last instant of the sticking phase and first instant of the slipping phase, respectively. These roles are played by $t_1$ and $t_2$ throughout the remainder of this work. In summary, Eqs. (1) and (3) govern stick-slip motions, the transitions occurring *without changes in the integration variables*, as indicated by Eqs. (4) and (5).

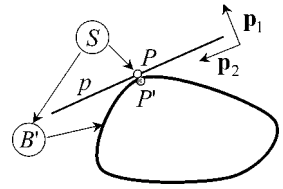



**Fig. 4 A particle in contact with the surface of a rigid body**

**Transition Conditions.** A key point in the simulations of motions involving stick-slip is the formulation of transition conditions, namely, conditions the satisfaction of which initiate a transition from one phase to another. To formulate these conditions, consider $B'$ and $P$ again, and let $P'$ be a point of $B'$ instantaneously in contact with $P$ in the slipping phase and constantly in contact with $P$ in the sticking phase, as shown in Fig. 4. Let $\mathbf{v}^P$ and $\mathbf{v}^{P'}$ be the velocities of $P$ and $P'$ in $N$. Then $\mathbf{v}_{\rm rel}$, defined by

$$\mathbf{v}_{\rm rel}=\mathbf{v}^P-\mathbf{v}^{P'} \tag{6}$$

must lie on $p$, the plane tangent to $B'$ at $P'$. If $\mathbf{v}_{\rm rel}$ does not lie on $p$, then either $P$ penetrates $B'$ or $P$ loses contact with $B'$. Finally, let $\mathbf{F}^{P'/P}$ be the force exerted by $P'$ on $P$, and note that $\mathbf{F}^{P'/P}\cdot\mathbf{p}_1>0$ at all times if $\mathbf{p}_1$ is a unit vector perpendicular to $p$ and directed as shown in Fig. 4. The two phases and the associated transition conditions can be described as follows.

Sticking phase (Eqs. (3)): $P$ remains in contact with $P'$. Then

$$\mathbf{v}_{\rm rel}=0, \tag{7}$$

a vector equation with scalar counterparts that can be cast in the form of Eqs. (2). Let $\mathbf{p}_2'$ be a unit vector aligned with the projection of $\mathbf{F}^{P'/P}$ on $p$, i.e.,

$$\mathbf{p}_2'\triangleq(\mathbf{F}^{P'/P}-\mathbf{F}^{P'/P}\cdot\mathbf{p}_1\mathbf{p}_1)/|\mathbf{F}^{P'/P}-\mathbf{F}^{P'/P}\cdot\mathbf{p}_1\mathbf{p}_1|. \tag{8}$$

Then $\mathbf{F}^{P'/P}\cdot\mathbf{p}_2'\mathbf{p}_2'$ is the friction force exerted on $P$ by $P'$. In connection with numerical analysis, the sticking phase persists unless

$$\mu_s\mathbf{F}^{P'/P}\cdot\mathbf{p}_1<|\mathbf{F}^{P'/P}\cdot\mathbf{p}_2'| \tag{9}$$

where $\mu_s$ is the static coefficient of friction.

Slipping phase (Eqs. (1)): $P$ moves relative to $P'$, i.e., $\mathbf{v}_{\rm rel}\neq 0$, except at $t=t_2$ (the first instant of the slipping phase), when $\mathbf{v}_{\rm rel}(t_2)=0$. Let $\mathbf{p}_2$ be a unit vector aligned with $\mathbf{v}_{\rm rel}$, namely,

$$\mathbf{p}_2\triangleq\mathbf{v}_{\rm rel}/(|\mathbf{v}_{\rm rel}|+d), \quad d|_{t=t_2}=1; d|_{t\neq t_2}=0, \tag{10}$$

where $d$ is defined so that $\mathbf{p}_2(t_2)=0$. Then, in accordance with Coulomb's friction law

$$\mathbf{F}^{P'/P}=\mathbf{F}^{P'/P}\cdot\mathbf{p}_1(\mathbf{p}_1-\mu_d\mathbf{p}_2^*), \quad \mathbf{p}_2^*\triangleq\mathbf{p}_2-d\mathbf{p}_2'(t_1), \tag{11}$$

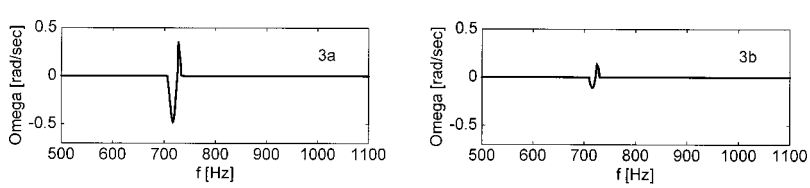



**Fig. 3 Ω as a function *f* for *α*=37 deg: simulation results for (*a*) *b*=14, (*b*) *b*=18**

where $\mu_d$ is the dynamical coefficient of friction between $P$ and $B'$. The definition of $\mathbf{p}_2^*$ indicates, in view of Eqs. (10), that $\mathbf{p}_2^* = \mathbf{v}_{\text{rel}}/|\mathbf{v}_{\text{rel}}|$ whenever $t \neq t_2$, and $\mathbf{p}_2^* = -\mathbf{p}_2'(t_1)$ when $t = t_2$. The rationale underlying the introduction of $\mathbf{p}_2^*$ is the following. By definition $\mathbf{v}_{\text{rel}}(t_2) = 0$, hence $\mathbf{v}_{\text{rel}}$ cannot be used to determine the direction of the friction force at $t_2$. However, at $t = t_1$ (the last instant of the sticking phase) the direction of the friction force is given by $\mathbf{p}_2'(t_1)$ (see Eq. (8)). Moreover, $t_2$ is infinitely close to $t_1$, hence the following assumption is made, namely, that $\mathbf{p}_2'(t_1)$ is also the direction of the friction force at $t = t_2$. Note that this assumption holds only for the first integration step following a transition from the sticking phase to the slipping phase. Then the direction of the friction force change instantaneously.

In connection with numerical analysis, the slipping phase persists unless

$$|\mathbf{v}_{\text{rel}}| < \delta, \quad \mathbf{a}_{\text{rel}} \cdot \mathbf{p}_2 < 0. \tag{12}$$

Here $\mathbf{a}_{\text{rel}} = d/dt(\mathbf{v}_{\text{rel}})$, where differentiation in $N$ is implied; and $\delta$ is a positive constant chosen to be less than the maximum expected value of $|\mathbf{a}_{\text{rel}} \cdot \mathbf{p}_2| \Delta t$ in the region $|\mathbf{v}_{\text{rel}}| < \varepsilon$, where $\varepsilon$ is a predetermined, problem-dependent constant, and $\Delta t$ is the integration time-step. The exact value assigned to $\delta$ is the outcome of a compromise between two conflicting requirements. These are (a) passages through $\mathbf{v}_{\text{rel}} = 0$ always lead to transitions to the sticking phase (hence $\delta$ should be large); and (b) transitions to the sticking phase occur when $\mathbf{v}_{\text{rel}} = 0$ (hence $\delta$ should be small). Now, if $\delta$ is too small or too large compared to $|\mathbf{a}_{\text{rel}} \cdot \mathbf{p}_2| \Delta t$, then chatters may arise. In the present context, chatters comprise series of nonphysical switches from one phase to another at consecutive time steps. These chatters lead, at best, to inefficient simulations, and frequently to inaccurate results. Chatters arise if $\delta$ is ''too small'': then passages of $\mathbf{v}_{\text{rel}}$ through zero may not be detected; or if $\delta$ is ''too large'': then transitions to the sticking phase may occur when $\mathbf{v}_{\text{rel}}$ is much larger than zero. Note that, in the absence of $\mathbf{a}_{\text{rel}} \cdot \mathbf{p}_2 < 0$ from condition (12) an additional cause for chatters arises if $\delta$ is ''too large.'' Then transitions to the slipping phase may lead, after one integration step, to $\mathbf{v}_{\text{rel}}$ which is smaller than $\delta$. Transition back to the sticking phase will follow, although the slipping phase should have persisted. It may be concluded that, to date, no closed-form expression for the determination of $\delta$ was found that guarantees the elimination of chatters. It will be shown, however, that a few numerical experiments lead to a choice of $\delta$ by which chatters can be avoided.

The following comments are in order:

(a) Satisfaction of conditions (12) indicates that $|\mathbf{v}_{\text{rel}}|$ is smaller than $\delta$ and becomes yet smaller (it may occur that $|\mathbf{v}_{\text{rel}}|$ is smaller than $\delta$, but becomes larger, i.e., following a transition from the sticking phase to the slipping phase).
(b) The condition $\mathbf{a}_{\text{rel}} \cdot \mathbf{p}_2 < 0$ can be replaced with $\mathbf{a}_{\text{rel}} \cdot \mathbf{v}_{\text{rel}} < 0$ since $|\mathbf{v}_{\text{rel}}| > 0$ (see Eqs. (10)).
(c) $\mathbf{F}^{P'/P} \cdot \mathbf{p}_1$ in Eqs. (11) and (9) and $\mathbf{F}^{P'/P} \cdot \mathbf{p}_2'$ in Eq. (9) can be found, e.g., with the aid of a procedure given by Kane and Levinson ([2], Sec. 4.9) for the determination of ''noncontributing'' forces.
(d) $\mu_d$ can be defined as a function of $|\mathbf{v}_{\text{rel}}|$, accommodating different friction models, such as Coulomb's model, Stribeck's model, the linear model, and the multivalue model. These models are defined, e.g., by Hsien-I and Jeng-Hong [4].
(e) The satisfaction of conditions (9) and (12) is examined ''off-line'' after each integration step. The outcome of the examinations is the identification of either Eqs. (1) or Eqs. (3) as the equations which should be integrated in the next integration step.

Inequalities (9) and (12) are the requisite transition conditions. Their use is illustrated in the following sections.

**Past Efforts.** Mechanical systems involving stick-slip motions are the subject of numerous investigations. A number of authors seeking numerical simulations of motions of such systems take the ''phase approach,'' assuming that the motion consists of a number of phases governed by Eqs. (1)–(3). Transition conditions frequently encountered comprise simplified versions of conditions (9) and (12). A widely used version of these conditions reads

$$TC_{\text{slip-stick}} \quad |\mathbf{v}| < \delta; \quad TC_{\text{stick-slip}} \quad \mu_s N < F, \tag{13}$$

where the self-explanatory notation stands for quantities which are carefully defined in connection with conditions (9) and (12). At times, the use of this version leads to chatters. Nevertheless it is used, e.g., by Hsien-I and Jeng-Hong [4] investigating a number of friction models, and by Huston et al. [5], studying the dynamics of a bowling ball. Kane and Levinson [6] deal with the motion of a symmetric top with a spherical peg in contact with a horizontal plane. Their transition conditions do not include the second of conditions (12); however, they prolong the use of $\mathbf{p}_2'$ in equations similar to Eqs. (11) until $|\mathbf{v}_{\text{rel}}|$ becomes greater than $\delta$. Thus, chatters are avoided at the cost of slight inaccuracies. Synnesvedt [3] suggests a dynamic partitioning method to generate equations governing phases of motion involving stick, which he applies to a number of examples. In addition, Synnesvedt points out the role of acceleration/deceleration in the determination of transition conditions. However, he does not cast his observation into a workable formulation as in conditions (12).

Karnopp [7] offers a different approach to the stick-slip problem. He formulates only one set of motion equations, namely, the set governing the motion in the slipping phase. However, the friction forces exerted on points where stick-slip occurs have dual expressions. One expression is of the ''$\mu_d N$'' type, valid when slip persists. The other is defined so as to cause $|\mathbf{v}_{\text{rel}}|$ to vanish, hence is valid when stick persists. Finally, Karnopp uses conditions (13) to determine transitions, calling $\mu_s N$ ''breakaway force.'' It is not clear from Karnopp's paper how chatters are eliminated. Nevertheless, his approach is extensively used, mainly by authors synthesizing control schemes to overcome stick-slip. For instance, Southward et al. [8] discuss a force compensation scheme, supplementing PD control applied to a rotary system. Jia-Yush et al. [9] prove stability of a pseudo-derivative feedback (PDF) controller applied to a system undergoing stick-slip motion, and synthesize an accurate PDF rotary controller. Lastly, Tan and Rogers [10] apply Karnopp's approach to a system comprising a ball moving on a horizontal plane and acted upon by springs and dampers. It seems that the authors using Karnopp's approach achieve ''numerical stability,'' i.e., chatter-free simulations, with a proper, experimental choice of $\delta$.

When a complex stick-slip problem is considered, it is difficult to determine in advance whether Karnopp's approach is advantageous as compared with the ''phase approach''; for, a number of steps are common to the two approaches, e.g., the generation of $\mathbf{F}^{P'/P} \cdot \mathbf{p}_1$ and $\mathbf{F}^{P'/P} \cdot \mathbf{p}_2'$. It may be concluded, however, that conditions (9) and (12) are equally applicable to both these approaches. In fact, these conditions are applicable to any problem involving stick-slip motions. With a proper choice of $\delta$, no chatters arise, and the accuracy of the formulation is extended to its limits.

## 3 A Particle in Contact With a Vibrating Slider

Let $P$ be a particle of mass $m$, sliding on a planar surface of body $B'$, which undergoes a predetermined motion in $N$, as shown in Fig. 5. Let $q_1$ and $x(t)$ be defined as in the figure, and let $u_1 \triangleq \dot{q}_1$. Then $\mathbf{v}_{\text{rel}} = (u_1 - \dot{x})\mathbf{n}_2$, and the equation governing the motion of $P$ in $N$ in the slipping phase is, in view of Eq. (11),

$$-m\dot{u}_1 + \mathbf{F}^{P'/P} \cdot \mathbf{n}_1 (\mathbf{n}_1 - \mu_d \mathbf{p}_2^*) \cdot \mathbf{n}_2 = 0, \tag{14}$$

where $\mathbf{n}_1$ replaces $\mathbf{p}_1$ since $p$ is fixed in $N$. Now, $\mathbf{F}^{P'/P} \cdot \mathbf{n}_1 = mg$, where $g$ is the gravitational acceleration. Also, $\mathbf{p}_2 = (u_1 - \dot{x})\mathbf{n}_2/(|u_1 - \dot{x}| + d)$ and $\mathbf{p}_2'(t_1) = \ddot{x}(t_1)\mathbf{n}_2/|\ddot{x}(t_1)|$; hence Eq. (14) becomes

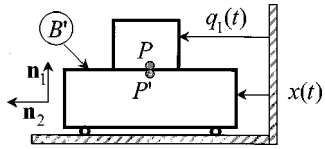



**Fig. 5 A particle in contact with a reciprocating base**

$$-m\dot{u}_1-\mu_d mg[(u_1-\dot{x})/(|(u_1-\dot{x})|+d)-d\ddot{x}(t_1)/|\ddot{x}(t_1)|]=0, \tag{15}$$

an equation playing the role of Eqs. (1) for $n=1$. Moreover, if, in accordance with conditions (12),

$$|u_1-\dot{x}|<\delta, \quad (\dot{u}_1-\ddot{x})(u_1-\dot{x})<0, \tag{16}$$

then transition to the sticking phase occurs, and the motion proceeds so that

$$u_1=\dot{x}(t). \tag{17}$$

Finally, $|\mathbf{F}^{P'/P}\cdot\mathbf{p}_2'|=m|\ddot{x}(t)|$ in the sticking phase, hence transition to the slipping phase occurs when

$$\mu_s mg<m|\ddot{x}(t)|, \tag{18}$$

a relation obtained by reference to condition (9). Note that in this case no equation plays the role of Eqs. (3) $(n=m=1)$, and Eq. (17) is a kinematical equation.

Suppose that $x=x_0\cos\omega t$. Then Figs. 6(*a*), 6(*b*), and 6(*c*) show the time histories of $q_1$, of $u_1$ and $\dot{x}$ (heavier line) and of the friction force (the term multiplied by $\mu_d$ in Eq. (15)), respectively, for $m=1$ kg, $\omega=3$ rad/sec, $x_0=0.35$ m, $\mu_d=0.15$, $\mu_s=0.3$, $\Delta t=0.0001$ sec and $\delta=0.0003$ m/sec $(|\mathbf{a}_{\rm rel}\cdot\mathbf{p}_2|_{\max}\Delta t=\ddot{x}_{\max}\Delta t=x_0\omega^2\Delta t=0.0003$ m/sec). Note that since $B'$ moves ''symmetrically'' back and forth, $P$ does not advance. Suppose now that the motion of $B'$ is induced by a crank-and-slider mechanism with a crank of radius $x_0$ and a connecting rod of length $kx_0$. Then $x=x_0\cos\omega t+kx_0\cos\{[\sin^{-1}[(\sin\omega t)/k]\}$; and, for $k=3$ and $x_0$ and $\omega$ as before, the time histories of $q_1$, of $u_1$ and $\dot{x}$ (heavier line) and of the friction force are as in Figs. 6(*d*), 6(*e*) and 6(*f*), respectively. These results were obtained with a numerical integrator based on Kutta-Merson, variable step-size algorithm with $\Delta t$ as the maximum time-step.

At times, stick-slip may occur at two points, say $\hat{P}$ and $\bar{P}$. Then two classes of problems can be identified. The first class involves cases where stick/slip at $\hat{P}$ implies stick/slip at $\bar{P}$, except, possibly, at singular configurations, and vice versa, as when $\hat{P}$ and $\bar{P}$ belong to the same rigid body. As before, this class of problems involves two phases of motion, namely, a sticking phase and a slipping phase; and was investigated, e.g., by Han et al. [11] and by Pfeiffer [12]. A second class of problems involves two points where stick-slip occurs independently, as when $\hat{P}$ and $\bar{P}$ belong to two distinct objects. This state of affairs calls for new definitions of phases of motion ([7]), because phases of motion are possible in which sticking occurs at one point and slipping occurs at the other. One example of such a system is that in Fig. 1, representing the problem of the spinning ball.

## 4 The Problem of the Spinning Ball

Here it will be shown that the spinning of the ball in Fig. 1 can be simulated if the stick-slip theory is applied to $\hat{P}$ and $\bar{P}$, points of $B$ instantaneously in contact with $\hat{P}'$ and $\bar{P}'$, points of $P_H$ and $P_B$, respectively. Stick-slip can occur at $\hat{P}$ and $\bar{P}$, independently; hence, four phases of motion are defined as follows. Phase 1: slip at both $\hat{P}$ and $\bar{P}$; Phase 2: stick at $\hat{P}$ and slip at $\bar{P}$; Phase 3: slip at $\hat{P}$ and stick at $\bar{P}$; and Phase 4: stick at both $\hat{P}$ and $\bar{P}$.

To obtain equations governing the motion of $B$ in $N$ one may start by writing motion equations of $B$ in Phase 1, when $B$ is in its ''minimally constrained state'' (the minimally constrained state of a system is defined as a state in which the motion is subject to constraints which are *not* removed throughout the motion ([13])). This can be done with the aid of $\mathbf{n}_i$ and $\mathbf{b}_i$ $(i=1,2,3)$, two sets of three dextral, mutually perpendicular unit vectors fixed in $N$ and $B$, respectively; and $u_r$ $(r=1,2,3)$, defined such that

$$\mathbf{v}^{B^*}=u_1\mathbf{n}_1+u_2\mathbf{n}_2, \quad \omega^B=u_3\mathbf{n}_3 \tag{19}$$

where $B^*$ is the mass center of $B$. Suppose that initially $\mathbf{b}_i$ are aligned with $\mathbf{n}_i$ $(i=1,2,3)$, that $\mathbf{v}^{B^*}=\omega^B=0$, that $B^*$ coincides with $O$, a point fixed in $N$; and define $q_r$ $(r=1,2,3)$ such that

$$\mathbf{p}^{O/B^*}=q_1\mathbf{n}_1+q_2\mathbf{n}_2, \quad \mathbf{n}_1\cdot\mathbf{b}_1=\cos q_3 \tag{20}$$

where $\mathbf{p}^{O/B^*}$ is the position vector from $O$ to $B^*$. Then $\dot{q}_r=u_r$ $(r=1,2,3)$.

Now, the only contact forces acting on $B$ are $\mathbf{F}^{\hat{P}'/\hat{P}}$ and $\mathbf{F}^{\bar{P}'/\bar{P}}$, exerted by $\hat{P}'$ on $\hat{P}$ and by $\bar{P}'$ on $\bar{P}$, respectively. These forces can be expressed, in accordance with Eq. (11), as

$$\mathbf{F}^{\hat{P}'/\hat{P}}=\hat{F}(-\mathbf{n}_1-\mu_d\hat{\mathbf{p}}_2^*), \qquad \hat{F}=F-k(y\cos\alpha-q_1)-b(\dot{y}\cos\alpha-u_1); \tag{21}$$

$$\mathbf{F}^{\bar{P}'/\bar{P}}=\bar{F}(\mathbf{n}_1-\mu_d\bar{\mathbf{p}}_2^*), \qquad \bar{F}=F+k(y\cos\alpha-q_1)+b(\dot{y}\cos\alpha-u_1). \tag{22}$$

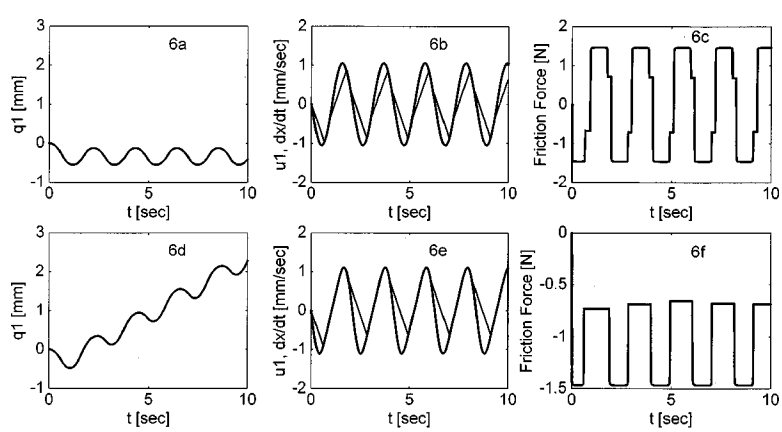



**Fig. 6 $q_1$, $u_1$ and friction force temporal behavior for symmetrical and nonsymmetrical reciprocating motions**

Here $\mathbf{p}_1=-\mathbf{n}_1$ and $\mathbf{p}_1=\mathbf{n}_1$ in connection with $\mathbf{F}^{\hat{P}'/\hat{P}}$ and $\mathbf{F}^{\overline{P}'/\overline{P}}$, respectively; $F$ is the preload of the springs (i.e., the magnitude of the force exerted by each of the springs on the ball when $q_1=y\cos\alpha$ and $u_1=\dot{y}\cos\alpha$); and $\hat{\mathbf{p}}_2^*$ and $\overline{\mathbf{p}}_2^*$ are unit vectors playing, in connection with $\hat{P}$ and $\overline{P}$, the role which $\mathbf{p}_2^*$ plays in connection with $P$ in Eqs. (11). Moreover, if $\hat{d}$ and $\overline{d}$ are defined similarly to $d$ in Eqs. (10), then $\hat{\mathbf{p}}_2^*$ and $\overline{\mathbf{p}}_2^*$ can be expressed as

$$\begin{aligned}\hat{\mathbf{p}}_2^*&=\hat{\mathbf{v}}_{\mathrm{rel}}/(|\hat{\mathbf{v}}_{\mathrm{rel}}|+\hat{d})-\hat{d}\mathbf{F}^{\hat{P}'/\hat{P}}\cdot\mathbf{n}_2\mathbf{n}_2/|\mathbf{F}^{\hat{P}'/\hat{P}}\cdot\mathbf{n}_2|\\ \overline{\mathbf{p}}_2^*&=\overline{\mathbf{v}}_{\mathrm{rel}}/(|\overline{\mathbf{v}}_{\mathrm{rel}}|+\overline{d})-\overline{d}\mathbf{F}^{\overline{P}'/\overline{P}}\cdot\mathbf{n}_2\mathbf{n}_2/|\mathbf{F}^{\overline{P}'/\overline{P}}\cdot\mathbf{n}_2|,\end{aligned}\tag{23}$$

in accordance with Eqs. (11), (10), and (8). Here $\mathbf{F}^{\hat{P}'/\hat{P}}-\mathbf{F}^{\hat{P}'/\hat{P}}\cdot\mathbf{n}_1\mathbf{n}_1=\mathbf{F}^{\hat{P}'/\hat{P}}\cdot\mathbf{n}_2\mathbf{n}_2$ and $\mathbf{F}^{\overline{P}'/\overline{P}}-\mathbf{F}^{\overline{P}'/\overline{P}}\cdot\mathbf{n}_1\mathbf{n}_1=\mathbf{F}^{\overline{P}'/\overline{P}}\cdot\mathbf{n}_2\mathbf{n}_2$; and

$$\hat{\mathbf{v}}_{\mathrm{rel}}\triangleq\mathbf{v}^{\hat{P}}-\mathbf{v}^{\hat{P}'},\quad \overline{\mathbf{v}}_{\mathrm{rel}}\triangleq\mathbf{v}^{\overline{P}}-\mathbf{v}^{\overline{P}'}.\tag{24}$$

Explicit expressions for $\hat{\mathbf{v}}_{\mathrm{rel}}$ and $\overline{\mathbf{v}}_{\mathrm{rel}}$ can be obtained with the aid of the position vectors from $B^*$ to $\hat{P}$ and $\overline{P}$, given by

$$\mathbf{p}^{B^*/\hat{P}}=R\mathbf{n}_1,\quad \mathbf{p}^{B^*/\overline{P}}=-R\mathbf{n}_1.\tag{25}$$

Hence the velocities of $\hat{P}$ and $\overline{P}$ are

$$\mathbf{v}^{\hat{P}}=u_1\mathbf{n}_1+(u_2+Ru_3)\mathbf{n}_2,\quad \mathbf{v}^{\overline{P}}=u_1\mathbf{n}_1+(u_2-Ru_3)\mathbf{n}_2;\tag{26}$$

the velocities of $\hat{P}'$ and $\overline{P}'$ are

$$\mathbf{v}^{\hat{P}'}=\mathbf{v}^{\hat{P}}\cdot\mathbf{n}_1\mathbf{n}_1+\dot{y}\sin\alpha\mathbf{n}_2,\quad \mathbf{v}^{\overline{P}'}=\mathbf{v}^{\overline{P}}\cdot\mathbf{n}_1\mathbf{n}_1+\dot{y}\sin\alpha\mathbf{n}_2\tag{27}$$

so that, by reference to Eqs. (24),

$$\hat{\mathbf{v}}_{\mathrm{rel}}=(u_2+Ru_3-\dot{y}\sin\alpha)\mathbf{n}_2,\quad \overline{\mathbf{v}}_{\mathrm{rel}}=(u_2-Ru_3-\dot{y}\sin\alpha)\mathbf{n}_2.\tag{28}$$

One can now write the motion equations for Phase 1. These are

$$-m\dot{u}_1-\hat{F}+\overline{F}=0\tag{29}$$

$$-m\dot{u}_2-\mu_d\hat{F}\hat{\mathbf{p}}_2^*\cdot\mathbf{n}_2-\mu_d\overline{F}\overline{\mathbf{p}}_2^*\cdot\mathbf{n}_2=0\tag{30}$$

$$-I\dot{u}_3-\mu_d\hat{F}R\hat{\mathbf{p}}_2^*\cdot\mathbf{n}_2+\mu_d\overline{F}R\overline{\mathbf{p}}_2^*\cdot\mathbf{n}_2=0,\tag{31}$$

where $m$ is the mass of $B$, and $I$ is the central moment of inertia of $B$. In Phases 2 and 3 the motion is subject to the constraints $\hat{\mathbf{v}}_{\mathrm{rel}}=0$ and $\overline{\mathbf{v}}_{\mathrm{rel}}=0$, respectively. These become, in view of Eqs. (28),

$$u_2+Ru_3-\dot{y}\sin\alpha=0,\quad u_2-Ru_3-\dot{y}\sin\alpha=0.\tag{32}$$

Solving each of Eqs. (32) for $u_3$ one has

$$u_3=(\dot{y}\sin\alpha-u_2)/R\tag{33}$$

$$u_3=-(\dot{y}\sin\alpha-u_2)/R.\tag{34}$$

Equations (33) and (34) play the role of Eqs. (2) in Phases 2 and 3, respectively. Direct substitutions in Eqs. (3) from Eqs. (29)–(31) and (33) lead to equations governing the motion of $B$ in Phase 2. These are

$$-m\dot{u}_1-\hat{F}+\overline{F}=0\tag{35}$$

$$-(m+I/R^2)\dot{u}_2+I\ddot{y}\sin\alpha/R^2-2\mu_d\overline{F}\overline{\mathbf{p}}_2^*\cdot\mathbf{n}_2=0.\tag{36}$$

Similar equations govern the motion of $B$ in Phase 3, namely,

$$-m\dot{u}_1-\hat{F}+\overline{F}=0\tag{37}$$

$$-(m+I/R^2)\dot{u}_2+I\ddot{y}\sin\alpha/R^2-2\mu_d\hat{F}\hat{\mathbf{p}}_2^*\cdot\mathbf{n}_2=0,\tag{38}$$

obtained by substitutions from Eqs. (29)–(31) and (34) in Eqs. (3). Finally, the motion in Phase 4 is subject to both of the constraints in Eqs. (32). Solving these equations simultaneously for $u_2$ and $u_3$ one gets

$$u_2=\dot{y}\sin\alpha,\quad u_3=0.\tag{39}$$

These equations, together with Eqs. (29)–(31) lead, in conjunction with Eqs. (3), to

$$-m\dot{u}_1-\hat{F}+\overline{F}=0,\tag{40}$$

an equation which governs the motion of $B$ in this phase.

Next, letting $TC_{i-j}$ denote transition conditions from Phase $i$ to Phase $j$ $(i,j=1,2,3,4;i\neq j)$ and applying conditions (9) and (12) to the four phases, one obtains

$$\begin{aligned}&TC_{1-2}\quad |\hat{\mathbf{v}}_{\mathrm{rel}}|<\delta,\quad \hat{\mathbf{a}}_{\mathrm{rel}}\cdot\hat{\mathbf{v}}_{\mathrm{rel}}<0\\ &TC_{1-3}\quad |\overline{\mathbf{v}}_{\mathrm{rel}}|<\delta,\quad \overline{\mathbf{a}}_{\mathrm{rel}}\cdot\overline{\mathbf{v}}_{\mathrm{rel}}<0\\ &TC_{1-4}\quad |\hat{\mathbf{v}}_{\mathrm{rel}}|<\delta,\quad \hat{\mathbf{a}}_{\mathrm{rel}}\cdot\hat{\mathbf{v}}_{\mathrm{rel}}<0,\quad |\overline{\mathbf{v}}_{\mathrm{rel}}|<\delta,\quad \overline{\mathbf{a}}_{\mathrm{rel}}\cdot\overline{\mathbf{v}}_{\mathrm{rel}}<0;\end{aligned}\tag{41}$$

$$\begin{aligned}&TC_{2-1}\quad \mu_s\hat{F}<|\mathbf{F}^{\hat{P}'/\hat{P}}\cdot\mathbf{n}_2|\\ &TC_{2-3}\quad \mu_s\hat{F}<|\mathbf{F}^{\hat{P}'/\hat{P}}\cdot\mathbf{n}_2|,\quad |\overline{\mathbf{v}}_{\mathrm{rel}}|<\delta,\quad \overline{\mathbf{a}}_{\mathrm{rel}}\cdot\overline{\mathbf{v}}_{\mathrm{rel}}<0\\ &TC_{2-4}\quad |\overline{\mathbf{v}}_{\mathrm{rel}}|<\delta,\quad \overline{\mathbf{a}}_{\mathrm{rel}}\cdot\overline{\mathbf{v}}_{\mathrm{rel}}<0\end{aligned}\tag{42}$$

where

$$\mathbf{F}^{\hat{P}'/\hat{P}}\cdot\mathbf{n}_2=I(\ddot{y}\sin\alpha-\dot{u}_2)/R^2-\mu_d\overline{F}\overline{\mathbf{p}}_2^*\cdot\mathbf{n}_2;\tag{43}$$

$$\begin{aligned}&TC_{3-1}\quad \mu_s\overline{F}<|\mathbf{F}^{\overline{P}'/\overline{P}}\cdot\mathbf{n}_2|\\ &TC_{3-2}\quad \mu_s\overline{F}<|\mathbf{F}^{\overline{P}'/\overline{P}}\cdot\mathbf{n}_2|,\quad |\hat{\mathbf{v}}_{\mathrm{rel}}|<\delta,\quad \hat{\mathbf{a}}_{\mathrm{rel}}\cdot\hat{\mathbf{v}}_{\mathrm{rel}}<0\\ &TC_{3-4}\quad |\hat{\mathbf{v}}_{\mathrm{rel}}|<\delta,\quad \hat{\mathbf{a}}_{\mathrm{rel}}\cdot\hat{\mathbf{v}}_{\mathrm{rel}}<0\end{aligned}\tag{44}$$

where

$$\mathbf{F}^{\overline{P}'/\overline{P}}\cdot\mathbf{n}_2=I(\ddot{y}\sin\alpha-\dot{u}_2)/R^2-\mu_d\hat{F}\hat{\mathbf{p}}_2^*\cdot\mathbf{n}_2;\tag{45}$$

and

$$\begin{aligned}&TC_{4-1}\quad \mu_s\hat{F}<|\mathbf{F}^{\hat{P}'/\hat{P}}\cdot\mathbf{n}_2|,\quad \mu_s\overline{F}<|\mathbf{F}^{\overline{P}'/\overline{P}}\cdot\mathbf{n}_2|\\ &TC_{4-2}\quad \mu_s\overline{F}<|\mathbf{F}^{\overline{P}'/\overline{P}}\cdot\mathbf{n}_2|\\ &TC_{4-3}\quad \mu_s\hat{F}<|\mathbf{F}^{\hat{P}'/\hat{P}}\cdot\mathbf{n}_2|\end{aligned}\tag{46}$$

where

$$\mathbf{F}^{\overline{P}'/\overline{P}}\cdot\mathbf{n}_2=\mathbf{F}^{\hat{P}'/\hat{P}}\cdot\mathbf{n}_2=0.5m\ddot{y}\sin\alpha.\tag{47}$$

Finally, note that, in view of Eqs. (43), (45), and (47), the definitions of $\hat{\mathbf{p}}_2^*$ and $\overline{\mathbf{p}}_2^*$ in Eqs. (23) do not reflect the fact that different expressions for $\mathbf{F}^{\hat{P}'/\hat{P}}\cdot\mathbf{n}_2$ and $\mathbf{F}^{\overline{P}'/\overline{P}}\cdot\mathbf{n}_2$ are valid at different phases. Thus, the phase preceding a transition to Phase 2, say, may be either Phase 3, in which event $\mathbf{F}^{\overline{P}'/\overline{P}}\cdot\mathbf{n}_2$ is given by Eq. (45); or Phase 4, in which event $\mathbf{F}^{\overline{P}'/\overline{P}}\cdot\mathbf{n}_2$ is given Eqs. (47). This fact can be taken into account if the definitions of $\hat{\mathbf{p}}_2^*$ and $\overline{\mathbf{p}}_2^*$ are modified as follows:

$$\begin{aligned}\hat{\mathbf{p}}_2^*=&\,\hat{\mathbf{v}}_{\mathrm{rel}}/(|\hat{\mathbf{v}}_{\mathrm{rel}}|+\hat{d}_{21}+\hat{d}_{23}+\hat{d}_{41}+\hat{d}_{43})\\ &-(\hat{d}_{21}+\hat{d}_{23})[\mathbf{F}^{\hat{P}'/\hat{P}}\cdot\mathbf{n}_2]_2\mathbf{n}_2/[|\mathbf{F}^{\hat{P}'/\hat{P}}\cdot\mathbf{n}_2|]_2-(\hat{d}_{41}+\hat{d}_{43})\\ &\times[\mathbf{F}^{\hat{P}'/\hat{P}}\cdot\mathbf{n}_2]_4\mathbf{n}_2/[|\mathbf{F}^{\hat{P}'/\hat{P}}\cdot\mathbf{n}_2|]_4\end{aligned}\tag{48}$$

$$\begin{aligned}\overline{\mathbf{p}}_2^*=&\,\overline{\mathbf{v}}_{\mathrm{rel}}/(|\overline{\mathbf{v}}_{\mathrm{rel}}|+\overline{d}_{31}+\overline{d}_{32}+\overline{d}_{41}+\overline{d}_{42})\\ &-(\overline{d}_{31}+\overline{d}_{32})[\mathbf{F}^{\overline{P}'/\overline{P}}\cdot\mathbf{n}_2]_3\mathbf{n}_2/[|\mathbf{F}^{\overline{P}'/\overline{P}}\cdot\mathbf{n}_2|]_3-(\overline{d}_{41}+\overline{d}_{42})\\ &\times[\mathbf{F}^{\overline{P}'/\overline{P}}\cdot\mathbf{n}_2]_4\mathbf{n}_2/[|\mathbf{F}^{\overline{P}'/\overline{P}}\cdot\mathbf{n}_2|]_4,\end{aligned}\tag{49}$$

where, in connection with a transition from Phase 4 to Phase 1, $\hat{d}_{41}|_{t=t_2}=1$ and $\hat{d}_{41}|_{t\neq t_2}=0$, and $[\mathbf{F}^{\hat{P}'/\hat{P}}\cdot\mathbf{n}_2]_4$ is given by Eqs. (47); etc. Furthermore, $\hat{\mathbf{p}}_2^*=\hat{\mathbf{v}}_{\mathrm{rel}}/|\hat{\mathbf{v}}_{\mathrm{rel}}|$ whenever $|\hat{\mathbf{v}}_{\mathrm{rel}}|\neq 0$, and $\hat{\mathbf{p}}_2^*$

$= -[\mathbf{F}^{\hat{P}'/\hat{P}} \cdot \mathbf{n}_2]_i \mathbf{n}_2 / [|\mathbf{F}^{\hat{P}'/\hat{P}} \cdot \mathbf{n}_2|]_i$ at $t_2$, the first instant of slipping at $\hat{P}$ in Phases 1 or 3—following a transition from either Phase 2 (then $i=2$) or Phase 4 (then $i=4$); etc.

A program simulating motions of $B$ in $N$ was constructed, based upon Eqs. (29)–(31), (35)–(36), (37)–(38), and (40), which govern the motion in Phases 1, 2, 3, and 4, respectively; and on transition conditions (41), (42), (44), and (46). Equations (29)–(31) were coded explicitly, whereas Eqs. (35)–(36), (37)–(38), and (40) were generated by the manipulation of terms of Eqs. (29)–(31), in conjunction with Eqs. (33), (34), and (39), respectively, as indicated by Eqs. (3) and (2). One second simulations were run for $f=f_1$, $f=f_1+1, \ldots, f=f_1+M$ Hz, where $f_1$ and $M$ were selected so that the associated spin differs from zero. The average angular speed $\Omega(f)$ was evaluated for each run and plotted in Fig. 3($a$) as a function of $f$, which seems to resemble $\Omega(f)$ in Fig. 2. Slight deviations can be expected since such factors as the inertial properties of $P_B$ and $P_H$, and the convexity of the surfaces of $P_B$ and $P_H$ when deformed were disregarded, as was the fact that $\mu_d$ in the neighborhood of $\hat{P}$ might be different from $\mu_d$ in the neighborhood of $\bar{P}$.

A number of experimental runs of the indicated program showed that the maximum expected value of $\mathbf{a}_{\text{rel}} \cdot \mathbf{p}_2$ is 200 m/sec$^2$, so that, with $\Delta t = 0.00002$ sec, $\delta$ should be smaller than 0.004 m/sec. In fact, it was found experimentally that with $\delta = 0.0001$ no chatters arise. Now, Edelstein [1] in his experiments used plates ($P_B$ and $P_H$) made of poly-methyl-methacrylate (Dupont's PMMA is commercially known as Perspex). Measuring the static and dynamic coefficients of friction in question, he found $\mu_s = 0.19 \pm 0.084$ and $\mu_d = 0.066 \pm 0.024$. Accordingly, Fig. 3($a$) was obtained with $\mu_s = 0.19$, $\mu_d = 0.066$, $b = 14$ N-sec/m($\xi = 0.007$) and $m = 0.224$ Kg, $I = 0.0000328$ Kg-m$^2$, $k = 2313000$ N-m. Interestingly, $q_2$ shows a cyclic behavior with an amplitude of four micrometers. One may conclude that lateral space in the order of a few micrometers suffices for the spinning ball effect to develop.

Additional runs show that $\Omega(f)$ changes only slightly if different values of $\mu_s$ or $\mu_d$ are used, and that the inclusion of gravity forces produces almost no effect on the simulation results. On the other hand, different values of $b$ affect the results considerably, as shown in Fig. 3($b$) for $b = 18$ N-sec/m. A plot similar to that in Fig. 3($b$) is obtained if the preload $F$ is increased to 420 N. These results point out ways to decay the spinning ball effect.

Finally, the reverse of the sign of $\Omega(f)$ associated with the passage through $f_n = 1/2\pi\sqrt{2k/m} = 720$ Hz deserves attention. It can be explained with the aid of Figs. 7 and 8, showing portions of the time histories of $q_1$ (thin line), $y$ and $\dot{y}$ (thick lines) for $f = 710$ Hz($< f_n$) and $f = 730$ Hz($> f_n$), respectively, and the associated phases of motion. Specifically, Fig. 7 shows that $q_1$ leads $y$ and that $q_{1\,\max}/y_{\max} \cong 26$ (note that $y$ in Fig. 7 is multiplied by 26). Hence, when $q_1 > 0$, spring $S_H$ is pressed ($\mathbf{F}^{\hat{P}'/\hat{P}} \cdot \mathbf{n}_1 \cong 650$ N) and spring $S_B$ is relaxed ($\mathbf{F}^{\bar{P}'/\bar{P}} \cdot \mathbf{n}_1 \cong 10$ N). Consequently, stick arises at $\hat{P}$ and slip arises at $\bar{P}$ (Phase 2). Moreover, when $q_1 > 0$, then $\dot{y} \lesssim 0$ (i.e., when $q_1 > 0$, then $\dot{y}$ is "mostly negative"), hence, $P_H$ "pulls" $\hat{P}$ to the right, making $B'$ spin clockwise about $\hat{P}$. Similarly, when $q_1 < 0$, spring $S_B$ is pressed and spring $S_H$ is relaxed. Consequently, there is stick at $\bar{P}$ and slip at $\hat{P}$ (Phase 3); and, since $\dot{y} \gtrsim 0$ (i.e., $\dot{y}$ is "mostly positive"), $P_B$ "pulls" $\bar{P}$ to the left, making $B'$ spin clockwise about $\bar{P}$. By the same token, Fig. 8 shows that $q_1$ leads $y$ an amount somewhat smaller than $\pi$. Also, $q_{1\,\max}/y_{\max} \cong 20$ (note that $y$ is Fig. 8 multiplied by 20). Hence, when $q_1 > 0$, spring $S_H$ is pressed and spring $S_B$ is relaxed. Again, sticking arises at $\hat{P}$ and slipping at $\bar{P}$ (Phase 2). However, now $\dot{y} \gtrsim 0$, hence, $P_H$ "pulls" $\hat{P}$ to the left, making $B'$ spin counterclockwise about $\hat{P}$. Finally, when $q_1 < 0$, spring $S_B$ is pressed and spring $S_H$ is relaxed. Stick at $\bar{P}$ and slip at $\hat{P}$ result in (Phase 3); and, since $\dot{y} \lesssim 0$, $P_B$ "pulls" $\bar{P}$ to the right, making $B'$ spin counterclockwise about $\bar{P}$. When $f \ll f_n$, then $q_{1\,\max}/y_{\max}$ comes close to one, and when $f \gg f_n$, then $q_{1\,\max}/y_{\max}$ comes close to zero. In either case the elastic deflections are small, preventing slip from taking place at either $\hat{P}$ or at $\bar{P}$ (Phase 4).

The phase plots in Figs. 7 and 8 indicate that the switches from one phase to another are free of chatters.

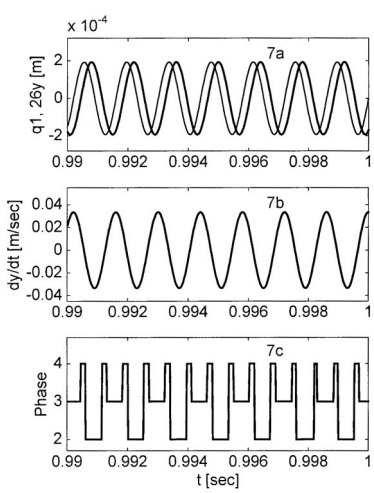



**Fig. 7 Portions of the time histories of $q_1$ (thin line), $y$, and $\dot{y}$ for $f = 710$ Hz**

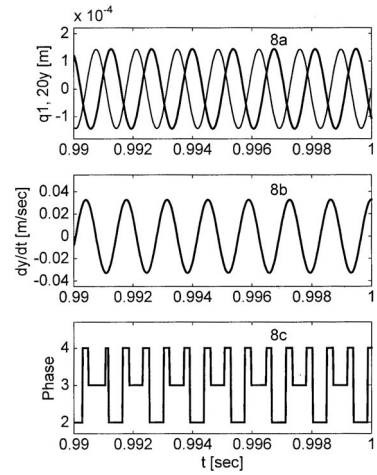



**Fig. 8 Portions of the time histories of $q_1$ (thin line) $y$, and $\dot{y}$ for $f = 730$ Hz**

## Conclusions

It may occur that simplified transition conditions suffice to capture stick-slip motions with fidelity. For example, this might be the case when no reciprocating motion is involved. By way of contrast, the modified conditions presented here are applicable to all stick-slip problems; and are shown to provide reasonable predictions of a complex stick-slip motion, such as that of the spinning ball.

It is finally worth pointing out that rolling elements of bearings may be subjected to the spinning ball effect, increasing wear even in the absence of rotational motion. The same effect underlies the following phenomenon, namely, that, at times, screws become loose when subject to linear vibrations. An understanding of the spinning ball problem may help in preventing these undesirable effects.

## References


[1] Edelstein, E., 1993, ''Rotation of a Spherical Ball Clamped Between Spherical Cavities Due to Linear Reciprocating Motion,'' Ph.D. dissertation, Technion, Haifa.

[2] Kane, T. R., and Levinson, D. A., 1985, *Dynamics: Theory and Application*, McGraw-Hill, New York.

[3] Synnestvedt, R. G., 1996, ''An Effective Method for Modeling Stiction in Multibody Dynamic Systems,'' ASME J. Dyn. Syst., Meas., Control, **118**, pp. 172–176.

[4] Hsien-I, Y., and Jeng-Hong, H., 1995, ''The Influence of Friction-Speed Relation on the Occurrence of Stick-Slip Motion,'' ASME J. Tribol., **117**, pp. 450–455.

[5] Huston, R. L., Passerello, C., Winget, J. M., and Sears, J., 1979, ''On the Dynamics of a Weighted Bowling Ball,'' Paper No. 79-WA/APM-17, ASME Winter Annual Meeting, Dec. 2–7, New York.

[6] Kane, T. R., and Levinson, D. A., 1978, ''A Realistic Solution of the Symmetric Top Problem,'' ASME J. Appl. Mech., **45**, pp. 903–909.

[7] Karnopp, D., 1985, ''Computer Simulation of Stick-Slip Friction in Mechanical Dynamic Systems,'' ASME J. Dyn. Syst., Meas., Control, **107**, pp. 100–103.

[8] Southward, S. C., Radcliff, C. J., and MacCluer, C. R., 1991, ''Robust Nonlinear Stick-Slip Friction Compensation,'' ASME J. Dyn. Syst., Meas., Control, **113**, pp. 639–645.

[9] Jia-Yush, Y., Chih-Jung, H., and Shu-Shung, L., 1997, ''Stability of PDF Controller With Stick-Slip Friction Device,'' ASME J. Dyn. Syst., Meas., Control, **119**, pp. 486–490.

[10] Tan, X., and Rogers, R. J., 1998, ''Simulation of Friction in Multi-degree of Freedom Vibration Systems,'' ASME J. Dyn. Syst., Meas., Control, **120**, pp. 144–146.

[11] Han, I., Gilmore, B. J., and Ogot, M. M., 1993, ''The Incoporation of Arc Boundaries and Stick/Slip Friction in a Rule-Based Simulation Algorithm for Dynamic Mechanical Systems with Changing Topologies,'' ASME J. Mech. Des., **115**, pp. 423–434.

[12] Pfeiffer, F., 1996, ''Complementary Problems of Stick-Slip Vibrations,'' ASME J. Vibr. Acoust., **118**, pp. 177–183.

[13] Djerassi, S., 1999, ''Dynamical Analysis of Systems with Varying Constraints,'' ASME J. Appl. Mech., **66**, pp. 786–793.

# The Fiber Composite With Nonlinear Interface—Part I: Axial Tension

**A. J. Levy**
Department of Mechanical, Aerospace
and Manufacturing Engineering,
Syracuse University,
Syracuse, NY 13244-1240
Mem. ASME

*This paper treats the effective axial tension response of a composite consisting of fibers that debond from the matrix according to nonlinear Needleman-type cohesive zones. A second, related paper (Part II) treats effective antiplane shear response. The composite cylinders representation of a representative volume element (RVE) is employed throughout. For axial tension loading a simple rotationally symmetric boundary value problem for a single composite cylinder is solved. Bounds on the total potential energy and the total complementary energy are shown to coincide and an exact solution for axial extension and Poisson contraction of an RVE of the composite is obtained. Nonlinear interfacial debonding, however, is shown to have a negligible effect on extensional response and only a small, though potentially destabilizing, effect on Poisson contraction response.* [S0021-8936(00)02004-3]

## 1 Introduction

This paper (Part I) is concerned with the effective tensile response of unidirectional randomly arrayed fiber composites containing interfaces that suffer nonlinear displacement discontinuities. A companion paper (Part II) treats antiplane shear of this class of composite. Previous work on the nonlinear interface composite has dealt with the case of loading transverse to the fiber direction. In particular, transverse bulk response has been analyzed using Hashin's composite cylinder model[1] ([1]) and transverse shear response has been studied employing the dilute estimate ([2]) and the Mori-Tanaka model ([3]). All of this work, including the present pair of articles, assumes that separation/slip phenomena at a fiber-matrix interface can be captured by uniform Needleman-type cohesive zones of vanishing thickness ([4]). Although our formulation can accommodate general interface constitutive relations in this paper we employ a model, derivable from a potential, which allows for interface failure in normal mode, the shear mode being inactive due to the nature of the geometry and the loading.

Formally, the axial tension problem is no different than the transverse bulk problem in that we employ the direct method of composite materials theory together with a geometrical model to estimate the uniform fields associated with the fiber phase. Here we employ the composite cylinder representation of a representative volume element (RVE) ([5]) exclusively. It can be shown that the constitutive relation for a single composite cylinder, subject to axial (or transverse radial) tension, is of the form

$$\sigma = F(\varepsilon, \chi^+, \chi^-, c, \omega_0),$$
$$G_0(\varepsilon, \chi^+, \chi^-, c, \omega_0, \xi_k) = 0. \tag{1}$$

In (1) $\sigma$, $\varepsilon$ are corresponding mean effective stress and strain components; $\chi^+, \chi^-$ are elastic properties of matrix and fiber, respectively; $c$ is the volume concentration of fiber; and the $\xi_k$, $k=1,2,\ldots$ are interface constitutive parameters. Relation $(1_1)$ is the effective property relation for a single composite cylinder. The internal variable $\omega_0$ is the uniform interface displacement discontinuity and appears in $(1_1)$ through expressions for the uniform strain in the fiber phase and a component of the displacement discontinuity tensor.[2] Relation $(1_2)$ is determined from the interface boundary condition for the composite cylinder and it depends explicitly on the interface constitutive relation prescribed in the form of an interface force-displacement discontinuity relation. The formulation resulting in (1), for a single composite cylinder, is then mapped onto an entire RVE via the composite cylinders construct. The elastic fields obtained from the uniform strain boundary value problem and the uniform stress boundary value problem are then used to bound the true response of the RVE. If the bounds coincide then there is an exact solution for the effective response of an RVE. Note that although the framework presented in (1) bears a resemblance to continuum damage constitutive models, the equations are not, however, postulated a priori. They arise naturally from the direct method of composite materials theory, a particular representation of an RVE and a particular interface constitutive relation.

In the next section we consider axial tension of a single composite cylinder and derive equivalent constitutive relations governing its response. In the section that follows we use variational arguments to show that equations developed for the composite cylinder apply to an RVE consisting of a volume-filling configuration of composite cylinders of varying diameters. In particular, the effective property relation and the single evolution equation governing the uniform interface separation variable are shown by variational bounding to characterize the exact solution for axial tension of the RVE. Sections on Poisson contraction, limiting cases, and predictions follow. Overall, the treatment is straightforward and the results extend those obtained for axial tension response of rigid interface composites ([5]), and linearly separating interface composites ([6]). We conclude the paper with a brief summary of pertinent results.

## 2 The Composite Cylinder

Consider the composite cylinder shown in Fig. 1. The fiber has elastic modulus $E^-$ and Poisson ratio $\nu^-$ and occupies domain $\Omega$. The matrix has elastic properties $E^+, \nu^+$ and occupies domain $B$-$\Omega$. Orient a Cartesian coordinate system with origin $\mathbf{o}$ and basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$, and a cylindrical coordinate system with basis

[1]It can be shown that the response is identical to that predicted by the three-phase model.

[2]The difference between the applied boundary strain and the mean strain of the composite cylinder.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, May 7, 1999; final revision, June 25, 2000. Associate Technical Editor: I. M. Daniel. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

 Copyright © 2000 by ASME 

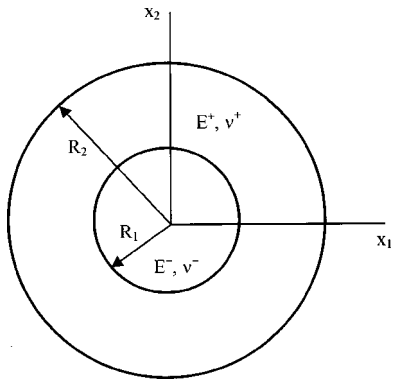



**Fig. 1 Composite cylinder geometry**

$\{\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_z = \mathbf{e}_3\}$, such that the fiber axis points in the $\mathbf{e}_3$ direction. Assume that the end caps are subjected to a displacement $\mathbf{E}_H\mathbf{r}$ where $\mathbf{E}_H$ is a uniform strain $\mathbf{E}_H = E_{rr}\mathbf{e}_r\otimes\mathbf{e}_r + E_{\theta\theta}\mathbf{e}_\theta\otimes\mathbf{e}_\theta + e\mathbf{e}_3\otimes\mathbf{e}_3$ and $\mathbf{r}(\mathbf{x})$ is the position vector from the origin $\mathbf{o}$ to a point $\mathbf{x}$ on the bases. Throughout this section we assume all elastic fields are rotationally symmetric so that the displacement field is of the form, $\mathbf{u} = u_r(r)\mathbf{e}_r + u_z(z)\mathbf{e}_z$. Rotationally symmetric solutions of Navier's equation, the stress-displacement relations, and the uniform strain boundary condition are given by

$$\begin{aligned}
&\mathbf{u} = a_1 r\mathbf{e}_r + ez\mathbf{e}_z,\\
&\mathbf{S} = (2(\lambda^- + \mu^-)a_1 + \lambda^- e)[\mathbf{e}_r\otimes\mathbf{e}_r + \mathbf{e}_\theta\otimes\mathbf{e}_\theta]\\
&\quad + (2\lambda^- a_1 + (\lambda^- + 2\mu^-)e)\mathbf{e}_z\otimes\mathbf{e}_z,\\
&r\in(0, R_1),\\
&\mathbf{u} = (a_2 r + a_3 r^{-1})\mathbf{e}_r + ez\mathbf{e}_z,\\
&\mathbf{S} = (2(\lambda^+ + \mu^+)a_2 - 2\mu^+ a_3 r^{-2} + \lambda^+ e)\mathbf{e}_r\otimes\mathbf{e}_r + (2(\lambda^+ + \mu^+)a_2\\
&\quad + 2\mu^+ a_3 r^{-2} + \lambda^+ e)\mathbf{e}_\theta\otimes\mathbf{e}_\theta + (2\lambda^+ a_2 + (\lambda^+ + 2\mu^+)e)\mathbf{e}_z\otimes\mathbf{e}_z,\\
&r\in(R_1, R_2),
\end{aligned} \tag{2}$$

where $(\lambda, \mu)$ are the Lamé moduli and $\mathbf{S}$ is the stress tensor. The constants $(a_1, a_2, a_3)$ appearing in (2) are obtained from interface boundary conditions at $r = R_1$ and the stress-free boundary condition on the lateral surface,

$$\begin{aligned}
&\operatorname{Lim}\mathbf{s}(\mathbf{e}_r) = \mathbf{s}_I(\mathbf{e}_r), \quad r\uparrow R_1,\\
&\operatorname{Lim}\mathbf{s}(-\mathbf{e}_r) = -\mathbf{s}_1(\mathbf{e}_r), \quad r\downarrow R_1,\\
&\mathbf{s}(\mathbf{e}_r) = \mathbf{0}, \quad r = R_2,
\end{aligned} \tag{3}$$

where $\mathbf{s}(\mathbf{n})$ is the traction vector on a surface with normal $\mathbf{n}$. The vector $\mathbf{s}_I(\mathbf{n})$ characterizes the constitutive characteristics of the interface and may be written in the form

$$\mathbf{s}_I(\mathbf{n}) = f(u)\mathbf{n} \tag{4}$$

where $u$ is the normalized interface displacement jump in normal mode, $u = [u_r]/R_1$, and $f$ is a prescribed function. The constitutive hypothesis (4) may be expressed in more general terms that include two additional tangential components of interface traction. All functions then depend on three displacement jump components and possibly the interface coordinates as well. Owing to rotationally symmetric solutions for the displacement and an assumed uniformity of interface these additional terms are absent from (4).

Boundary conditions (3) and (4) can be employed in a straightforward manner to determine constants $(a_1, a_2, a_3)$,

$$a_1 = \frac{f(u) - \lambda^- e}{2(\lambda^- + \mu^-)}, \quad a_2 = \frac{-cf(u)}{2(\lambda^+ + \mu^+)(1-c)} - \frac{\lambda^+ e}{2(\lambda^+ + \mu^+)}, \tag{5}$$

$$a_3 = \frac{f(u)}{2\mu^+(R_2^{-2} - R_1^{-2})},$$

where $c$, the fiber volume concentration, is given by $(R_1/R_2)^2$. With the aid of (5) the elastic fields (2) can be written in terms of the unknown interface displacement jump $u$. In particular, the equation governing this quantity follows directly from the relation

$$u = R_1^{-1}[\operatorname{Lim}_{r\downarrow R_1} u_r(r) - \operatorname{Lim}_{r\uparrow R_1} u_r(r)] \tag{6}$$

and the elastic field solution for $\mathbf{u}\cdot\mathbf{e}_r$ in the matrix and in the fiber. The resulting equation is of the form

$$\begin{aligned}
&u + \alpha f(u) - he = 0,\\
&\alpha = \frac{c}{2(\lambda^+ + u^+)(1-c)} + \frac{1}{2\mu^+(1-c)} + \frac{1}{2(\lambda^- + \mu^-)},\\
&h = \frac{\lambda^-}{2(\lambda^- + \mu^-)} - \frac{\lambda^+}{2(\lambda^+ + \mu^+)} = \nu^- - \nu^+.
\end{aligned} \tag{7}$$

With the prescription of $f$ and the solution of (7) the elastic field in the composite cylinder is completely determined.

An equivalent[3] axial constitutive relation for the composite cylinder may be obtained by requiring a homogeneous cylinder of identical geometry and loading to have the same axial mean stress, i.e., $\sigma = \mathbf{e}_z\cdot\overline{\mathbf{S}}\mathbf{e}_z$, where $\overline{\mathbf{S}}$ is the mean stress of the composite cylinder defined by

$$\overline{\mathbf{S}} = \frac{1}{\operatorname{vol}(B)}\int_B \mathbf{S}\,dv. \tag{8}$$

Utilizing (2) and (5) results in the equivalent axial constitutive relation for the composite cylinder,

$$\sigma = E_M e + 2c(\nu^- - \nu^+)f(u), \tag{9}$$

where

$$E_M = \lfloor cE^- + (1-c)E^+ \rfloor. \tag{10}$$

(Note that the second term on the right-hand side of (9) is proportional to the difference in Poisson ratio of matrix and fiber and, when it vanishes, the resulting expression is the rule of mixtures.) Thus, (7) and (9) completely characterize the equivalent axial constitutive response of the composite cylinder provided a suitable form of interface force-separation relation is prescribed. In the next section we show that the equivalent constitutive model (7) and (9) is in fact the effective constitutive relation for an entire RVE provided certain geometrical constraints are satisfied.

The analysis of the traction boundary value problem for the composite cylinder is carried out in similar fashion. Note that although we utilize a uniform traction boundary condition ($\mathbf{s}(\mathbf{n}) = \sigma\mathbf{e}_z$) applied to the end caps to compute the equivalent response of the composite cylinder, we may take the previous solution (2), (5) (for displacement boundary conditions) as the solution to the traction problem. We justify this on the basis of Saint-Venant's principle, which asserts the validity of solution (2), (5) away from the cylinder ends. Thus, we interpret $\sigma\mathbf{e}_z$ as the mean applied traction on the end caps. To compute the equivalent response use the $\mathbf{e}_z\otimes\mathbf{e}_z$ component of the effective property relation ([7,8])

$$e = \frac{1}{E^+}\sigma + c\left(\frac{1}{E^-} - \frac{1}{E^+}\right)S_{zz}^- - c\left(\frac{\nu^-}{E^-} - \frac{\nu^+}{E^+}\right)(S_{rr}^- + S_{\theta\theta}^-) + A, \tag{11}$$

where $\sigma$ is the mean applied axial stress on the cylinder ends, $e$ is the axial mean effective strain,[4] $S_{rr}^-, S_{\theta\theta}^-, S_{zz}^-$, are the cylindrical components of the mean stress in the fiber phase, and $A$ is the $\mathbf{e}_z\otimes\mathbf{e}_z$ component of the displacement discontinuity tensor (this

[3]Distinct from an effective constitutive relation for an RVE.

[4]Generally the sum of the mean strain and the displacement discontinuity tensor.

term is exactly zero for the given loading). Equations (2), (5), and (11) imply that the equivalent axial constitutive relation, for traction boundary conditions, is identical to that computed for displacement boundary conditions ((9)).

## 3 Variational Bounding

Now the composite cylinders model requires the geometry of the RVE to consist of long cylindrical fibers each of which is surrounded by a cylindrical cage of matrix material. In order to obtain a volume filling configuration composite cylinders are assumed to vary in diameter down to the vanishingly small. This geometry, along with the requirement of identical mean stresses and strains in each composite cylinder composing the RVE, necessitates the satisfaction of certain geometrical constraints. In particular, the cylinders must be geometrically similar, i.e., the ratio $R_1/R_2$ must be fixed. In addition, because the interface force law (4) generally requires an additional length scale for its prescription, additional constraints may need to be prescribed in terms of it. We will comment on this later on in the sequel (Section 6) after we specify a form of interface constitutive function $f$. For now, assume that all geometrical constraints are satisfied so that mean stress and strain fields computed from (7) and (9) are identical for all composite cylinders forming the volume filling composite cylinders geometry.

In order to prove that relations (7), (9) apply to an entire RVE of the composite we follow the analysis of the transverse bulk response of nonlinear interface composites ([1]). This treatment utilizes the energetic arguments of Hashin and Rosen [5]. Thus we show that upper and lower bounds to the true strain energy of the composite, based on admissible fields obtained from the composite cylinder displacement and traction problems, coincide. First, assume that the displacement field in a composite cylinder, with displacement vector $\mathbf{u}=\mathbf{E}_H\mathbf{r}$ prescribed on its end caps, forms a kinematically admissible displacement field[5] for the displacement in the entire RVE of the composite. This follows from the fact that a rigid-body displacement superimposed on the boundary displacement of a composite cylinder, while leaving the strain energy unaffected, renders homogeneous boundary displacements for the RVE. Now the potential energy $\Phi\{e\}$ stored in a single composite cylinder subject to the displacement boundary condition can be calculated from (2), (5), and (7) and is given by

$$\Phi\{e\}=\left\{\frac{1}{2}[(1-c)E^{+}+cE^{-}]e^{2}+\frac{c}{\pi}\int_{0}^{2\pi}\varepsilon_{I}d\theta\right.$$
$$+\left(\frac{c(1+\nu^{-})(1-2\nu^{-})}{E^{-}}+\frac{c^{2}(1+\nu^{+})(1-2\nu^{+})}{(1-c)E^{+}}\right.$$
$$\left.\left.+\frac{c(1+\nu^{+})}{(1-c)E^{+}}\right)f^{2}\right\}v(B), \tag{12}$$

where $\varepsilon_1$ is the interface potential defined by

$$\varepsilon_{I}=\int^{u}f\,d\tau, \tag{13}$$

and $v(B)$ is the volume of the composite cylinder. (Note that the vanishing of the derivative $D_u\Phi$ implies relation (7) while $\sigma=D_e[v^{-1}(B)\Phi\{e\}]$ implies relation (9).) An upper bound to the potential energy for the representative volume element of the composite, based on the admissible displacement field, is $\Phi\{e\}v/v(B)$ where $v$ is the volume of the representative volume element.

Next, assume that the homogeneous stress field, produced by the traction boundary condition $\mathbf{s}(\mathbf{n})=\sigma\mathbf{e}_z$ applied to the end caps of the composite cylinder, is a statically admissible stress field[6] for the entire representative volume element. Then the complementary energy for a single composite cylinder may be determined in a manner similar to $\Phi$ and is given by

$$\Psi\{\sigma\}=\left\{\frac{1}{2[(1-c)E^{+}+cE^{-}]}\sigma^{2}+\frac{c}{\pi}\int_{0}^{2\pi}\varepsilon_{IC}d\theta\right.$$
$$-\frac{2c(\nu^{-}-\nu^{+})}{[(1-c)E^{+}+cE^{-}]}\sigma f+\left(\frac{2c^{2}(\nu^{-}-\nu^{+})^{2}}{[(1-c)E^{+}+cE^{-}]}\right.$$
$$+\frac{c^{2}(1+\nu^{+})(1-2\nu^{+})}{(1-c)E^{+}}+\frac{c(1+\nu^{+})}{(1-c)E^{+}}$$
$$\left.\left.+\frac{c(1+\nu^{-})(1-2\nu^{-})}{E^{-}}\right)f^{2}\right\}v(B), \tag{14}$$

where $\varepsilon_{IC}$ is defined by

$$\varepsilon_{IC}=\int^{f}u\,d\tau, \tag{15}$$

and use has been made of (2). (The vanishing of the derivative $D_f\Psi$ yields the relation

$$u+\left(\alpha+\frac{2ch^{2}}{E_{M}}\right)f(u)-\frac{h}{E_{M}}\sigma=0, \tag{16}$$

where $\alpha$, $h$ are defined in (7) and $E_M$ is defined in (10) while $e=D_\sigma[v^{-1}(B)\Psi\{\sigma\}]$ yields (9).) An upper bound to the complementary energy for the representative volume element of the composite, based on the statically admissible stress field, is then given by $\Psi\{\sigma\}v/v(B)$. The potential (complementary) energy given by (12) ((14)) may be regarded as the strain (stress) energy for an equivalent homogeneous cylinder owing to the displacement (traction) condition applied on its end caps. Upper and lower bounds on the true strain energy ($U$) of the displacement problem follow from the bounding theorem ([9]).

$$\int_{\partial B}\overline{\mathbf{s}}\cdot\hat{\mathbf{u}}\,da-U_{K}\{\overline{\mathbf{S}}\}\leqslant U\leqslant U_{C}\{\tilde{\mathbf{E}}\}, \tag{17}$$

where $\tilde{\mathbf{E}}$ is a kinematically admissible strain field, $\overline{\mathbf{S}}$ is a statically admissible stress field with $\overline{\mathbf{s}}$ the corresponding surface traction, and $U_K\{\overline{\mathbf{S}}\}(U_C\{\tilde{\mathbf{E}}\})$ the stress (strain) energy. For the present problem (17) may be written as

$$2\sigma e v-\Psi\{\sigma\}v/v(B)\leqslant U\leqslant\Phi\{e\}v/v(B), \tag{18}$$

where $\Phi$, $\Psi$ are given by (12), (14). We need to show that the bounds on $U$ shown in (18) coincide. First note that, as required, combining (9) and (16) recovers (7) so the constitutive relation for a single composite cylinder, based on uniform strain and uniform stress boundary conditions, is the same. Next, to show equivalence of bounds simply substitute (9), (16) into (14) and note that

$$\varepsilon_{IC}=uf-\varepsilon_{I}. \tag{19}$$

Thus, the true strain energy of an RVE of the composite is given by $U=\Phi\{e\}v/v(B)$ and the stress-strain relation (9) and (7) (or (16)) follows from

$$\sigma=\frac{1}{v}\frac{\partial U}{\partial e},\quad 0=\frac{1}{v}\frac{\partial U}{\partial u}. \tag{20}$$

## 4 Effective Poisson Contraction

The lateral response of the composite, when subjected to axial loading, can be obtained in a straightforward manner since the upper and lower bounds to the strain energy of the RVE coincide. Thus, from (2) and (5), compute radial strain $E_{rr}$, or circumfer-

[5]One that satisfies the displacement boundary condition and is twice continuously differentiable.

[6]One that satisfies the equilibrium equation and traction boundary condition.

ential strain $E_{\theta\theta}$, and evaluate it on the outer boundary $r=R_2$ of the composite cylinder,

$$E_{rr}=E_{\theta\theta}=\nu^+ e+\frac{2c}{1-c}\frac{1-\nu^{+2}}{E^+}f(u). \tag{21}$$

Equation (21), when coupled to (7), completely characterizes the transverse strain of the RVE due to axial tension.

## 5 Limiting Cases

Limiting conditions of rigid interface behavior may be obtained as follows. Set the interface displacement jump $u$ to zero so that the interface force becomes, by (7), $f(0)=he/\alpha$, where $h$ and $\alpha$ are defined in (7). When this value of $f(0)$ is substituted into (9) and (21) we get

$$\begin{aligned} E^*&=cE^-+(1-c)E^+ + \frac{c(1-c)E^-E^+(\nu^- - \nu^+)^2}{c(1+\nu^+)(1-2\nu^+)E^- + (1+\nu^+)E^- + (1+\nu^-)(1-2\nu^-)E^+},\\ \nu^*&=c\nu^- + (1-c)\nu^+ + \frac{c(1-c)[(1+\nu^+)(1-2\nu^+)E^- - (1+\nu^-)(1-2\nu^-)E^+](\nu^- - \nu^+)}{c(1+\nu^+)(1-2\nu^+)E^- + (1+\nu^+)E^- + (1-c)(1+\nu^-)(1-2\nu^-)E^+}, \end{aligned} \tag{22}$$

where $E^*(=\sigma/e)$, $\nu^*(=-E_{rr}/e)$ are the effective stiffness and Poisson ratio, respectively. These results are the same as those obtained by Hashin and Rosen [5]. For Poisson contraction the limiting condition of the voided composite follows in similar fashion by substituting $f(u)=0$ into (21). However, for axial extension the condition of voided composite obviously does not follow by putting $f(u)=0$ in (9). When there is vanishing interface force (9) implies that axial response follows a rule of mixtures exactly.

## 6 Predictions

Predictions of effective axial tension response follow from (7), (9), (10), and (21) provided we specify the normal interface force separation relation $f$. For the remainder of this section we utilize the physically based model of Ferrante et al. [10] which allows for decohesion in normal mode. (Needleman has utilized generalizations of this model in numerous studies of debonding. See, for example, Needleman [11] for a numerical unit cell analysis of the plane strain problem.) Strictly, the interface model is derived for crystalline materials and the parameters arising in it have a precise physical meaning. Here we employ the model in a more general context and interpret the parameters as phenomenological, to be determined from experiment. In this work the interface force law will be taken in the form

$$f(u)=\exp(1)\sigma_{\mathrm{MAX}}(u/\rho)\exp(-u/\rho), \tag{23}$$

where $\rho$ is the characteristic force length ratio (characteristic force length normalized with respect to $R_1$) and $\sigma_{\mathrm{MAX}}$ is the interface strength. Note that the maximum value of $f$ is $\sigma_{\mathrm{MAX}}$, which occurs at $u=\rho$, and the interface traction $f$ vanishes asymptotically with increasing $u$ (Fig. 2). Furthermore, if (23) is substituted into (7) and (9), and the normalized displacement jump $u$ is written as $\rho\omega$, then it is clear that in order to ensure equivalent response of composite cylinders of varying sizes the ratios $c(=(R_1/R_2)^2)$, and $\rho$ must be held fixed. This unphysical result, which has been noted in Hashin [12] in the context of linearly separating interfaces, requires that the characteristic force length of a particular composite cylinder interface must vary in direct proportion to the fiber radius. Nevertheless from hereon we assume that this is true.

Owing to the fact that the axial stress of the composite RVE depends on the product of the interface force and the difference in Poisson ratio of matrix and fiber (see (9)) the effect of nonlinear interface response can be expected to be small. Clearly the greatest deviation from the rule of mixtures occurs when the difference in Poisson ratio $\nu^- - \nu^+$ is at its largest positive value. However, even when this occurs plots of response reveal only a negligible deviation from the rule of mixtures.

Graphs of effective Poisson contraction reveal a more significant interface effect. Figure 3 shows plots of radial strain versus axial strain at different values of fiber volume concentration for interface force length parameter $\rho=.001$. All other properties are fixed at the values

$$\nu^-=.45,\quad \nu^+=.25,\quad \frac{\sigma_{\mathrm{MAX}}}{E^+}=.002,\quad \frac{E^-}{E^+}=21. \tag{24}$$

Also shown is pure matrix response and pure fiber response. Consider the curve drawn at $c=.5$ in Fig. 3 and note that the fiber

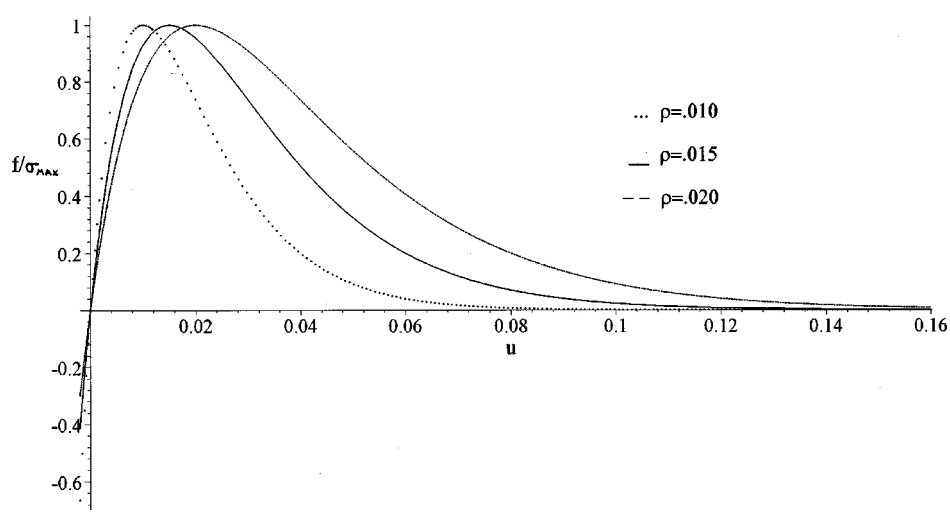



**Fig. 2 The normal interface force law**

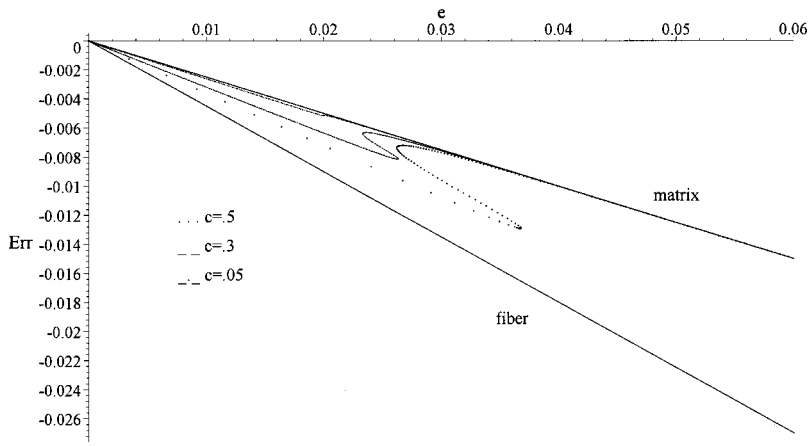



**Fig. 3 Effect of concentration on Poisson contraction response, $\rho = 0.001$**

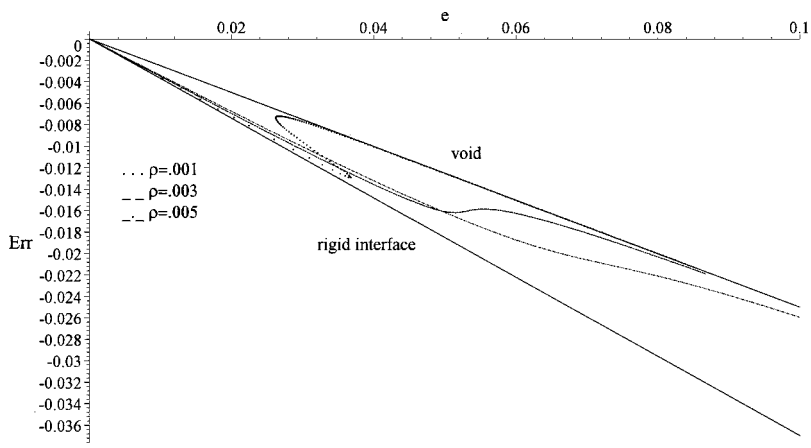



**Fig. 4 Effect of force length ratio on Poisson contraction response, $c = 0.5$**

Poisson ratio exceeds the matrix value. For small values of axial strain the radial contraction of the RVE exceeds that of the matrix value owing to the relatively large radial tensile tractions that develop on the interfaces. With increasing axial strain the fiber pulls away from the matrix. The interface tractions decrease and the overall radial contraction of the composite approaches that of the voided composite or matrix. The interface parameters chosen permit brittle decohesion and the process involves an abrupt transition to matrix behavior. Naturally the magnitude of the phenomenon decreases with decreasing fiber volume concentration.

Consider now a situation in which the matrix Poisson ratio is greater then the fiber value (graph not shown). For small values of axial strain the radial contraction of the RVE is smaller than the matrix value, owing to the relatively large radial compressive tractions that develop on the interfaces. Because the interface force law effectively eliminates fiber-matrix interpenetration we obtain behavior, which closely follows rigid interface response (discrepancies arise due to small amounts of interpenetration at the interfaces (see Fig. 2)). Clearly, as the fiber volume concentration decreases we approach matrix response.

Finally, Fig. 4 shows a graph of radial strain versus axial strain at different values of interface force length parameter $\rho$. The fiber volume concentration is assumed fixed at 0.5 and all other properties are as given in (24). In all cases a transition occurs from rigid interface response, on initial application of the load, to void behavior as the interface fully separates. The effect of decreasing $\rho$ on behavior is clearly demonstrated as the response changes from gradual contraction under increasing axial strain to an abrupt decrease in the magnitude of the radial contraction as the interface unloads.

## 7 Concluding Remarks

An analysis of the axial tension response of a composite has been carried out assuming that matrix and fiber interact through a normal, nonlinear interface force-displacement discontinuity relation. The particular homogenization method employed to pass from local response to aggregate response is the composite cylinders construct ([5]). Upper and lower bounds to the total potential energy of a composite RVE have been shown to coincide so an exact solution exists for the composite cylinders geometry. Overall, however, the effect of interface characterization on extensional response is virtually nil and essentially follows a ''rule of mixtures.'' The effect of the interfaces on Poisson contraction is somewhat greater with the nature of the behavior largely dependent on the sign of the difference in Poisson ratio of matrix and fiber.

Some features of Poisson contraction response are that increasing fiber volume concentration ($c$) (i) increases the magnitude of the lateral strain jump at brittle decohesion, (ii) delays the onset of brittle decohesion, and (iii) increases the initial Poisson contraction over pure matrix response. Additionally, decreasing the force length ratio ($\rho$) (i) precipitates a transition from ductile to brittle composite response, (ii) encourages the onset of brittle decohesion, and (iii) increases the initial Poisson contraction.

Note that these results (and analogous results for other loadings) have been demonstrated to hold only for a particular interface force law and specific parameter values. Furthermore, they are based solely on the composite cylinders geometry. Although the sensitivity of the overall behavior to variations in microstructure and microstructural arrangement is unknown, we should ex-

pect that the sensitivity of the local behavior to these factors to likely be high. Consider the fact that cavity formation by interfacial decohesion occurs in a nonuniform progressive manner throughout the composite and that this feature is not captured in the present model. Further development of these kinds of models in this regard is probably worth the effort given their ability to predict global damage behavior based solely on the constitutive characteristics of the constituents. Naturally such development must proceed alongside a vigorous experimental program.

## References


[1] Levy, A. J., 1996, ''The Effective Dilatational Response of Fiber Reinforced Composites With Nonlinear Interface,'' ASME J. Appl. Mech., **63**, pp. 357–364.

[2] Levy, A. J., and Dong, Z., 1998, ''Effective Transverse Response of Fiber Composites With Nonlinear Interface,'' J. Mech. Phys. Solids, **46**, pp. 1279–1300.

[3] Dong, Z., and Levy, A. J., 2000, ''Mean Field Estimates of the Response of Fiber Composites With Nonlinear Interface,'' Mech. Mater., to appear.

[4] Needleman, A., 1987, ''A Continuum Model for Void Nucleation by Inclusion Debonding,'' ASME J. Appl. Mech., **54**, pp. 525–531.

[5] Hashin, Z., and Rosen, B. W., 1964, ''The Elastic Moduli of Fiber Reinforced Materials,'' ASME J. Appl. Mech., **31**, pp. 223–232.

[6] Hashin, Z., 1990, ''Thermoelastic Properties of Fiber Composites With Imperfect Interface,'' Mech. Mater., **8**, pp. 333–348.

[7] Benveniste, Y., 1985, ''The Effective Mechanical Behavior of Composite Materials With Imperfect Contact Between the Constituents,'' Mech. Mater., **4**, pp. 197–208.

[8] Aboudi, J., 1991, *Mechanics of Composite Materials*, Elsevier, Amsterdam.

[9] Gurtin, M. E., 1984, ''The Linear Theory of Elasticity,'' *Mechanics of Solids*, Vol. 2, C. Truesdell, ed., Springer-Verlag, Heidelberg, pp. 1–273.

[10] Ferrante, J., Smith, J. R. and Rose, J. H., 1982, ''Universal Binding Energy Relations in Metallic Adhesion,'' *Microscopic Aspects of Adhesion and Lubrication*, J. M. Georges, ed., Elsevier, Amsterdam, pp. 19–30.

[11] Needleman, A., 1993, ''Void Nucleation by Inclusion Debonding in a Crystal Matrix,'' Modell. Simul. Mater. Sci. Eng., **1**, pp. 111–132.

[12] Hashin, Z., 1991, ''Thermoelastic Properties of Particulate Composites With Imperfect Interface,'' J. Mech. Phys. Solids, **39**, pp. 745–762.

# The Fiber Composite With Nonlinear Interface—Part II: Antiplane Shear

**A. J. Levy**
Department of Mechanical, Aerospace
And Manufacturing Engineering,
Syracuse University,
Syracuse, NY 13244-1240
Mem. ASME

*This paper treats the effective antiplane shear response of a composite consisting of fibers that interact with the matrix through nonlinear Needleman-type cohesive zones. The first paper (Part I) examines effective axial tension response. The composite cylinders representation of a representative volume element (RVE) is employed throughout. For antiplane shear loading the elastic field solution for a single composite cylinder is found in the form of a series expansion whose coefficients are governed by an infinite set of nonlinear equations. Bounds on the total potential energy and the total complementary energy of an RVE do not coincide although they are shown to differ by a term of order $O(c^4)$ where c is the fiber volume concentration. Interaction effects due to finite volume concentration, coupled with nonlinear interface characterization, are shown to precipitate instability in composite response.* [S0021-8936(00)02104-8]

## 1 Introduction

This paper (Part II) presents a study of the effective antiplane shear response of randomly arrayed fiber composites containing interfaces that can slip nonlinearly. The first related paper (Part I) examines axial tension response. In both parts we assume that fiber-matrix interface discontinuities are governed by uniform Needleman-type cohesive zones of vanishing thickness ([1]). In Part I the elastic fields of a single composite cylinder subject to axial tension were used along with the composite cylinders construct ([2]) and variational bounding to obtain an exact solution for the response of the representative volume element (RVE). In this paper we employ the same homogenization method assuming a potential interface constitutive model which allows for interface failure in shear mode.

The more complicated antiplane shear problem cannot be put in the simple framework of the axial tension problem (see (1) of Part I). This is because the slip discontinuity at the fiber matrix interface is a function of interface coordinate whereas the interface separation discontinuity under axial tension is uniform, independent of interface coordinate. It will be shown, however, that the constitutive relation for a single composite cylinder may be written in the form

$$\tau=F(\gamma,\chi^+,\chi^-,c,\omega_i),$$
$$G_i(\gamma,\chi^+,\chi^-,c,\omega_j,\xi_k)=0,\quad i=0,1,2,\ldots, \tag{1}$$

where $\tau$, $\gamma$ are corresponding mean effective shear stress and strain components; $\chi^+$, $\chi^-$ are elastic properties of matrix and fiber, respectively; $c$ is the volume concentration of fiber; and the $\xi_k$, $k=1,2,\ldots$ are interface constitutive parameters. Relation $(1_1)$ is the effective property relation for the composite cylinder. The set of internal variables $\omega_i$, $i=0,1,2,\ldots$ are the expansion coefficients of the interface slip discontinuity field and appear in $(1_1)$ through expressions for the mean strain in the fiber phase and the displacement discontinuity tensor.[1] Relations $(1_2)$ are determined from the interface boundary condition for the composite cylinder and depend explicitly on the interface constitutive relation prescribed in the form of a tangential interface force-slip discontinuity relation. Homogenization via the composite cylinders representation and variational bounding then follow in a manner similar to that in Part I.

In the next two sections we consider antiplane shear of a single composite cylinder and derive equivalent constitutive relations governing its response. The treatment extends the results obtained for antiplane shear response of rigid interface composites ([2]), and linearly separating interface composites ([3]). In particular, Section 2 is concerned with the displacement boundary value problem while Section 3 treats the traction boundary value problem. In Section 4 we use variational arguments to bound the response of an RVE consisting of a volume-filling configuration of composite cylinders of varying diameters. Unlike its rigid or linear interface counterparts the bounds do not coincide so an exact solution does not exist. The results obtained under uniform strain and uniform stress boundary conditions must therefore be interpreted as bounds on overall behavior. These bounds, however, are shown to differ by a term of order $O(c^4)$ so that the stress-strain response may be accurately predicted by uniform stress or strain boundary conditions provided $c$ is not very large. We conclude the paper with a brief discussion on the utility of modeling composite damage by homogenizing the response of a single composite cylinder.

## 2 The Displacement Boundary Value Problem

Hashin and Rosen [2] have solved the problem of antiplane shear of a single composite cylinder composed of a cylindrical fiber rigidly bonded to an annular cylinder of matrix material. Here we generalize that analysis to include nonlinear interface response.

Consider the composite cylinder shown in Fig. 1. Following Hashin and Rosen [2] we begin with Navier's equation for isochoric deformation ($\operatorname{div}\mathbf{u}=0$),

$$\Delta\mathbf{u}=\mathbf{0}, \tag{2}$$

subject to a uniform strain boundary condition on the outer cylindrical surface,

$$\mathbf{E}_H=\frac{\gamma}{2}(\mathbf{e}_1\otimes\mathbf{e}_3+\mathbf{e}_3\otimes\mathbf{e}_1),\quad \text{on } r=R_2. \tag{3}$$

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, May 7, 1999; final revision, June 25, 2000. Associate Technical Editor: I. M. Daniel. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

[1]The difference between the applied boundary strain and the mean strain.

 Copyright © 2000 by ASME 

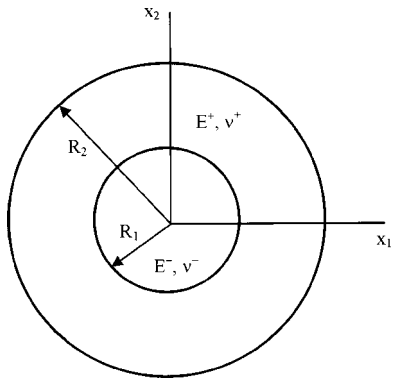



**Fig. 1 Composite cylinder geometry**

(The traction boundary value problem will be considered in the following section.) If the displacement due to the rigid rotation field

$$\mathbf{W}=\frac{\gamma}{2}(\mathbf{e}_1\otimes\mathbf{e}_3-\mathbf{e}_3\otimes\mathbf{e}_1) \tag{4}$$

is superimposed on the displacement resulting from (3), then the resulting boundary displacement may be written in polar coordinates in the form

$$\mathbf{u}=\gamma z\cos\theta\mathbf{e}_r-\gamma z\sin\theta\mathbf{e}_\theta,\quad \text{on } r=R_2. \tag{5}$$

We seek solutions to the problem (2) and (5) in the fiber domain $\Omega=\{(r,\theta):r\in(0,R_1),\theta\in(0,2\pi)\}$, and in the matrix domain $B-\Omega=\{(r,\theta):r\in(R_1,R_2),\theta\in(0,2\pi)\}$ for the case where the interface boundary condition is a prescribed nonlinear function of tangential interface displacement discontinuity. Thus, in addition to (5) prescribe

$$\begin{aligned}\mathbf{s}(\mathbf{e}_r;r,\theta)&=g(\theta)\mathbf{e}_z,\quad \text{on } r\uparrow R_1,\quad r\in\Omega\\ \mathbf{s}(-\mathbf{e}_r;r,\theta)&=-g(\theta)\mathbf{e}_z,\quad \text{on } r\downarrow R_1,\quad r\in B-\Omega\end{aligned} \tag{6}$$

where $g$ is taken to be an even function of $\theta$.[2] The problem is now completely posed, provided we insure bounded solutions by enforcing regularity at the origin

$$\text{Lim } \mathbf{u} \text{ bounded},\ r\downarrow 0. \tag{7}$$

Now assume a form of solution, valid in both domains,

$$\begin{aligned}u_r&=\sum_{n=0}^{\infty}d_n(z)\cos n\theta,\quad u_\theta=\sum_{n=0}^{\infty}h_n(z)\sin n\theta,\\ u_z&=\sum_{n=0}^{\infty}f_n(r)\cos n\theta.\end{aligned} \tag{8}$$

It can be shown (see the Appendix) that (8) will satisfy Navier's equation provided that the sets of functions $\{d_n\}$, $\{h_n\}$, $\{f_n\}$ are such that (8) may be expressed in the form

$$\begin{aligned}u_r&=(A_{11}+A_{12}z)\cos\theta,\\ u_\theta&=-(A_{11}+A_{12}z)\sin\theta,\\ u_z&=C_{10}+C_{20}\log r+\sum_{n=0}^{\infty}(C_{1n}r^n+C_{2n}r^{-n})\cos n\theta,\end{aligned} \tag{9}$$

where $A_{11}$, $A_{12}$, $C_{10}$, $C_{20}$, $\{C_{1n}\}$, $\{C_{2n}\}$ are constants. It remains to determine the constants for the fiber and matrix domains from the boundary and regularity conditions.

[2]Ultimately, $g$ will depend on $u_z$, which is an even function of $\theta$.

For the fiber domain $\Omega$ the boundary conditions ($6_1$) and (7) fix the constants $C_{20}$, $\{C_{1n}\}$, $\{C_{2n}\}$ so that the displacement field is given by

$$\begin{aligned}u_r^-&=(A_{11}+A_{12}z)\cos\theta,\\ u_\theta^-&=-(A_{11}+A_{12}z)\sin\theta,\\ u_z^-&=C_{10}+\frac{R_1}{\mu^-}\sum_{n=1}^{\infty}\frac{1}{n}g_n\left(\frac{r}{R_1}\right)^n\cos n\theta,\end{aligned} \tag{10}$$

where $\mu^-$ is the shear modulus of the fibers and $g_n$ are the mode multipliers in the expansion for $g$, i.e.,

$$\begin{aligned}g(\theta)&=\sum_{n=0}^{\infty}g_n\cos n\theta,\quad g_0=\frac{1}{2\pi}\int_0^{2\pi}g(\theta)d\theta,\\ g_n&=\frac{1}{\pi}\int_0^{2\pi}g(\theta)\cos n\theta d\theta.\end{aligned} \tag{11}$$

(Note that the constant term $g_0$ in the expansion is required to vanish in order to insure overall global equilibrium of the fiber.) The remaining constants appearing in (10), i.e., $A_{11}, A_{12}, C_{10}$, can be identified with the rigid-body displacement,

$$\mathbf{u}_R=\mathbf{u}_0+\mathbf{W}\mathbf{r},\quad \mathbf{u}_0=u_{01}\mathbf{e}_1+u_{03}\mathbf{e}_3,\quad \mathbf{W}=\omega(\mathbf{e}_1\otimes\mathbf{e}_3-\mathbf{e}_3\otimes\mathbf{e}_1), \tag{12}$$

so that the elastic field in the fiber may be written as

$$\begin{aligned}u_r^-&=(u_{01}+\omega z)\cos\theta,\\ u_\theta^-&=-(u_{01}+\omega z)\sin\theta,\\ u_z^-&=u_{03}^--\omega r\cos\theta+\frac{R_1}{\mu^-}\sum_{n=1}^{\infty}\frac{1}{n}g_n\left(\frac{r}{R_1}\right)^n\cos n\theta,\\ S_{rz}^-&=\sum_{n=1}^{\infty}g_n\left(\frac{r}{R_1}\right)^{n-1}\cos n\theta,\\ S_{\theta z}^-&=-\sum_{n=1}^{\infty}g_n\left(\frac{r}{R_1}\right)^{n-1}\sin n\theta.\end{aligned} \tag{13}$$

For the matrix domain $B$-$\Omega$ the boundary conditions (5) and ($6_2$) determine the constants appearing in (9) so that the elastic field in the matrix is

$$\begin{aligned}u_r^+&=\gamma z\cos\theta,\\ u_\theta^+&=-\gamma z\sin\theta,\\ u_z^+&=\left(\frac{-\gamma c}{1+c}\right)\left(r-\frac{R_2^2}{r}\right)\cos\theta\\ &\quad+\frac{R_2}{\mu^+}\sum_{n=1}^{\infty}\frac{c^{(n+1)/2}}{n(1+c^n)}g_n\left[\left(\frac{r}{R_2}\right)^n-\left(\frac{R_2}{r}\right)^n\right]\cos n\theta,\\ S_{rz}^+&=\mu^+\gamma\cos\theta+\sum_{n=1}^{\infty}\frac{c^{(n+1)/2}}{(1+c^n)}g_n\left[\left(\frac{r}{R_2}\right)^{n-1}+\left(\frac{R_2}{r}\right)^{n+1}\right]\cos n\theta,\\ S_{\theta z}^+&=-\mu^+\gamma\sin\theta-\sum_{n=1}^{\infty}\frac{c^{(n+1)/2}}{(1+c^n)}g_n\left[\left(\frac{r}{R_2}\right)^{n-1}-\left(\frac{R_2}{r}\right)^{n+1}\right]\sin n\theta,\end{aligned} \tag{14}$$

where $\mu^+$ is matrix shear modulus and $c$ is the fiber volume concentration defined to be $(R_1/R_2)^2$. Now $g$ is the tangential interface force, which is generally a function of normalized tangential displacement discontinuity at the fiber-matrix interface. Thus, $g(w(\theta))$[3] where $w$ is defined by

[3]We assume the interface is uniform, i.e., $g$ is dependent on coordinate $\theta$ implicitly only through $w$.

$$w=[\mathbf{u}]\cdot\mathbf{e}_z/R_1=(u_z^+-u_z^-)/R_1. \tag{15}$$

Radial and angular displacement jump components $[\mathbf{u}]\cdot\mathbf{e}_r/R_1=(u_r^+-u_r^-)/R_1$, and $[\mathbf{u}]\cdot\mathbf{e}_\theta/R_1=(u_\theta^+-u_\theta^-)/R_1$ may be assumed to vanish implying that $u_{01}=0$ and $\omega=\gamma$ by (13) and (14).

Because $w$ is unknown we need to solve for it from an interfacial integral equation. This can be obtained by subtracting $(13_3)$ from $(14_3)$, equating the result to (15), and noting the expansion (11). Thus,

$$w(\theta)=w_0+\frac{2}{1+c}\gamma\cos\theta-\int_0^{2\pi}K_\gamma(\theta-\theta')g(w(\theta'))d\theta',$$
$$K_\gamma(\theta-\theta')=\frac{1}{\pi}\sum_{n=1}^{\infty}\left(\frac{1-c^n}{1+c^n}\frac{1}{\mu^+}+\frac{1}{\mu^-}\right)\frac{1}{n}\cos n(\theta-\theta'). \tag{16}$$

In (16) the constant term in the expansion for $w$, i.e., the uniform slip mode $w_0$, is indeterminate. A determinate set of equations governing $w$ is obtained by coupling to (16) the integral equation resulting from rigid-body force equilibrium

$$\frac{1}{2\pi}\int_0^{2\pi}g(w(\theta))d\theta=0. \tag{17}$$

The solution $w$, to the nonlinear integral Eqs. (16) and (17), is obtained in the form of a Fourier expansion

$$w(\theta)=w_0+\sum_{n=1}^{\infty}w_{2n-1}\cos n\theta. \tag{18}$$

Expansion (18), together with orthogonality of eigenfunctions, reduces the governing integral equations to an infinite set of nonlinear algebraic equations[4]

$$\frac{1}{2\pi}\int_0^{2\pi}g(w_i;\theta)d\theta=0,$$
$$w_1+\frac{1}{\pi}\left(\frac{1-c}{1+c}\frac{1}{\mu^+}+\frac{1}{\mu^-}\right)\int_0^{2\pi}g(w_i;\theta)\cos\theta d\theta=\frac{2}{1+c}\gamma,$$
$$w_{2n-1}+\frac{1}{n\pi}\left(\frac{1-c^n}{1+c^n}\frac{1}{\mu^+}+\frac{1}{\mu^-}\right)\int_0^{2\pi}g(w_i;\theta)\cos n\theta d\theta=0,$$
$$n=2,3,\ldots. \tag{19}$$

Throughout the remainder of the paper we assume that interfacial slip $w$ is antisymmetric with respect to the $\mathbf{e}_2$-$\mathbf{e}_3$ plane so that $w$ is of the form

$$w(\theta)=\sum_{n=1}^{\infty}w_{4n-3}\cos(2n-1)\theta. \tag{20}$$

It is a straightforward matter to show that solution (20) satisfies (19) with $w_0=w_{4n-1}=0$, $n=1,2,\ldots$ Thus, we work with the reduced set

$$w_1+\frac{1}{\pi}\left(\frac{1-c}{1+c}\frac{1}{\mu^+}+\frac{1}{\mu^-}\right)\int_0^{2\pi}g(w_i;\theta)\cos\theta d\theta=\frac{2}{1+c}\gamma,$$
$$w_{4n-3}+\frac{1}{(2n-1)\pi}\left(\frac{1-c^{2n-1}}{1+c^{2n-1}}\frac{1}{\mu^+}+\frac{1}{\mu^-}\right)$$
$$\times\int_0^{2\pi}g(w_i;\theta)\cos(2n-1)\theta d\theta=0,\quad n=2,3,\ldots. \tag{21}$$

With the prescription of $g$ the solution to (21) can proceed and the complete elastic field follows from (11), (13), and (14).

[4]This procedure has been utilized by the author in a different context, most recently in Dong and Levy [4].

An equivalent antiplane constitutive relation for the composite cylinder may be obtained from the $\mathbf{e}_1\otimes\mathbf{e}_3$ component of the effective stiffness relation ([5,6])

$$\tau=\mu^+\gamma+2c(\mu^--\mu^+)\frac{S_{13}^-}{2\mu^-}-2\mu^+A_{13}, \tag{22}$$

where $\tau$ is the mean shear stress on the lateral surface of the cylinder, $\gamma$ is twice the applied shear strain, $S_{13}^-$ is the Cartesian component of the mean stress in the fiber phase (obtained from (13) or the mean stress theorem), and $A_{13}$ is the $\mathbf{e}_1\otimes\mathbf{e}_3$ component of the displacement discontinuity tensor ($\mathbf{A}=c/(2\,\mathrm{vol}(\Omega))\int_{\partial\Omega}([\mathbf{u}]\otimes\mathbf{n}+\mathbf{n}\otimes[\mathbf{u}])da$) given by

$$A_{13}=\frac{c}{2\pi}\int_0^{2\pi}w\cos\theta d\theta. \tag{23}$$

Equations (22), (23), and (19) yield the equivalent antiplane shear constitutive relation for the composite cylinder (subject to a displacement boundary condition),

$$\tau=\mu^+\frac{1-c}{1+c}\gamma+\frac{2c}{1+c}\frac{1}{\pi}\int_0^{2\pi}g(w_i;\theta)\cos\theta d\theta. \tag{24}$$

## 3 The Traction Boundary Value Problem

Assume now that the composite cylinder of Fig. 1 is subject to the traction boundary condition

$$\mathbf{s}(\mathbf{e}_r)=\tau\cos\theta\mathbf{e}_z,\quad \text{on } r=R_2,$$

and, additionally, is subject to the interface boundary condition (6) and the regularity condition (7). An analysis of this problem, similar to the one given above for the displacement problem, yields the elastic field in the matrix:

$$u_r^+=(u_{01}+\omega z)\cos\theta,$$
$$u_\theta^+=-(u_{01}+\omega z)\sin\theta,$$
$$u_z^+=u_{03}^+-\omega r\cos\theta+\frac{1}{1-c}\left[1+c\left(\frac{R_2}{r}\right)^2\right]\frac{\tau}{\mu^+}r\cos\theta$$
$$-\frac{R_2}{\mu^+}\sum_{n=1}^{\infty}\frac{c^{(n+1)/2}}{n(1-c^n)}g_n\left[\left(\frac{r}{R_2}\right)^n-\left(\frac{R_2}{r}\right)^n\right]\cos n\theta,$$
$$S_{rz}^+=\frac{1}{1-c}\left[1-c\left(\frac{R_2}{r}\right)^2\right]\tau\cos\theta$$
$$-\sum_{n=1}^{\infty}\frac{c^{(n+1)/2}}{(1-c^n)}g_n\left[\left(\frac{r}{R_2}\right)^{n-1}-\left(\frac{R_2}{r}\right)^{n+1}\right]\cos n\theta,$$
$$S_{\theta z}^+=\frac{-1}{1-c}\left[1+c\left(\frac{R_2}{r}\right)^2\right]\tau\sin\theta$$
$$+\sum_{n=1}^{\infty}\frac{c^{(n+1)/2}}{(1-c^n)}g_n\left[\left(\frac{r}{R_2}\right)^{n-1}+\left(\frac{R_2}{r}\right)^{n+1}\right]\sin n\theta, \tag{25}$$

where $u_{01},\omega,u_{03}^+$ represent a rigid-body displacement (12). The elastic field in the fiber is identical to that of the displacement problem and is given by (13). The integral equations governing the tangential displacement discontinuity (15) are obtained from (13) and (25),

$$w(\theta)=w_0+\frac{2}{1-c}\frac{\tau}{\mu^+}\cos\theta-\int_0^{2\pi}K_\tau(\theta-\theta')g(w(\theta'))d\theta',$$
$$K_\tau(\theta-\theta')=\frac{1}{\pi}\sum_{n=1}^{\infty}\left(\frac{1+c^n}{1-c^n}\frac{1}{\mu^+}+\frac{1}{\mu^-}\right)\frac{1}{n}\cos n(\theta-\theta'), \tag{26}$$

where $w_0$ is determined from (17) which ensures rigid-body force equilibrium. An expansion of the form (18) reduces (26) to the set

$$\frac{1}{2\pi}\int_0^{2\pi} g(w_i;\theta)d\theta=0,$$

$$w_1+\frac{1}{\pi}\left(\frac{1+c}{1-c}\frac{1}{\mu^+}+\frac{1}{\mu^-}\right)\int_0^{2\pi} g(w_i;\theta)\cos\theta\, d\theta=\frac{2}{1-c}\frac{\tau}{\mu^+}, \tag{27}$$

$$w_{2n-1}+\frac{1}{n\pi}\left(\frac{1+c^n}{1-c^n}\frac{1}{\mu^+}+\frac{1}{\mu^-}\right)\int_0^{2\pi} g(w_i;\theta)\cos n\theta d\theta=0,$$

$$n=2,3,\ldots,$$

which we reduce further by assuming $w$ in the form (20), i.e., consider the reduced set consisting of $(27_3)$ with $n$ odd.

For the traction boundary value problem an equivalent constitutive relation for the composite cylinder may be obtained from the $\mathbf{e}_1\otimes\mathbf{e}_3$ component of the effective compliance relation ([5,6])

$$\gamma=\frac{\tau}{\mu^+}+c\left(\frac{1}{\mu^-}-\frac{1}{\mu^+}\right)S_{13}^-+A_{13}, \tag{28}$$

with $A_{13}$ in the form (23). Combining (13), (23), (27), and (28) the equivalent constitutive for the composite may be written as

$$\gamma=\frac{1+c}{1-c}\frac{\tau}{\mu^+}-\frac{2c}{\mu^+(1-c)}\frac{1}{\pi}\int_0^{2\pi} g(w_i;\theta)\cos\theta d\theta. \tag{29}$$

## 4 Variational Bounding

In this section we derive bounds on the effective response of an RVE and show that in general (27), (29), of the traction boundary value problem, does not predict the same response as (19) and (24) of the displacement boundary value problem. Throughout we employ the volume filling composite cylinders representation of an RVE introduced by Hashin and Rosen [2]. That is, a distribution of geometrically similar composite cylinders of varying sizes down to the vanishingly small such that all composite cylinders have identical mean stress and strain fields (see Section 6 for a precise statement of the constraints that insure this).

First, assume that the displacement field in a composite cylinder with displacement vector $\mathbf{u}=\gamma z\cos\theta\mathbf{e}_r-\gamma z\sin\theta\mathbf{e}_\theta$ prescribed on the surface $r=R_2$ forms a kinematically admissible displacement field[5] for the displacements in the entire RVE of the composite. Now the potential energy $\Phi\{\gamma\}$ stored in a single composite cylinder subject to the displacement condition can be calculated from (13), (14), and (21) and is given by

$$\Phi\{\gamma\}=\frac{c}{2}\left\{\sum_{n=2}^{\infty}\frac{(2n-1)\mu^+\mu^-(1+c^{2n-1})}{(1-c^{2n-1})\mu^-+(1+c^{2n-1})\mu^+}(w_{4n-3}^{\gamma})^2\right.$$
$$+\frac{1-c}{c(1+c)}\mu^+\gamma^2+\frac{\mu^+\mu^-(1+c)}{(1-c)\mu^-+(1+c)\mu^+}$$
$$\left.\times\left[\frac{2\gamma}{(1+c)}-w_1^{\gamma}\right]^2+\frac{2}{\pi}\int_0^{2\pi}\varepsilon_I d\theta\right\}v(B), \tag{30}$$

where $\varepsilon_I$ is the interface potential defined by

$$\varepsilon_I=\int^{w} g dz, \tag{31}$$

$v(B)$ is the volume of the composite cylinder and a superscript $\gamma$ indicates a quantity obtained from displacement boundary conditions. (Note that the vanishing of the derivatives $D_{w_i}\Phi$ imply relations (21) while $\tau=D_\gamma[v^{-1}(B)\Phi\{\gamma\}]$ and $(21_1)$ imply relation (24). An upper bound to the potential energy for the RVE of the composite, based on the admissible displacement field, is $\Phi\{\gamma\}v/v(B)$ where $v$ is the volume of the RVE.

Next, assume that the homogeneous stress field, produced by the traction boundary condition $\mathbf{s}(\mathbf{e}_r)=\tau\cos\theta\mathbf{e}_z$ applied to the boundary $r=R_2$ of the composite cylinder, is a statically admissible stress field[6] for the entire RVE. Then the complementary energy for a single composite cylinder is given by

$$\Psi\{\tau\}=\frac{c}{2}\left\{\sum_{n=2}^{\infty}\frac{(2n-1)\mu^+\mu^-(1-c^{2n-1})}{(1+c^{2n-1})\mu^-+(1-c^{2n-1})\mu^+}(w_{4n-3}^{\tau})^2\right.$$
$$+\frac{1+c}{c(1-c)}\frac{\tau^2}{\mu^+}+\frac{\mu^+\mu^-(1-c)}{(1+c)\mu^-+(1-c)\mu^+}$$
$$\left.\times\left[\frac{-4\tau^2}{\mu^{+2}}\frac{1}{(1-c)^2}+(w_1^{\tau})^2\right]+\frac{2}{\pi}\int_0^{2\pi}\varepsilon_{IC}d\theta\right\}v(B), \tag{32}$$

where a superscript $\tau$ indicates a quantity obtained from a traction boundary condition and use has been made of (13), (25), and the reduced form of (27). Note that $\varepsilon_{IC}$ is defined by

$$\varepsilon_{IC}=\int^{g} w dz. \tag{33}$$

(The vanishing of the derivatives $D_{g_i}\Psi$ imply the reduced form of relations (27), while $\gamma=D_\tau[v^{-1}(B)\Psi(\tau)]$ yields (29). An upper bound to the complementary energy for the RVE of the composite, based on the statically admissible stress field, is then given by $\Psi\{\tau\}v/v(B)$. The potential (complementary) energy given by (30) ((32)) may be regarded as the strain (stress) energy for an equivalent homogeneous cylinder owing to the displacement (traction) condition applied to the boundary. Upper and lower bounds on the true strain energy ($U$) of the displacement problem follow from the bounding theorem ([7]):

$$\int_{\partial B}\overline{\mathbf{s}}\cdot\hat{\mathbf{u}}da-U_K\{\overline{\mathbf{S}}\}\leqslant U\leqslant U_C\{\tilde{\mathbf{E}}\}, \tag{34}$$

where $\tilde{\mathbf{E}}$ is a kinematically admissible strain field, $\overline{\mathbf{S}}$ is a statically admissible stress field with $\overline{\mathbf{s}}$ the corresponding surface traction, and $U_K\{\overline{\mathbf{S}}\}(U_C\{\tilde{\mathbf{E}}\})$ the stress (strain) energy. For the present problem (34) may be written as

$$\gamma^{\gamma}\tau^{\tau}v-\Psi\{\tau^{\tau}\}v/v(B)\leqslant U\leqslant\Phi\{\gamma^{\gamma}\}v/v(B), \tag{35}$$

where $\Phi$, $\Psi$ are given by (30), (32).

Upper and lower bounds of $U$ in (35) can be compared by first noting that $\varepsilon_I$ of (31) and $\varepsilon_{IC}$ of (33) are related according to

$$\int_0^{2\pi}\varepsilon_{IC}d\theta=\pi\sum_{i=1}^{\infty}w_{4i-3}g_{2i-1}-\int_0^{2\pi}\varepsilon_1 d\theta, \tag{36}$$

where eigenfunction expansions for $g$ ((11)) and $w$ ((20)) have been employed along with orthogonality of eigenfunctions. Now assume $\gamma^{\tau}=\gamma^{\gamma}=\gamma$. Substitute $(27_2)$ into (29) to eliminate $g_1$ (of (11)), and substitute the result into (35) to eliminate $\tau^{\tau}$. If we utilize (36) then the difference between the upper bound $U_+$ and the lower bound $U_-$ may be written as

[5]One that satisfies the displacement boundary condition and is twice continuously differentiable.

[6]One that satisfies the equilibrium equation and traction boundary condition.

$$U_+ - U_- = -\frac{2c\mu^+\mu^-[w_1^\gamma - w_1^\tau]\gamma}{(1-c)\mu^- + (1+c)\mu^+} + \frac{1}{2}\frac{\mu^+\mu^- c(1+c)[(w_1^\gamma)^2 - (w_1^\tau)^2]}{(1-c)\mu^- + (1+c)\mu^+} + \frac{c}{2}\sum_{n=2}^{\infty} \times \frac{(2n-1)\mu^+\mu^-(1-c^{2n-1})[(w_{4n-3}^\gamma)^2 - (w_{4n-3}^\tau)^2]}{(1+c^{2n-1})\mu^- + (1-c^{2n-1})\mu^+} - 2c\sum_{n=2}^{\infty}\mu^+\mu^- \frac{(2n-1)c^{2n-1}(w_{4n-3}^\gamma)^2}{(\mu^- - \mu^+)^2 c^{2(2n-1)} - (\mu^- + \mu^+)^2}. \tag{37}$$

Furthermore, it is not hard to show that (21), the reduced form of (27), and (29) imply that

$$(w_1^\tau - w_1^\gamma) + \left(\frac{(1-c)\mu^- + (1+c)\mu^+}{\pi\mu^+\mu^-(1+c)}\right) \times \int_0^{2\pi}[g_\tau(w_i^\tau;\theta) - g_\gamma(w_i^\gamma;\theta)]\cos\theta\, d\theta = 0,$$

$$(w_{4n-3}^\tau - w_{4n-3}^\gamma) + \left(\frac{(1-c^{2n-1})\mu^- + (1+c^{2n-1})\mu^+}{\pi(2n-1)\mu^+\mu^-(1+c^{2n-1})}\right) \times \int_0^{2\pi}[g_\tau(w_i^\tau;\theta) - g_\gamma(w_i^\gamma;\theta)]\cos(2n-1)\theta\, d\theta + \frac{4c^{2n-1}}{\pi(2n-1)\mu^+(1+c^{2n-1})(1-c^{2n-1})}\int_0^{2\pi} g_\tau(w_i^\tau;\theta) \times \cos(2n-1)\theta\, d\theta = 0, \quad n = 2,3,\ldots. \tag{38}$$

Clearly if $c=0$ in (38), corresponding to a single fiber embedded in an unbounded matrix, then $w_1^\tau = w_1^\gamma$, $w_{4n-3}^\tau = w_{4n-3}^\gamma$. For $c \neq 0$ the last term on the left-hand side of ($38_2$) is $O(c^{2n-1})$ and the implicit function theorem[7] implies that $w_{4n-3}^\tau = w_{4n-3}^\gamma + O(c^3)$ for all $n=1,2,\ldots$. Thus, from (37) the difference between the upper and lower bounds is a term of $O(c^4)$.

Note that for both rigid interface behavior and void behavior the upper and lower bounds coincide ([2]). These well-known results may be obtained as follows. For rigid interface behavior directly substitute $w_{4n-3}^\tau = w_{4n-3}^\gamma = 0$, $n=1,2,\ldots$ into (37). For void behavior, $g_{2n-1}^\tau = g_{2n-1}^\gamma = 0$, $n=1,2,\ldots$ so that by (21) $w_{4n-3}^\gamma = 0$, $n=2,3,\ldots$ and by (38) $w_{4n-3}^\tau = w_{4n-3}^\gamma$, $n=1,2,\ldots$. Then it follows from (37) that the bounds coincide.

[7]This argument fails near bifurcation points of (21) and the reduced form of (27).

## 5 Limiting Cases

The equations governing the effective response of the limiting cases of rigid interface behavior ($w_i=0$, $i=1,2,\ldots$) and void behavior ($g_j=0$, $j=1,2,\ldots$) can be recovered directly from (21) and (24) or the reduced form of (27) and (29). Thus, for void response $g_i=0$, $i=1,2,\ldots$ so that by (24)

$$\tau = \mu^+\frac{1-c}{1+c}\gamma. \tag{39}$$

Rigid interface response is obtained by substituting $g_1 = 2\gamma\mu^-\mu^+/[(1-c)\mu^- + (1+c)\mu^+]$ from (21) into (24). The result is

$$\tau = \left[\frac{(1+c)\mu^- + (1-c)\mu^+}{(1-c)\mu^- + (1+c)\mu^+}\right]\mu^+\gamma. \tag{40}$$

Both results (39) and (40) are the same as that of Hashin and Rosen [2].

## 6 Predictions

In this paper we employ an interface constitutive relation, proposed by Needleman [8], which allows for shear failure of the interface. The two-parameter potential model is of the form

$$g(w) = e^{1/2}\tau_{\mathrm{MAX}}\left(\frac{w}{\rho}\right)e^{-1/2(w/\rho)^2}, \tag{41}$$

where $\rho$ is the force length ratio (force length normalized with respect to fiber radius $R_1$) and $\tau_{\mathrm{MAX}}$ is the interfacial shear strength. Note that $\tau_{\mathrm{MAX}}$ is attained when the (normalized) tangential interface displacement discontinuity $w$ is equal to $\rho$. A graph of (41) for various values of parameter $\rho$ is shown in Fig. 2. Note that in order to satisfy the requirement of identical mean stress and strain in each composite cylinder composing the RVE the ratios $c(=(R_1/R_2)^2)$ and $\rho$ must be held fixed. This can be seen by introducing the variables $w_i = \rho\omega_i$ into (21), (24), and (41).

Equation (41) together with the governing Eqs. (21), (24) for displacement boundary conditions, or the reduced form of (27), (29) for traction boundary conditions, may be solved to obtain the bounds on shear stress-shear strain response. In this work a computer program, employing the Newton-Raphson method together with the composite Simpson 1/3 Rule, was written to integrate the equations numerically. Computations were carried out for an $n$ value of 9, which corresponds to nine shear modes ($\cos(2j-1)\theta$, $j=1,2,\ldots,9$). Figure 3 is a graph depicting these bounds for parameter values,

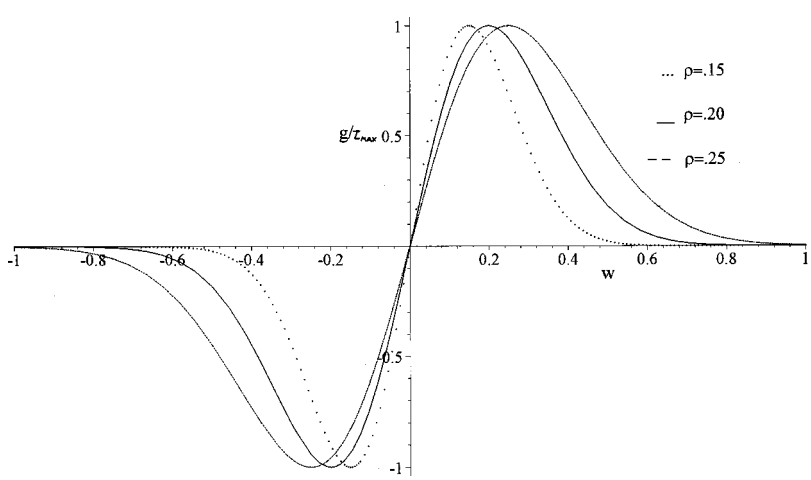



**Fig. 2 The tangential interface force law**

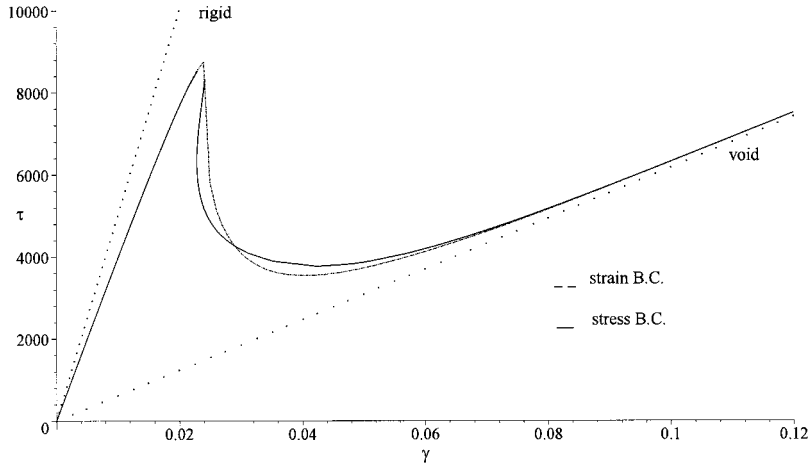



**Fig. 3 Bounds on antiplane shear stress-strain response, $c=0.5$, $\rho=0.01$**

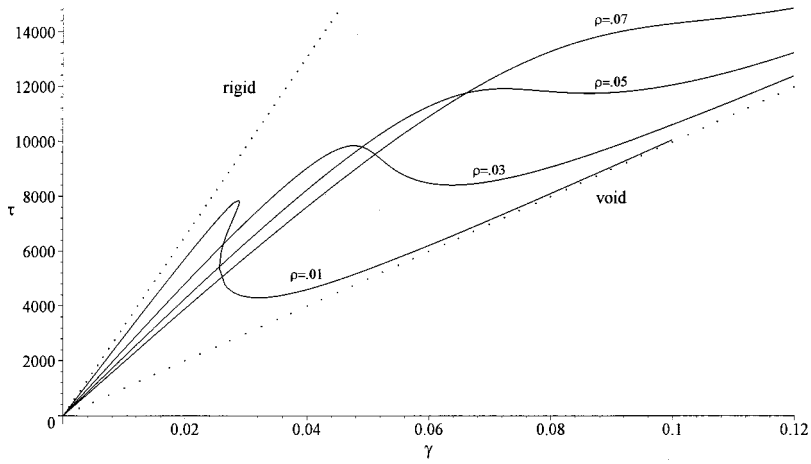



**Fig. 4 Effect of force length ratio on antiplane shear stress-strain response, $c=0.3$**

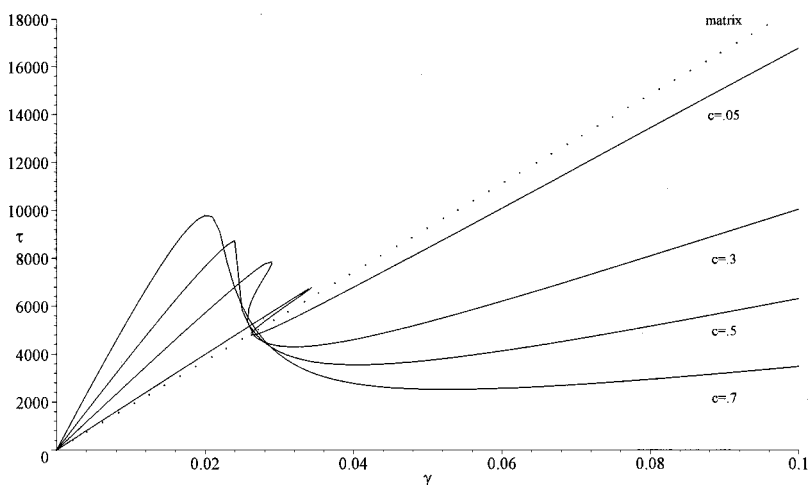



**Fig. 5 Effect of concentration on antiplane shear stress-strain response, $\rho=0.01$**

$$\mu^{-}=4375000\ \text{psi}, \quad \mu^{+}=185185\ \text{psi}, \quad \tau_{\text{MAX}}=10{,}000\ \ \text{psi},$$

which are taken to hold for the remainder of the paper. In addition, in Fig. 3, the fiber volume concentration ($c$) is assumed to be 0.5 and the force length ratio ($\rho$) is taken as 0.01. At these parameter values there is an abrupt transition in response under increasing stress corresponding to shear interface failure. The difference in the bounds at this value of $c$ is inconsequential since it is generally small and it occurs on a branch that is unreachable by a continuous increase in stress from the reference state. At smaller values of $c$ the difference between the bounds decreases further and at values of $c$ corresponding to the dilute estimate the curves

essentially coincide. From here on we study composite response obtained from governing equations for displacement boundary conditions, i.e., (21) and (24).

Figure 4 is a graph of antiplane shear response at various values of force length parameter $\rho$. The fiber volume concentration is assumed fixed at $c=0.3$. All curves undergo a transition from bonded interface response at initial application of load to void response corresponding to complete decohesion of the interface. Under decreasing $\rho$ this transition changes from gradual to abrupt. Curves similar to those in Fig. 4 have been obtained for transverse response of fiber composites where the operative mechanism is normal separation under transverse equibiaxial load ([9]), normal separation under transverse shear load ([10]), or normal separation coupled to interfacial slip under transverse shear load ([4]). This is not surprising since the interface models used in these studies are derivable from a potential and have ascending branches and descending branches that vanish asymptotically at infinity.

Figure 5 is a graph of antiplane shear response at various values of fiber volume concentration ($c$) for fixed force length parameter $\rho=0.01$. The effect of increasing $c$ is to (i) increase the initial stiffness of composite response, (ii) elevate the shear strength of the composite, (iii) increase the magnitude of the jump in shear strain at that value of stress equal to the shear strength.

Finally, although we have not considered solution convergence explicitly in this paper it has been considered in Levy [11] for the related problem of transverse response of a solitary fiber which debonds from the matrix according to a purely normal exponential separation law. In that work it was demonstrated that the solution converges quite rapidly and, in fact, the essential qualitative features of behavior under equibiaxial load are captured in only a two-term approximation.

## 7 Concluding Remarks

An analysis of the effective antiplane shear response of a composite has been carried out assuming that matrix and fiber interact through a tangential, nonlinear interface force-displacement discontinuity relation. The particular homogenization method employed to pass from local response to aggregate response is the composite cylinders construct ([2]). For antiplane shear, bounds on the total potential energy of the composite have been computed and they do not coincide although they have been shown to differ by a term of order $O(c^4)$.

For both antiplane shear and axial tension the effect of finite volume concentration on response is indicated in Fig. 5 of this paper and Fig. 3 of Part I ([12]), respectively. The relationship of the predicted behavior to physical reality is somewhat murky and underlines the inherent difficulties in modeling local failure mechanisms based on homogenizing the results for a single composite cylinder. Clearly, the stochastic, progressive nature of fiber debonding coupled to the sensitivity of the process to local perturbations from the assumed symmetry indicate the need for further refinement of the model presented here and in Part I. Another promising approach is numerical simulation based on the exact equations governing an N-particle system. Recently Sangani and Mo [13] have carried out an analysis of this type for linearly separating interfaces. For nonlinear interfaces, this approach is not without difficulties owing in part to bifurcation and stability issues associated with multiple equilibrium solutions.

## Appendix

We seek a general solution to $\Delta\mathbf{u}=\mathbf{0}$ in an annular or solid cylindrical domain. In polar coordinates the governing equations are

$$\Delta\mathbf{u}\cdot\mathbf{e}_z=\frac{\partial^2 u_z}{\partial r^2}+\frac{1}{r}\frac{\partial u_z}{\partial r}+\frac{1}{r^2}\frac{\partial^2 u_z}{\partial\theta^2}+\frac{\partial^2 u_z}{\partial z^2}=0,$$

$$\Delta\mathbf{u}\cdot\mathbf{e}_r=\frac{\partial^2 u_r}{\partial r^2}+\frac{1}{r}\frac{\partial u_r}{\partial r}+\frac{1}{r^2}\frac{\partial^2 u_r}{\partial\theta^2}+\frac{\partial^2 u_r}{\partial z^2}-\frac{2}{r^2}\frac{\partial u_\theta}{\partial\theta}-\frac{1}{r^2}u_r=0, \tag{A1}$$

$$\Delta\mathbf{u}\cdot\mathbf{e}_\theta=\frac{\partial^2 u_\theta}{\partial r^2}+\frac{1}{r}\frac{\partial u_\theta}{\partial r}+\frac{1}{r^2}\frac{\partial^2 u_\theta}{\partial\theta^2}+\frac{\partial^2 u_\theta}{\partial z^2}+\frac{2}{r^2}\frac{\partial u_r}{\partial\theta}-\frac{1}{r^2}u_\theta=0.$$

Assume a solution in the form

$$u_z=\sum_{n=0}^{\infty} f_n(r)\cos n\theta,\quad u_r=\sum_{n=0}^{\infty} d_n(z)\cos n\theta,$$
$$u_\theta=\sum_{n=0}^{\infty} h_n(z)\sin n\theta. \tag{A2}$$

Equation ($A1_1$) implies that

$$u_z=C_{10}+C_{20}\log r+\sum_{n=1}^{\infty}(C_{1n}r^n+C_{2n}r^{-n})\cos n\theta, \tag{A3}$$

where $C_{1n}, C_{2n}$ are constants. Equations ($A1_2$), ($A1_3$) imply that

$$u_r=(A_{11}+A_{12}z)\cos\theta,$$
$$u_\theta=-(A_{11}+A_{12}z)\sin\theta. \tag{A4}$$

## References


[1] Needleman, A., 1987, ''A Continuum Model for Void Nucleation by Inclusion Debonding,'' ASME J. Appl. Mech., **54**, pp. 525–531.
[2] Hashin, Z., and Rosen, B. W., 1964, ''The Elastic Moduli of Fiber Reinforced Materials,'' ASME J. Appl. Mech., **31**, pp. 223–232.
[3] Hashin, Z., 1990, ''Thermoelastic Properties of Fiber Composites With Imperfect Interface,'' Mech. Mater., **8**, pp. 333–348.
[4] Dong, Z., and Levy, A. J., 2000, ''Mean Field Estimates of the Response of Fiber Composites with Nonlinear Interface,'' Mech. Mater., to appear.
[5] Benveniste, Y., 1985, ''The Effective Mechanical Behavior of Composite Materials With Imperfect Contact Between the Constituents,'' Mech. Mater., **4**, pp. 197–208.
[6] Aboudi, J., 1991, *Mechanics of Composite Materials*, Elsevier, Amsterdam.
[7] Gurtin, M. E., 1984, ''The Linear Theory of Elasticity,'' *Mechanics of Solids* Vol. 2, C. Truesdell, ed., Springer-Verlag, Heidelberg, pp. 1–273.
[8] Needleman, A., 1992, ''Micromechanical Modelling of Interfacial Decohesion,'' Ultramicroscopy, **40**, pp. 203–214.
[9] Levy, A. J., 1996, ''The Effective Dilatational Response of Fiber Reinforced Composites With Nonlinear Interface,'' ASME J. Appl. Mech., **63**, pp. 357–364.
[10] Levy, A. J., and Dong, Z., 1998, ''Effective Transverse Response of Fiber Composites With Nonlinear Interface,'' J. Mech. Phys. Solids, **46**, pp. 1279–1300.
[11] Levy, A. J., 1997, ''On the Nucleation of Cavities in Planar Elasticity,'' Philos. Trans. R. Soc. London, Ser. A, **355**, pp. 2417–2458.
[12] Levy, A. J., 2000, ''The Fiber Composite With Nonlinear Interface—Part I: Axial Tension,'' ASME J. Appl. Mech., **67**, pp. 727–732.
[13] Sangani, A. S., and Mo, G., 1997, ''Elastic Interactions in Particulate Composites With Perfect as Well as Imperfect Interfaces,'' J. Mech. Phys. Solids, **45**, pp. 2001–2031.

**M. Ghandehari**
Department of Civil Engineering,
Polytechnic University,
Brooklyn, NY 11201

**S. Krishnaswamy**
Mem. ASME
Department of Engineering,
Northwestern University,
Evanston, IL 60208

**S. Shah**
Department of Civil Engineering
Northwestern University,
Evanston, IL 60208

# Bond-Induced Longitudinal Fracture in Reinforced Concrete

*Splitting of concrete caused by pullout of deformed rebars is investigated. The influence of specimen cross section size and geometry on the relationship between the components of stress and relative displacement at the interface is evaluated. Phase measurement interferometry is used for accurate mapping of the splitting cracks. The measured crack profiles, material model, and a fracture criterion are used in a hybrid experimental/numerical approach to evaluate the unknown normal component of traction at the interface.* [S0021-8936(00)03603-5]

## 1 Introduction

A considerable amount of research has been carried out in order to understand the splitting of concrete due to bond action, and to better predict the ultimate bond capacity of rebars embedded in concrete. Nevertheless, the experimental results reported by the research community indicate big differences in estimates of bond strength. Nonlinear fracture mechanics analysis of partly cracked cylinders has been able to predict the splitting of concrete with more success than methods based on strength criteria. However, a rational link between the splitting of a cylinder caused by internal pressure and bond strength is still missing. That is, the relation between the components of stress (normal and tangential) generated by the rebar lugs needs to be determined.

The interface shear stress and the associated relative tangential displacement between the steel rebar and concrete have been termed ''bond'' and ''slip,'' respectively. Due to the material stiffness mismatch and the relatively poor tensile strength of concrete, the interaction of the two materials leads to cracking of concrete. In general, the extent of elastic and inelastic deformations in the steel and concrete depends on the type of specimen (pullout, tension, etc.) in laboratory experiments, and on the characteristics of the rebar surface (lugs, etc.).

Bond-induced tensile cracking of concrete is classified in two main categories. When the rebar axis is in the plane of fracture, the crack is called ''longitudinal,'' and when the rebar axis passes through the fracture plane it is called ''transverse.'' The relative development of the shear stress $\tau$ and normal stress $\sigma$ is an important factor governing the sequence of the development of longitudinal versus transverse cracks. The development of these crack types is not entirely independent of each other, adding to the complexity in predicting bond cracks ([1]).

In this study, bond-induced longitudinal cracking (splitting) is investigated. These cracks initiate from the steel-concrete interface and propagate toward the concrete surface (Fig. 1). Considering a typical pullout test that fails by splitting, the slip in the initial stage is associated with the overall elastic deformation in addition to inelastic/crushing of concrete in front of the ribs (Fig. 1(*a*)). In the second stage, longitudinal cracks develop at the interface and propagate (Fig. 1(*b*)). Crack bridging stress, and the uncracked concrete ligament resist the internal pressure induced by the rebar lugs. The cracks then propagate through the entire concrete cross section (Fig. 1(*c*)). At this point, nearly all the bond capacity provided so far by the concrete hoop stresses is exhausted. In order to achieve greater strength, transverse reinforcement or externally applied transverse stress is required for confinement ([2–4]).

Tepfers in 1979 developed one of the first models for studying longitudinal cracking of concrete with an embedded rebar. The radial component of the bond stress was regarded as hydrostatic pressure acting against a thick wall cylinder. The splitting of the cylinder was evaluated based on the lower bound and the upper bound limits of the strength criteria. In the former, the criterion is satisfied when the tangential stress of the cylinder inside surface reaches the material tensile strength, and in the later it is satisfied when the tensile strength across the entire thickness of the cylinder is exhausted. A partly cracked model was also formulated but it did not account for the traction across the longitudinal cracks. Other studies by Tepfers [5] considered an externally confined concrete specimen with rebar in pullout. The normal components of traction at the rebar-concrete interface was evaluated by measuring the strain in the confining steel ring.

Prediction of the load capacity of a concrete cylinder subject to internal pressure using the cohesive crack formulation has been gaining popularity recently ([6–9]). The concrete cylinder is considered as an intact elastic outer ring and a cracked cohesive (softening) inner ring, subject to hydrostatic pressure acting on the circumference of the inner surface. An integral equation of equilibrium is solved, by expressing the crack opening in terms of the tangential stress ($\sigma_t$) according to the material post peak response in tension. In order to describe the tangential strain variation in the radial direction, certain assumptions are made regarding the tangential displacement variation in the cracked zone.

Den Uijl and Bigaj [10] incorporated the radial displacements of concrete in a fracture mechanics model similar to that described above. The relative magnitude of the displacement components (*dn* and *dt*) at the interface are shown to depend on the mode of failure. A linear relation between *dt* and *dn* is suggested for the splitting type failure. On the other hand when the anchorage failure is by shear/pullout, *dt* is shown to increase at a much faster rate after reaching the peak load vicinity.

In this study, an experimental program is carried out to evaluate the stages of splitting of concrete in the absence of confining effects other than the concrete cover. Keeping the rebar size constant, the influence of cross section size and geometry on the splitting of concrete is examined. An interferometer is devised for accurate detection of cracking of concrete during rebar pullout. It is commonly believed that longitudinal cracks initiate regardless of specimen size, and that the extent of propagation of these cracks depends on the concrete cover. Based on the interferometric images, it is shown that longitudinal cracks are not initiated if

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, May 21, 1999; final revision, Feb. 16, 2000. Associate Technical Editor: K. T. Ramesh. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

 **Copyright © 2000 by ASME** 

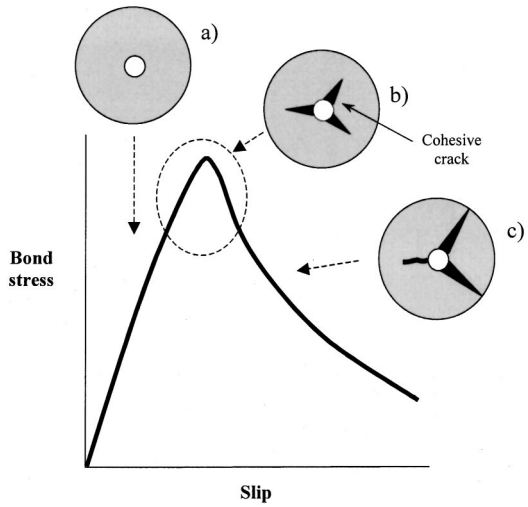



**Fig. 1 Bond stress and longitudinal splitting of concrete**

the mode of failure is changed from splitting to pullout. It must be added that there is always a possibility that longitudinal cracks may have initiated on the back surface of the specimen. It is further shown that this apparent size-dependent failure mode tendency also depends on changes in the specimen geometry. A study of the bond in reinforced concrete is carried out with particular attention to the influence of the cross section size and geometry. The influence of these parameters on cracking of concrete and, in particular, on the relationship between the normal and tangential stress components is identified.

## 2 Experimental Program

The experimental program is designed to observe the process of splitting that occurs in pullout test specimens. The cracking of concrete is monitored while measuring and evaluating the states of stress and displacement at the bar-concrete interface. The quantity of interest that cannot directly be measured is the normal component of traction at the interface. A hybrid numerical-experimental procedure is adopted where the measured crack profile, a material model and a fracture criterion are used to estimate the normal stress.

Since concrete splitting by rebar pullout is caused by the wedging action of the rebar lugs, specially configured wedge specimens are also tested to validate the procedure by which the rebar splitting force is obtained. This validation is possible as the normal component of the wedge force is independently available for this configuration.

Finally, the proposed model is validated by evaluating its effectiveness in simulating the shear strength reported from tests on specimens of other geometries conducted at McGill University and The University of Texas.

**2.1 Speckle Interferometer.** In concrete, crack-opening displacements near peak load are typically less than 50 microns ([11]). Speckle pattern interferometry ([12]) is a powerful technique for detection and measurement of such small displacements. This technique is powerful because the measurement gauge or sensitivity is on the order of a fraction of wavelength ($\lambda = 633$ nm) of the monochromatic light commonly used. In order to obtain an accurate measurement of the crack length and an accurate assessment of the cohesive stress bridging the cracks, displacement measurements of high resolution are required. Phase measurement speckle interferometry ([13]), is used for accurate evaluation of crack lengths and crack-opening displacements ([12]). The experimental setup is shown in Fig. 2. The specimen is installed such that the crack propagates vertically, and so the interferometer is configured to measure the surface displacements in the horizontal direction to obtain the crack-opening displacement. Light from a He-Ne laser is passed through a beam splitter. One leg contains a mirror attached to a piezoelectric transducer used for phase stepping. Prior to illuminating the object, each beam is expanded by passing through an objective lens and a spatial filter. The two wavefronts are scattered from the rough specimen surface and the scattered beams are recombined at the CCD array of a camera. The observed random/speckled intensity variation is due to the interference of the scattered wavefronts with random phase. Phase-stepped speckle intensity patterns of the specimen in its reference state are first recorded and then subtracted from the phase-stepped intensity pattern of the object in its deformed states. The resulting intensity fringe patterns associated with the specimen surface displacement are

$$I_i = I_0 + \{1 + \gamma_0 \cos[\phi(x,y) + \alpha_i]\} \quad i = 1,2,3 \tag{1}$$

for the three phase steps (Fig. 2).
from which the phase can be extracted:

$$\phi(x,y) = \tan^{-1}\{[\sqrt{3}(I_3 - I_2)]/[2I_1 - I_2 - I_3]\}. \tag{2}$$

The displacement $d$ mapped by each phase fringe is

$$d = \lambda/2 \sin\theta \tag{3}$$

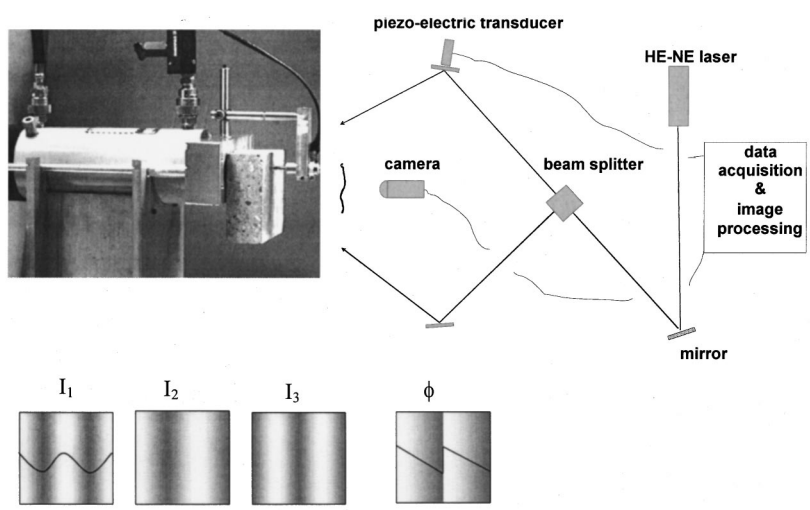



**Fig. 2 Loading apparatus and the interferometric setup**

**Table 1 Specimen Details**

rebar diam. $(d_b)=9$ mm, compressive strength $(f'c)=33$ MPa

| Specimen | Dimensions (Height×Width) (mm) |
|---|---|
| Edge rebar, small | 50×75 |
| Edge rebar, medium | 100×150 |
| Edge rebar, large | 150×225 |
| Center rebar, small | 50×75 |
| Center rebar, medium | 100×150 |
| Center rebar, Large | 150×225 |
| Edge wedge, small | 50×75 |
| Edge wedge, medium | 100×150 |
| Edge wedge, large | 150×225 |

where $\lambda=0.633\ \mu$m (He-Ne wavelength) $\theta=45$ deg (angle between the illuminating beams and the specimen normal).

**2.2 Rebar Specimens.** Normal strength concrete with water, cement, fine and coarse aggregate content proportions, respectively, of 0.5, 1, 2, and 2 parts by weight was used for all the specimens. The maximum aggregate size was 8 mm. The specimens were cured in 90 percent relative humidity for 28 days. A compressive strength of 33 MPa and a modulus of elasticity of 25,000 MPa were obtained as measured with 76 mm × 152 mm cylindrical test specimens. A tensile strength of 2.5 MPa was obtained using 100 × 30 × 20 mm prism specimens.

The pullout specimen dimensions are shown in Table 1 and Fig. 3. Two different specimen geometries were used keeping the same rebar size. In the first geometry (edge specimens), the rebar was cast near the specimen edge with a notch cut on the top so that the splitting crack propagates as an edge crack in one direction only. In the second geometry (center specimens), the rebar was cast in the center of the specimen. The aspect ratio (width/height) of all specimens was 1.5. The specimen dimensions were 50 × 75 (small), 100 × 150 (medium), and 150 × 225 (large) millimeters. All specimens were 50 mm thick, and the rebar diameter was 8 mm with ribs at 90 deg with respect to the longitudinal axis. In order to induce pressure in the horizontal direction for obtaining a reliable measure of the normal stress, the rebar was partly milled along the top and bottom surfaces of the longitudinal lugs.

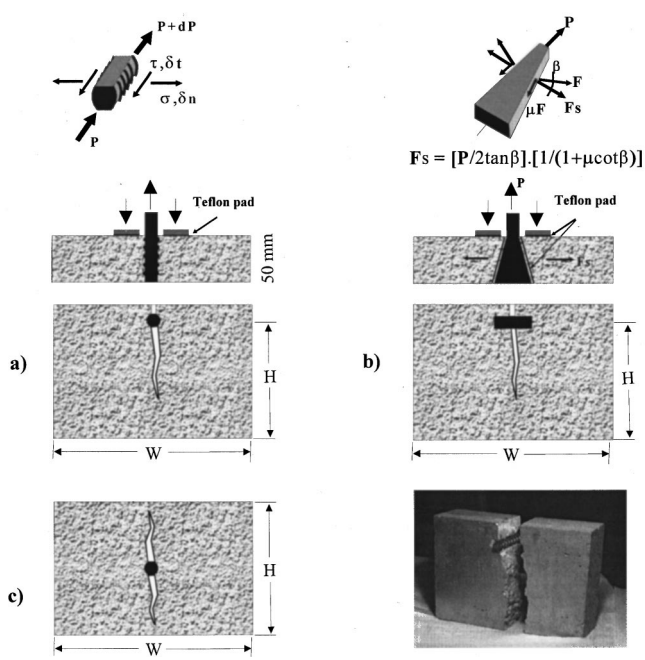



**Fig. 3 Pullout splitting tests**

Load was applied by a 200-kN center bore Enerpac hydraulic actuator (Fig. 2). The threaded end of the rebar was attached to a steel bar passing through the actuator. The pullout load was measured with a 25-kN Sensotec load cell. The slip was measured using an LVDT at the free end of the rebar. Teflon pads were placed at the specimen bearing reactions in order to limit the effect of reaction friction against splitting.

**2.3 Wedge Specimens.** Wedge specimens were tested in order to validate the procedure by which the normal stress for the rebar specimens is evaluated. The relation between the force components acting on the wedge is shown in Fig. 3. The dimensions of the wedge specimens are the same as the edge rebar specimens. Wedge splitting tests with this configuration were chosen in order to simulate the bearing boundary condition of the rebar pullout test. Two layers of thin Teflon pads were placed between each wedge-concrete interface establishing a reliable/known coefficient of friction. The coefficient of friction $\mu$ of Teflon against Teflon is 4.5 percent, smaller than the coefficient of friction of Teflon against concrete or steel. Teflon pads were also placed at the specimen bearing reactions in order to limit the effect of bearing friction against splitting.

## 3 Numerical Simulation

**3.1 Edge Rebar Specimens.** Figures 4 and 5 show typical wrapped phase maps representing the surface displacement of the edge and center rebar specimens. Since only the crack-opening displacement and crack lengths are of primary interest, automated two-dimensional unwrapping procedures ([13]) to obtain a displacement map were deemed unnecessary. Rather, by monitoring the extent of fringe discontinuities (reflecting displacement discontinuities across the crack) the crack length was directly measured. Manual unwrapping of the phase images along either side of the crack was performed, and the results were used to obtain the crack-opening displacement profiles.

Crack propagation was observed to be stable in the edge specimens all of which failed by splitting of the concrete. A typical crack profile for an edge rebar specimen is shown in Fig. 6. The crack-opening displacement was translated to a corresponding crack flank traction (Fig. 7) through an exponential function relating stress to crack opening ([11]) (Appendix A). The measured crack length at each load step was incorporated into the finite element mesh ([14]), and the crack flank traction was used as load input to the program. The stress intensity factor $Kl$ was evaluated using singular crack-tip elements. The crack propagation criterion used in the numerical simulations is that of the stress intensity factor reaching the critical stress intensity factor $(Kl_c)$, where $Kl_c$ is based on an effective elastic model (two-parameter model by [15]) (Appendix B). The normal stress (splitting force) at the rebar required to satisfy this fracture criterion was then calculated. The crack was incrementally propagated to the next experimentally measured length, and the procedure was repeated. The numerical simulation was performed in two dimensions, assuming plane-strain conditions.

**3.2 Center Rebar Specimens.** Unlike the edge rebar specimens, the mode of failure of the center rebar specimens was size dependent. The small and medium center specimens failed abruptly by splitting of concrete upon developing small longitudinal cracks. However, the large center specimen failed by pullout of rebar with no longitudinal cracks of any extent observable. The interface normal stress $\sigma$ at peak for the small and medium center specimens were evaluated in the same way as that outlined above for the edge specimens. The normal stress $\sigma$ in the elastic range was evaluated for both the center and the edge specimens by numerical simulation of the dilatation at the interface and correlating the results with the experimentally measured dilatation.

Since stable crack propagation was not achieved, no optically measured crack-opening profile or crack length data were avail-

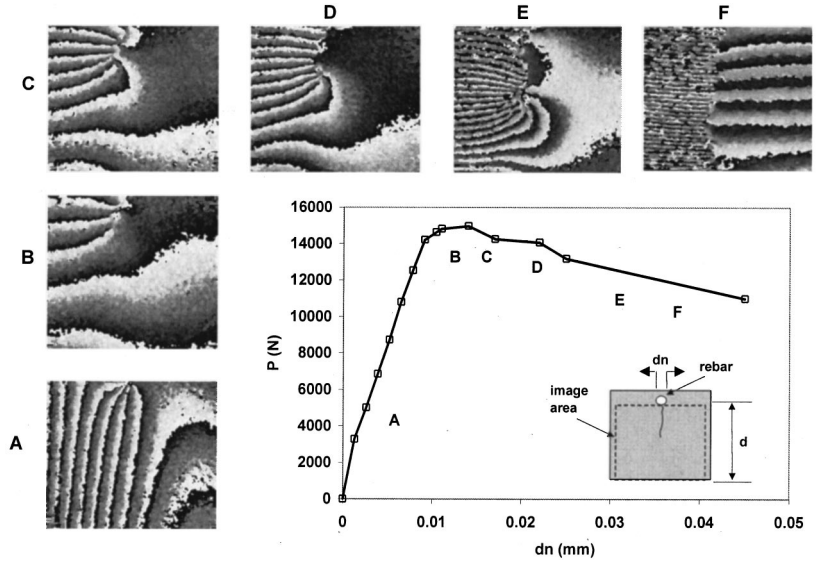



**Fig. 4 Crack propagation in the edge rebar specimen ($d$=150 mm)**

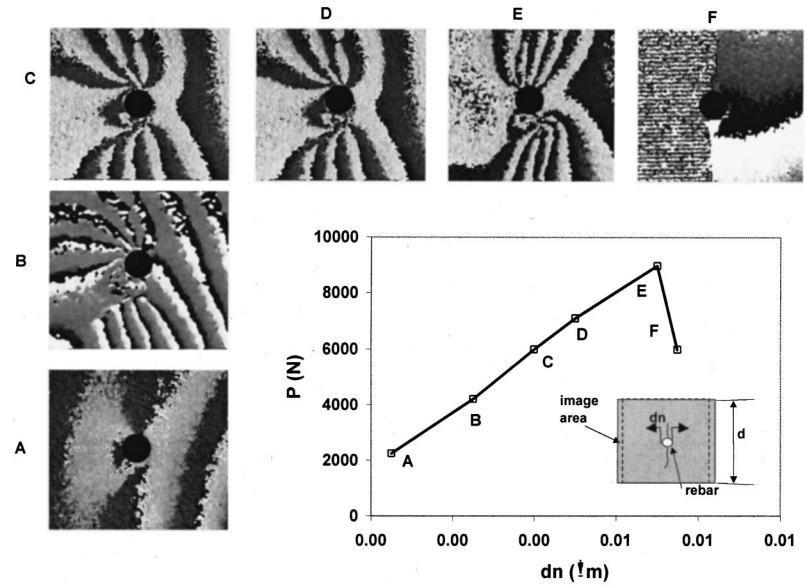



**Fig. 5 Crack propagation in the center rebar specimen ($d$=50 mm)**

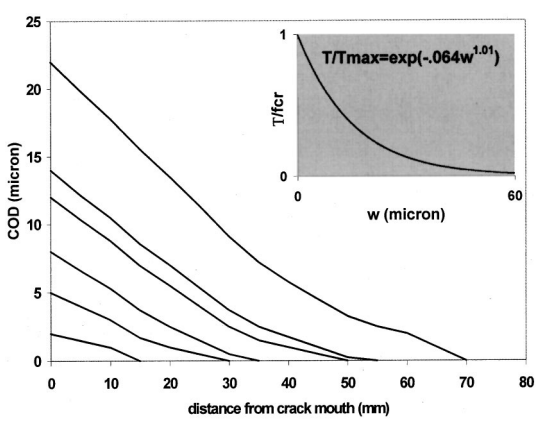



**Fig. 6 Measured crack profile**

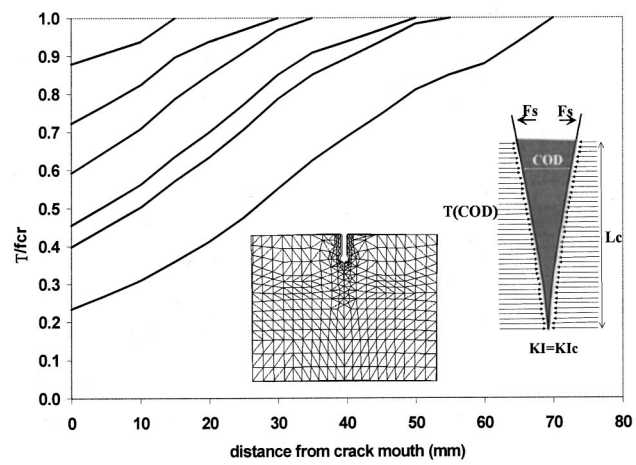



**Fig. 7 Calculated crack flank traction used as input for numerical simulation**

able in this case. Numerical evaluation of the splitting force for the center rebar specimens was therefore performed using an iterative scheme. The scheme is as follows:

1 Start with some small initial guess for the crack length.
2 Guess an initial crack-tip opening displacement (CTOD).
3 Assume linear crack profile.
4 Apply crack flank traction $T$ based on the material model discussed in Section 3.1.
5 Apply and increase the splitting force such that the crack tip stress intensity factor $Kl = Kl_c$.
6 Check the corresponding numerical crack profile (COD).
   If the output crack profile (COD) and the input $T$ satisfy the material model to within a stipulated accuracy, then consistency is achieved.
   Grow the crack to a new length and repeat Steps 2 to 6.
   Else guess a new CTOD and repeat Steps 3 to 6 using the new crack profile as given by the current iteration.

That this iterative algorithm leads to consistent results was established by adopting the same scheme for other specimens (where experimental data were available for comparison) as discussed in Section 4.2.

## 4 Results and Discussions

**4.1 Rebar Specimens.** Figure 8 shows the relation between the pullout force $P$ and the normal displacement $dn$. In general, increase in the size of the specimens which failed by splitting, led to an increase in the pullout strength. This is due to enhancement in the confinement. In the case of the edge specimens, this effect was pronounced. In the case of the center specimens it was less pronounced. As a matter fact, when the size of the center specimen is increased from medium to large, the pullout strength was observed to decrease as the failure mode changed from splitting to pullout. These trends are based on performing tests on three specimens of each type.

Figure 9 shows the relation between the two components of relative displacements at the rebar-concrete interface. The points of crack initiation (as detected on the specimen surface) and the peak load are marked with a cross and a star, respectively. Cracks initiate at approximately 6 to 10 microns of relative normal displacement ($dn$), but at larger relative shear displacement ($dt$) with increasing confinement. The increase in confinement is associated

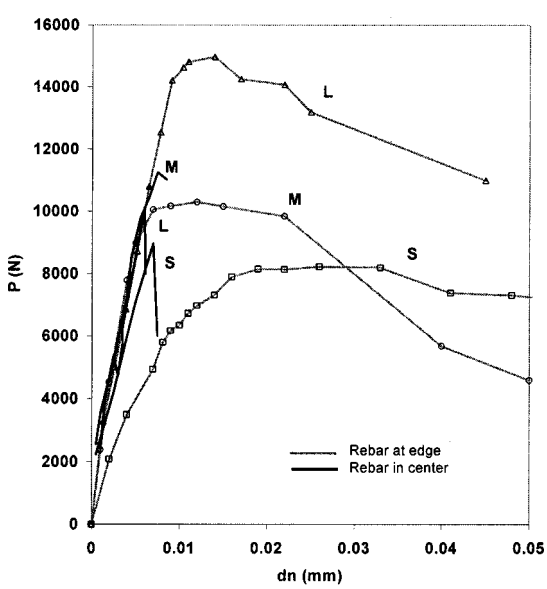



**Fig. 8 Rebar pullout force versus concrete expansion at the rebar**

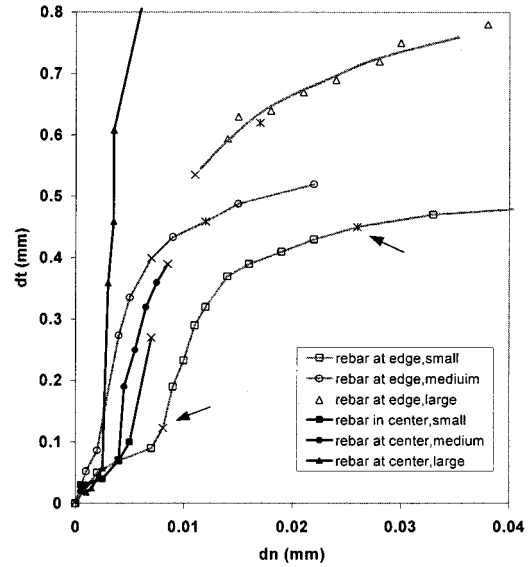



**Fig. 9 Components of rebar-concrete relative displacement**

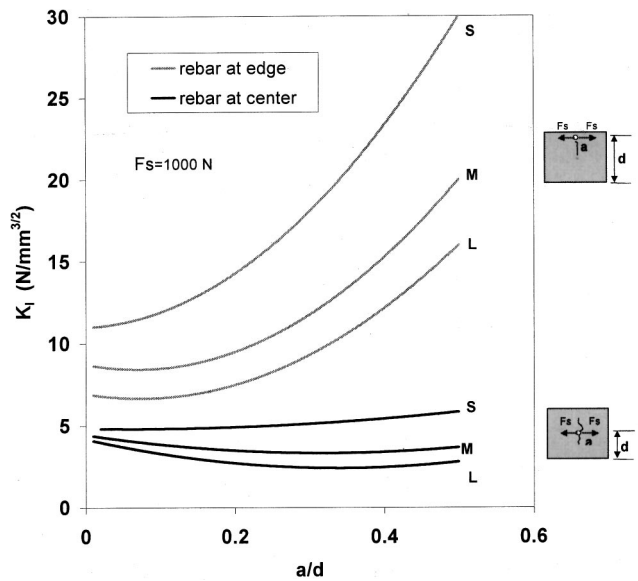



**Fig. 10 Stress Intensity factor ($KI$) for rebar specimens**

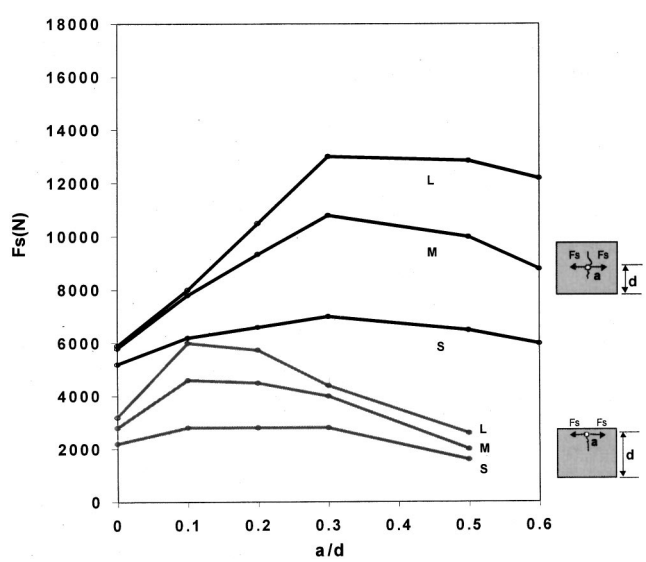



**Fig. 11 Numerical simulation of 9-mm rebar pullout**

either with increasing the specimen size or with changing the specimen geometry from edge to center. In the case of small and medium center rebar specimens, the point of crack initiation and peak pullout loads coincide.

Figure 10 shows the influence of specimen dimensions (size, and location of rebar) on the stress intensity factor ($Kl$) as obtained from a linear fracture mechanics analysis. $Kl$ is shown as a function of the nondimensional crack length ($a/d$) with constant applied splitting force $Fs$. Comparing the relative values of $Kl$ in the figure, it can be expected that the edge specimens would exhibit a larger splitting tendency than the center rebar specimens. The same tendency is expected as the specimen size is decreased.

The evaluated splitting force ($\sigma \times$ projected area) for both edge and the center geometries are shown in Fig. 11. The influence of increasing size on pullout load (Fig. 8) is consistent with trends observed on the increase of splitting force (Fig. 11) and decrease of stress intensity factor (Fig. 10).

**4.2 Wedge Specimens and Validation of the Hybrid Method.** Wedge-splitting tests were carried out primarily to validate the selected fracture criterion. Figure 12 shows the splitting force versus the crack-tip opening displacement (CTOD), and phase maps for a typical wedge specimen. The splitting force for the small, medium, and large wedge specimens are also compared. The results show that, as expected, the peak load and the corresponding CTOD increase with increasing specimen size.

Figures 13 and 14 show the crack profiles (COD), and crack flank traction at various load steps during crack propagation of the large wedge specimen. The prediction of the splitting force for the wedge specimens is performed in the same way as for the rebar specimens. Figure 15 shows the comparison of the simulated with the measured splitting force. It is shown that after accurate measurement of the crack profiles, the numerically predicted wedge splitting force at peak agrees with the experimental results when the fracture criterion $Kl = Kl_c$ is used. The alternative fracture criterion $Kl = 0$ and the tensile strength criterion were also considered. However, those simulations predicted forces much lower than that measured experimentally. The overprediction of the pre-peak response observed in the figure is attributed to the effect of

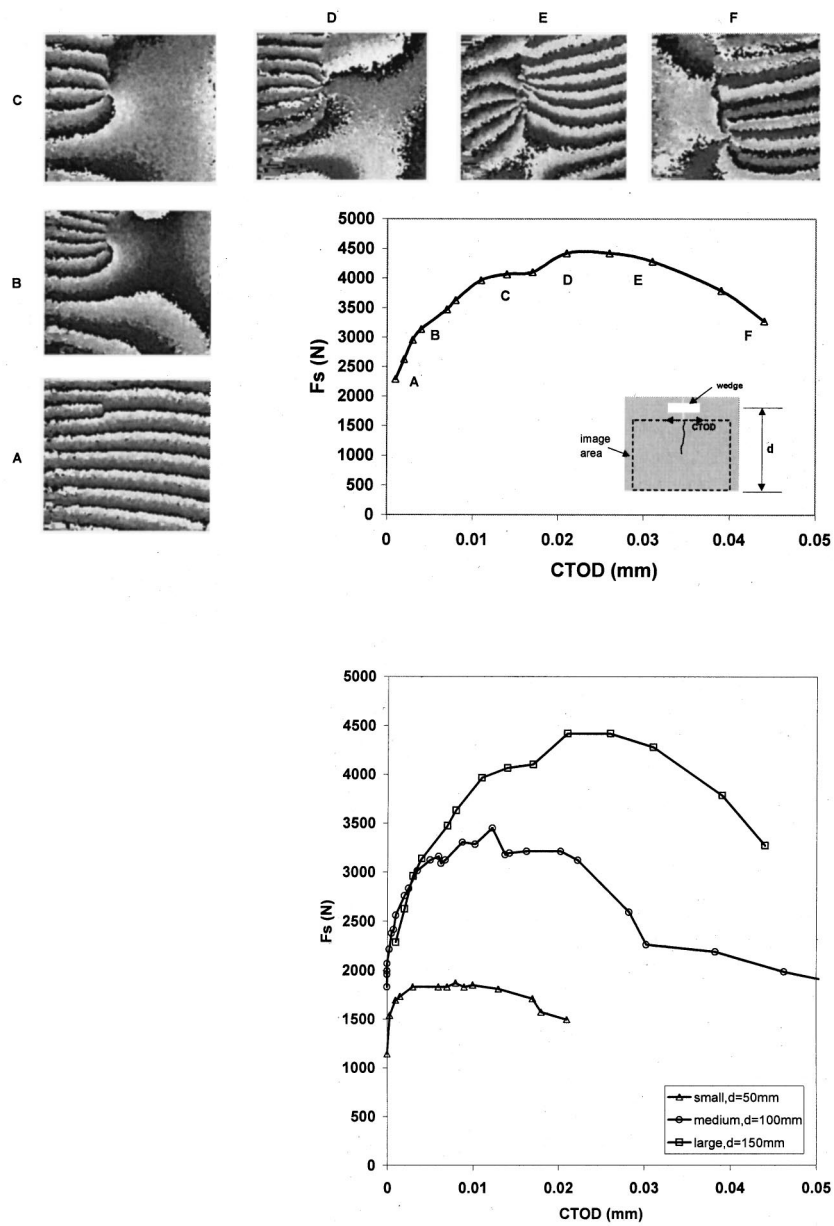



**Fig. 12 Splitting by wedge pullout**

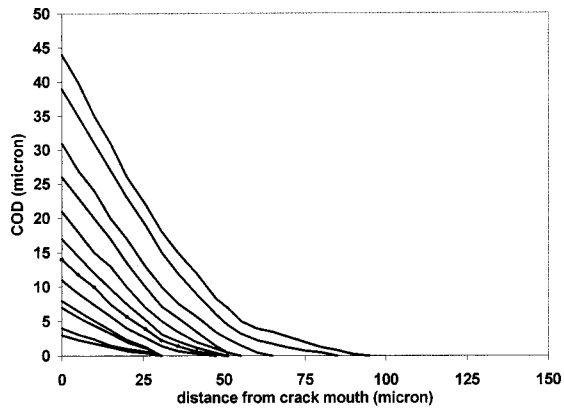



**Fig. 13 Measured crack profile**

the distributed microcracking at the crack tip which is not accounted for in the analysis. It is also partly attributed to possible deviations of the crack front profile (assumed straight in a plane-strain analysis) through the specimen thickness. These issues require further study.

**4.3 Confinement Parameter $\alpha$.** The measured shear stress and the evaluated normal stress for the rebar specimens are used here to establish a confinement parameter $\alpha = \arctan \sigma/\tau$ which is shown in Fig. 16. In the case of the edge rebar specimens, the full trace of $\alpha$ is shown as a function of slip $dt$. In the case of the center rebar specimens, $\alpha$ is only shown at peak, as stable fracture was not achieved. It is observed that $\alpha$ at peak load is about 60 deg for the center rebar specimens, and is approximately 40 deg for the edge rebar specimens.

From the above, it appears that $\alpha$ is geometry dependent, but essentially size independent. The later is based only on observations of edge and center rebar specimens. It is proposed that the variation of $\alpha$ with specimen geometry be functionally parametrized using concrete cover ratio $c1/c2$ (Fig. 16). Furthermore, it is proposed that the stress intensity factor ($Kl$) be used as a parameter for determining $\alpha$ at other cover ratios (that is when the rebar is located in between the edge and the center geometries). The rationale for this is drawn based on the effectiveness of the stress intensity factor (Fig. 10) in following the trends of splitting strength (Fig. 11). Numerical evaluation of the stress intensity factor for specimens of different geometry (based on rebar location) is performed and the stress intensity factor calculated. The form of stress intensity variation as a function of the rebar location is then adopted for scaling the confinement parameter $\alpha$. The proposed $\alpha$ shown in Fig. 16 demonstrates that the influence of

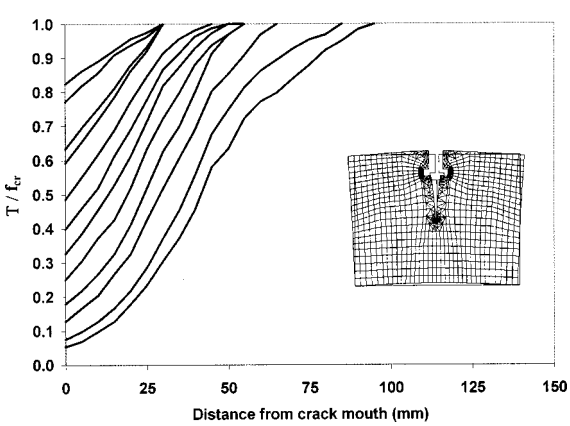



**Fig. 14 Calculated crack flank traction used as input for numerical simulation**

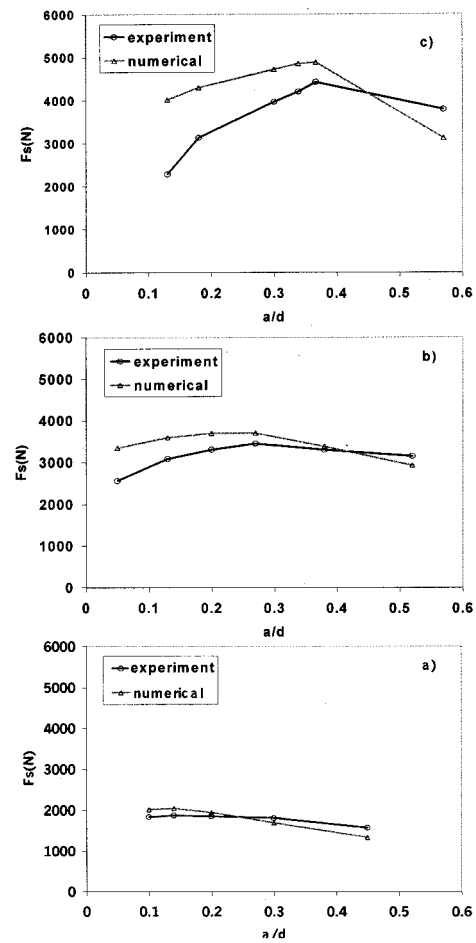



**Fig. 15 Numerical simulation of the wedge splitting test**

increasing the top cover on $\alpha$ is substantial when the ratio of top to bottom cover ($c1/c2$) is small. With increasing $c1/c2$ (as the rebar approaches the center of the specimen) this influence is shown to diminish.

**4.4 Model Validation.** The proposed confinement parameter $\alpha$ is used to predict/simulate tests performed at McGill University and the University of Texas on other specimen geometries. When testing a pullout specimen a nonuniform bond stress distribution is expected along the rebar. A uniform distribution can be assumed for a short specimen (specimen length $L<5$ bar diameter $d_b$, such as that used in this work). A unique procedure was developed at McGill in order to generate uniform bond stress along a relatively long rebar embedded in concrete ($L=15db$) Abrishami [16]. The influence of specimen size and rebar size was investigated. It was reported that increasing the specimen size resulted in an increase in the pullout strength, and increasing the rebar diameter resulted in a decrease in the pullout strength. The iterative numerical procedure outlined in Section 3.2 along with a confinement parameter of $\alpha = 60$ deg valid for center geometry are used to obtain an estimate for the peak bond stress measured at McGill. An approximate fracture toughness corresponding to the reported concrete strength ($f'c = 25$ Mpa) is chosen for the analysis. It was reported by the McGill researchers that the specimens failed with formation of a splitting crack passing through the rebar. That crack pattern is simulated accordingly. Specimen dimensions and the results of the numerical simulation are shown in Table 2 indicating reasonably good agreement with the experimental results.

The influence of the position of the rebar on bond strength was investigated at the University of Texas ([17]). A special loading

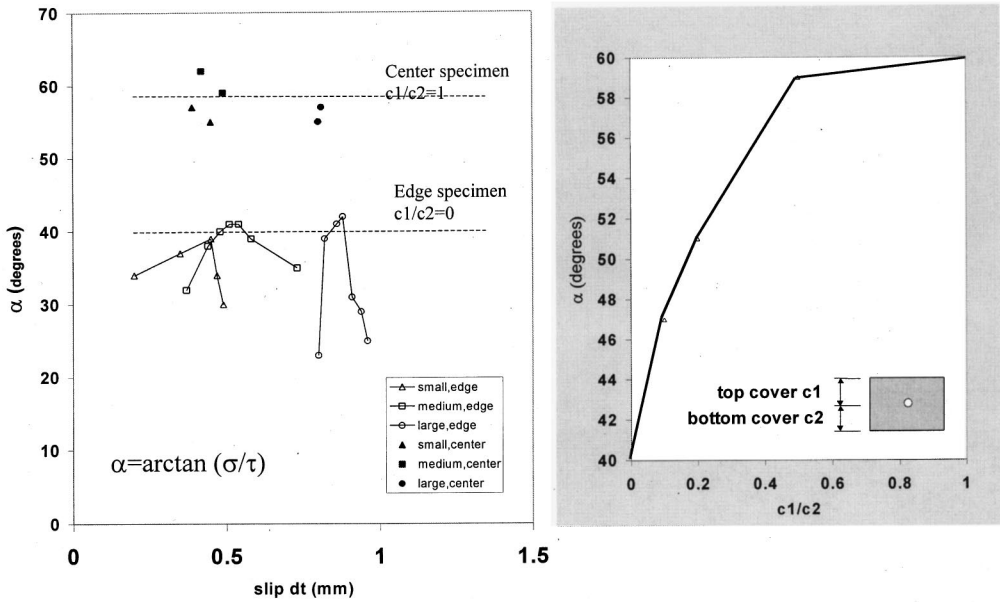
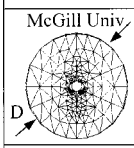
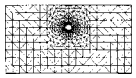



**Fig. 16 Effect of specimen size and geometry on the confinement parameter**

**Table 2 Prediction of tests by McGill University and University of Texas**

| Specimen shape | Specimen size (mm) & geometry | rebar db (mm) | $\tau_{max}$ experiment | NU prediction. |
|---|---|---|---|---|
| McGill Univ.  D | D=200 | 36 | 3.6 | 4 |
| | D=200 | 20 | 5.6 | 7.9 |
| | D=150 | 20 | 4.2 | 6.2 |
| Univ. of Texas  | 230x450 c1/c2=0.1 (top cover c1) (bottom cover c2) | 20 | 5 | 5.4 |
| | 230x450 c1/c2=0.45 | 20 | 6.25 | 6.7 |

configuration was used in order to generate uniform bond stress as indicated by strain gauge instrumented rebars. The specimens exhibited longitudinal cracks appearing on the concrete surface. Increasing the ratio of top to bottom concrete cover was found to increase the pullout strength. Specimen dimensions and the results of the numerical simulation for these tests are also shown in Table 2. Two different concrete cover ratios $c1/c2$ (0.1 and 0.45) were simulated using the proposed confinement parameter and the iterative procedure in Section 3.2. Once again reasonably good agreement is obtained with the experimental results lending confidence to the validity of the proposed formulation. Experimental scatter of tests performed at McGill and University of Texas are not available for assessing the quality of the predicted results.

## 5 Conclusions

Nonlinear fracture mechanics is used to model the bond/splitting failure in reinforced concrete. Singular cohesive crack modeling is shown to provide good prediction of the experimental results.

Edge rebar specimens were observed to fail by splitting. Increasing the specimen size has a pronounced effect on increasing the pullout load.

Small and medium center rebar specimens failed by splitting, but increasing the specimen size had a less pronounced effect on increasing the pullout load. Stable post-peak response was not obtained for the splitting center rebar specimens. Closed loop testing of this experiment, is planned for the future.

A large center rebar specimen was observed to fail by pullout without development of longitudinal cracks of any measurable extent.

A confinement parameter $\alpha = \arctan(\sigma/\tau)$ is found to be specimen geometry dependent but essentially size independent. $\alpha$ is experimentally determined to be 60 deg for the center specimen and 40 deg for the edge specimen.

Finally, it must be stated that the present model is valid for steel to concrete material interface with standard rebar proportions. For example, the application of epoxy coating to the rebar for the purpose of corrosion protection or alteration of the lug geometry will change the confinement parameter $\alpha$.

## Acknowledgments

Support from the National Science Foundation Center For Science and Technology of Advanced Cement Based Materials (ACBM), as well as the Department of Mechanical Engineering at Northwestern University is gratefully acknowledged. The author would also like to thank Dr. Liu Sheng Wang for his advice with the image processing.

## Appendix A

**A Constitutive Model for Crack Flank Traction in Concrete.** Concrete is a quasi-brittle material with substantial pre-peak nonlinearity ([18]). For example, the stress-elongation of a concrete prism subject to tension may be divided into four stages. In the first stage (up to 30 percent of peak load), the propagation of the internal voids is negligible. In the second stage (up to 80 percent of peak), internal voids propagate into isolated cracks randomly distributed through out the specimen volume. In this stage the distribution of internal strain is relatively uniform over the specimen length. Strain localization occurs in the third stage (up to peak load), followed by formation of a major crack in the fourth stage. A post-peak descending stage was estimated by Gopalaratnam [11], with a stress versus crack opening displacement relation shown in Fig. 6.

## Appendix B

**A Fracture Mechanic Model for Concrete.** Fracture of concrete is accompanied with a considerable crack flank traction and/or inelastic process zone. For that reason linear elastic fracture mechanics (LEFM) cannot be directly applied. An effective elastic crack governed by LEFM can, however, be introduced. The equivalence between the actual and the effective crack can be prescribed in the model such that the effective crack exhibits the same unloading compliance as the actual structure. Since stable crack propagation before peak load is present in concrete, the critical effective crack length is different form the initial. In order to predict the failure condition, at least two conditions must exist to determine the critical load and the corresponding crack length. Most effect crack models use two or more fracture parameters to define the fracture process. Jenq and Shah proposed the following criteria for an unstable crack:

$$Kl = Ki_c$$

$$\mathrm{CTOD} = \mathrm{CTOD}_C$$

Where $Kl$ and CTOD are the stress intensity factor and the crack-tip opening displacement. These can calculated based on LEFM. $Kl_c$ and $\mathrm{CTOD}_C$ are evaluated experimentally ([15]).

## References


[1] Lutz, L. A., and Gergely, P., 1967, ''Mechanics of Bond and Slip of Deformed Bars in Concrete,'' ACI J., November, pp. 711–721.

[2] Gambarova, P. G., Rosati, G. P., and Zasso, B., 1989, ''Steel-to-Concrete Bond After Splitting: Test Results,'' Mater. Struct., **22**, No. 127, pp. 36–47.

[3] Modena, C., 1992, ''Theoretical Prediction of the Ultimate Bond Strength Between a Reinforcing Bar and Concrete,'' *Bond in Concrete*, Riga Technical University, Riga, Latvia.

[4] Plizzari, G., Schumm, C., and Giuriani, E., 1987, ''The Effect of Residual Tensile Strength of Cracked Concrete on the Local Bond-Slip Law After Splitting,'' Studi E Ricerche, **9**, pp. 129–153.

[5] Tepfers, R., 1979, ''Cracking of Concrete Cover Along Anchored Deformed Reinforcing Bars,'' Magazine of Concrete Research, **31**, No. 106, pp. 3–12.

[6] Noghabai, K. 1995, Ph.d. disc., ''Splitting of Concrete in the Anchoring Zone of Deformed Bars,'' Lulea University of Technology.

[7] Reinhardt, H. W., 1992, ''Bond of Steel to Strain-Softening of Concrete Taking Account of Loading Rate,'' *Fracture Mechanics of Concrete Structures* (FraMCoS1), Breckenridge, CO, Elsevier, New York.

[8] Rosati, G., and Schumm, C., 1992, ''Modeling of Local Bond to Concrete Bond in Reinforced Concrete Beams,'' *Bond in Concrete*, Vol. 3, A. Skudra and A. Tepfers, eds., Riga Technical University, Riga, Latvia, pp. 34–43.

[9] Van der Veen, C., 1990, ''Cryogenic Bond-Stress-Slip Relationship,'' Ph.d. disc., Delft University of Technology.

[10] Den Uijl, J. A., and Bigaj, A. J., 1996, ''A Bond Model for Ribbed Bars Based on Concrete Confinement,'' HERON, **41**, No. 3, pp. 201–226.

[11] Gopalaratnam, V. S., and Shah, S. P., 1985, ''Softening Response of Plain Concrete in Direct Tension,'' ACI J., **82**, No. 3, pp. 310–323.

[12] Ghandehari, M., Krishnaswamy, S., and Shah, S. P., 1999, ''A Technique for Evaluating the Interactions at the Interface in Reinforced Concrete,'' J. Eng. Mech., **125**, No. 2, pp. 234–241.

[13] Creath, K., 1985, ''Phase-shifting Speckle Interferometry,'' Appl. Opt., **24**, No. 18, pp. 3053–3058.

[14] FRANC2D, 1995, ''A Two Dimensional Crack Propagation Simulator,'' Vol. 2.7, Cornell University, Ithaca, NY.

[15] Jenq, Y. S., and Shah, S. P., 1985, ''A Two-Parameter Fracture Model for Concrete,'' J. Eng. Mech., **111**, No. 4, pp. 1227–1241.

[16] Abrishami, H., 1993, ''Studies on Bond and Cracking of Structural Concrete,'' Ph.d. disc, McGill University.

[17] Ferguson, P. M., Thompson, J. N., 1965, ''Development Length for Large High strength Reinforcing Bars,'' ACI J., January, pp. 71–91.

[18] Shah, S. P., Swartz, S. E., and Ouyang, C., 1995, *Fracture Mechanics of Concrete*, John Wiley and Sons, New York.

**G. Rajesh**[1]
Beverage Technical Center,
American National Can,
2250 Lively Boulevard,
Elk Grove Village, IL 60007

**R. B. Bhagat**
Engineering Science and
Mechanics Department,
227 Hammond Building

**K. A. Fichthorn**
Chemical Engineering Department,
164 Fenske Building

The Pennsylvania State University,
University Park, PA 16802

# Reactive Flow in a Porous Medium: Formulation for Spatially Periodic Hexagonally Packed Cylinders

*This study develops an integrated micro-macro model of reactive flow in a porous medium consisting of spatially periodic hexagonal array of solid reacting cylinders. The micro model describes the growth of reaction product on the solid reactant surface. The macro flow of the infiltrant fluid is described by Darcy's law. The transient permeability and thus advancement of the infiltration front are determined as a function of process parameters from the micro model. Crucial process parameters that influence the advance of the fluid front are identified. The results from this investigation can be used to optimize the manufacture of ceramic-matrix composites.* [S0021-8936(00)02703-3]

## 1 Introduction

This study deals with the modeling aspect of reactive melt infiltration and estimates permeability as a function of time for a preform geometry of hexagonally packed array of cylindrical (specifically, carbon-coated silicon carbide) fibers. The overall objective of this research is to determine permeability, a macro flow parameter, in terms of manufacturing process parameters such as heat of reaction, initial temperature of fibers, etc., which finally would give information on the infiltration front. By achieving this objective, we will be able to optimize the parameters to maximize manufacturing efficiency as well as assure fully dense, high-quality ceramic composite products that is a main requirement in the industry.

The modeling of the complete process has been structured into two stages, namely, micro and macro modeling. At the micro level, a cylindrical geometry is considered for the fiber reinforcements and solutions to the problem of coupled heat and mass transfer with chemical reaction and volume expansion are derived and solved numerically. The porosity of a representative volume element (RVE) is calculated as a function of time from the growth of the SiC layer on the carbon-coated reinforcement. At the macro level, the permeability of the preform is then estimated as a function of the porosity in model. After determining the transient permeability, the position of the infiltration front at each instant of time is obtained by a simple integration procedure. The following section provides a literature review on reactive infiltration.

## 2 Background on Reactive Infiltration

**2.1 Process Models.** The method of reactive melt infiltration has been studied by Hillig and co-workers ([1–4]) for the purpose of fabricating silicon/silicon carbide composites for high temperature applications. Experiments have primarily been directed towards the silicon/carbon system due to the wide range of applicability of SiC/SiC composites. Ness and Page [5] extensively studied the evolution of the microstructure of $\alpha$-SiC and $\beta$-SiC during the manufacture of reaction bonded SiC. The work of Popper [6], Forrest et al. [7], and Ness and Page [5] laid the foundations for the detailed study of the infiltration of molten Si in carbon-coated SiC fiber preforms. With regard to the macroscopic flow studies during infiltration for the manufacture of composites, a detailed review has been given by Åström et al. [8].

Process models to help better understand infiltration are mainly based on the macroscopic flow of fluids through porous media (in our case, porous preforms). Darcy's law describes the flow of a fluid through a porous medium and is given by

$$q=-\frac{K}{\mu}\frac{\Delta P}{L} \tag{1}$$

where ''$q$'' is the macroscopic velocity of the fluid in the medium, ''$K$'' is defined as the permeability of the medium, $\mu$ is the coefficient of viscosity, and $\Delta P/L$ is the pressure gradient established over the length $L$. The permeability depends on the geometry of the medium.

Models developed to help understand the infiltration process focus mainly on the macro flow of the reactive melt. Edwards and Olson [9] and Martins et al. [10] presented a macro model based on the flow of the reactive melt using Darcy's law for flow through a porous medium for the manufacture of metal-matrix composites. Fitzer and Gadow [11,12] modeled the growth of the reaction product during infiltration using a micro model. *There is still a need to unite the macro and micro model developments for a complete description of infiltration.* One way of uniting micro and macro phenomena is to use statistical theories to describe infiltration ([13–16]).

**2.2 Permeability of Spatially Periodic Structures: Macro Modeling.** Chiang et al. [17] and Messner and Chiang [18] developed a macro model for the transient permeability of a porous preform using the Hagen-Poiseuille law for flow through a bundle of capillaries based on the model of Brittin [19]. In 1946 Brittin [19] developed a model for the infiltration process in an isotropic porous medium. The same assumption has been made in recent publications ([9,10]) dealing with the development of models for the infiltration kinetics in metal-matrix composites. The anisotropy of the porous preform has been considered by several investigators ([20–23]) using a permeability tensor in Darcy's law for porous media. To account for the temporal change of permeability in the preform, Messner and Chiang [24] proposed a one-dimensional model relating the permeability to the pore ''radius'' in a system of capillaries. Other models to predict permeability were proposed by Gebart [20], Cai [21,22], and Chan et al. [23].

[1]To whom correspondence should be addressed.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Oct. 21, 1997; final revision, Oct. 12, 1999. Associate Technical Editor: D. A. Signier. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

 Copyright © 2000 by ASME 

Perhaps the earliest attempt to express the permeability of a porous preform comprised of a bed of spherical particles in terms of its geometrical properties was made by Kozeny, as stated in Ref. [8]. Another approach adopted by Watson [25] for studies of flows in soils made use of the Navier-Stokes equations in two dimensions to describe the flow in a ''representative volume element'' (RVE). This study considered an RVE made of square blocks which was a crude representation of the actual geometry. A recent article on yet another attempt to obtain the permeability of a unidirectional fiber bed was given by Westhuizen and Du Pleiss [26].

**2.3 Micro Modeling.** Micro modeling in this study involves the investigation of the heterogeneous chemical reaction of the infiltrant fluid with a single cylindrical fiber (i.e., a solid reacting cylinder). In the literature on heterogeneous chemical reactions ([27–34]) (there are extensive treatises on heterogeneous reactions; only the references pertinent to our topic have been cited here), we can find extensive studies of gas-solid reactions with solid reaction products ([13–16]). Levenspiel [28] gives descriptions of heterogeneous reactions of solids when they react with fluids. These equations are also applicable to solid-liquid reactions. A model called the unreacted-core shrinking (URCS) model used by Yagi and Kunii [32] and also referred to and modified by Ausman and Watson [33] has been applied to solid-fluid reactions where the reaction is primarily diffusion controlled and a ''sharp'' interface is clearly visible between the solid reactant and the solid reaction product formed. The studies of Wen and Wang [34] for solid-gas reactions led to a modified definition of the effectiveness factor for nonisothermal URCS models for spheres and helped gain deeper insight into the thermal and geometric instabilities in a solid reacting particle during reaction.

**2.4 Integration of Micro and Macro Models.** One of the common approaches to solve problems occurring at a macro scale has been the continuum approach ([35,36]). Statistical theories have played a major role in including effects of chemical reactions in pores during infiltration of a porous preform by a gas ([17–20]). The porous preform is considered a ''pore network'' ([19]) similar to a crystal lattice network. This network model was used by Rege and Fogler [37] for carbonate cores infiltrated by acidic $FeCl_3$ solutions. This is essentially the idea of a spatially periodic medium where the geometry of the preform is invariant under a translation of magnitude equal to the unit cell dimension.

Statistical theories can be avoided for increasing the scale of the RVE to a macro scale since we are dealing with a spatially periodic geometry by following an approach called the method of multiple scales ([36]). The complete mathematical framework for the method of multiple scales approach is still in its developmental stages. The most well-known approach, however, is a ''physical approach'' ([36]) where macroscopic ''conductivities'' are found for a desired spatially periodic geometry. The ''conductivity'' may be an effective diffusivity, a permeability or a thermal conductivity.

## 3 Technical Approach

**3.1 Micro Modeling.** The following section first outlines the standard development of the unreacted-core shrinking (URCS) model as given by Levenspiel [28] and then extends the URCS model to include volume expansion of the solid reaction product.

*Standard Unreacted-Core Shrinking Model.* Let us consider a single spherical pellet of solid reactant material ''*B*'' which comes in contact and reacts with the fluid reactant ''*A*.'' The solid reaction product ''*C*'' thus formed on the spherical pellet grows on the remaining unreacted core which shrinks with more and more reaction. The chemical reaction can be written as

$$aA + bB \Rightarrow cC \tag{1a}$$

where $a$, $b$, and $c$ are stoichiometric coefficients for $A$, $B$, and $C$, respectively. The final expression in terms of the moving boundary of the unreacted core $r_c$ and the outer radius $R_o$ of the pellet (which does not change with time) is given by Levenspiel [28] as follows:

$$-\rho_B\left[\int_{r_c=R_o}^{r_c=r_c}\left(\frac{1}{r_c}-\frac{1}{R_o}\right)r_c^2 dr_c\right] = bD_{eA}C_{Ag}\int_{t=0}^{t=t} dt \tag{2}$$

where $\rho_B$ is the density of $B$, $D_{eA}$ is the effective diffusivity of $A$ within the core, $C_{Ag}$ is the concentration of reactant $A$, $b$ is the stoichiometric coefficient of $B$, and $t$ is time. *An assumption inherent in the above development is that there is no shrinkage or net volume expansion due to reactions*. Thus we have to account for volume expansion in the URCS model. A modification of the URCS model is outlined next.

*Modified Unreacted-Core Shrinking Model for Volume Expansion.* Let us assume we have identical conditions prevailing as in the ash-diffusion-controlled URCS model (presented by Levenspiel [28]) with the additional knowledge that now the outer moving radius with more reaction is not a constant, but is increasing with more volume expansion. Then, if $R$ is the outer radius of a sphere and $r_c$ the radius at the reaction front, we write $R = R(r_c)$, i.e., the outer radius $R$ depends on some fundamental manner on $r_c$ ([38–43]). From the initial reaction equation

$$A_f + B_s \rightarrow D_s \tag{3}$$

(where the subscripts ''*f*'' and ''*s*'' are fluid and solid, respectively), we have the rates of reaction of $A$, $B$, and $D$, respectively, related as follows:

$$-\text{rate}_A = -\text{rate}_B = \text{rate}_D\,. \tag{4}$$

Then

$$-\frac{dN_B}{dt} = \frac{dN_D}{dt} \tag{5}$$

where $N_B$ and $N_D$ are the number of moles of $B$ and $D$, respectively. From the above equation, after substituting for the number of moles as the molar density times the volume, then integrating and simplifying we get

$$\left(\frac{R}{R_o}\right)^3 = \alpha + (1-\alpha)\left(\frac{r_B}{R_o}\right)^3 \tag{6}$$

where $\alpha$ is the ratio of the molar density of the solid reactant $B$ to the molar density of the reaction product $D$, $r_B$ is the radius of unreacted core of $B$ (i.e., radius of the moving boundary$=r_c$ in Eq. (2)), $R$ is the outer moving or variable radius of the pellet, and $R_o$ is the initial outer radius. Setting the ratios to nondimensional radii, we have

$$\xi_{\text{out}}^3 = \alpha + (1-\alpha)\xi_{\text{in}}^3 \tag{7}$$

where $\xi_{\text{out}}$ and $\xi_{\text{in}}$ are $(R/R_o)$ and $(r_B/R_o)$, respectively. Following an identical procedure for cylindrical pellets (and considering only the radial growth of the reaction product), we would get

$$\xi_{\text{out}}^2 = \alpha + (1-\alpha)\xi_{\text{in}}^2. \tag{8}$$

Equations (7) and (8) for spherical and cylindrical geometries have to be solved simultaneously with Eq. (2) for $\xi_{\text{in}}$. Thus, mathematically, the main modification to the URCS model is that the integration in Eq. (2) cannot be carried out assuming $R_o$ a constant with respect to $r_c$ now.

*Nonisothermal Unreacted-Core Shrinking Models.* For exothermic or endothermic reactions, we have to consider the effects due to steep temperature gradients within the solid particles during reaction and diffusion. This would be an extension of the URCS models with and without volume expansion of the reaction product along with solutions of the heat conduction equation.

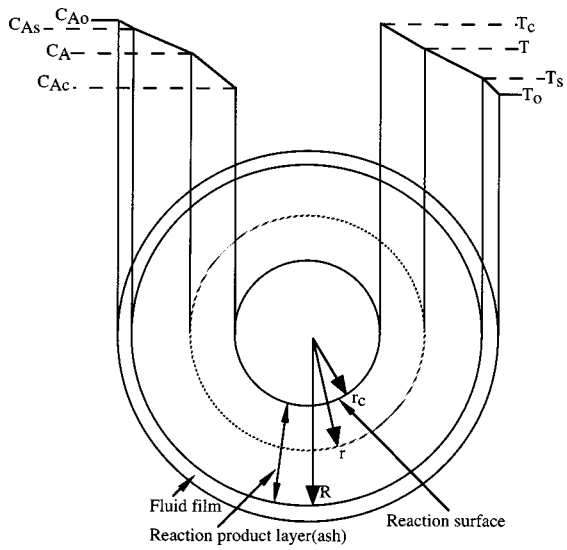



**Fig. 1 General schematic of concentration and temperature profiles in nonisothermal unreacted-core shrinking models after Wen and Wang [34]**

These equations now offer no closed-form solution due to the coupling of heat and mass transfer (from the Arrhenius form of the chemical reaction rate) and have to be solved numerically. Figure 1 shows the general concentration and temperature profiles that are obtained from the nonisothermal URCS models. The following developments are based on Wen and Wang's [34] nonisothermal URCS models which they developed for spheres without volume expansion.

*Nonisothermal Solutions for Cylinders Including Volume Expansion.* Let us denote $C_A$ to be the concentration of $A$ and $r$ the radius of the cylinder. We start from the basic pseudo-steady-state equation for the mole fraction of the fluid reactant (similar to Wen and Wang [34] which was developed for spheres) given by

$$\frac{\partial^2 C_A}{\partial r^2}+\frac{1}{r}\frac{\partial C_A}{\partial r}=0 \quad \forall \quad r_c<r<R(r_c) \tag{9}$$

with the boundary conditions

$$@r=R \quad D_{eA}\frac{\partial C_A}{\partial r}\bigg|_{\substack{r=R\\T=T_o}}=k_{mA}(C_{Ao}-C_{As}) \tag{10}$$

$$@r=r_c \quad D_{eA}\frac{\partial C_A}{\partial r}\bigg|_{\substack{r=r_c\\T=T_c}}=ak_sC_{so}C_{Ac} \tag{11}$$

$$@r=r_c \quad -D_{eA}|_{T_o}\frac{\partial C_A}{\partial r}\bigg|_{r=r_c}=aC_{so}\frac{dr_c}{dt} \tag{12}$$

and the initial condition

$$@t=0 \quad R=R_o \text{ and } r=R_o. \tag{13}$$

In the above equations, $D_{eA}$ is the effective diffusivity of $A$, the subscript ''$o$'' stands for the bulk fluid infiltrant, subscript ''$s$'' is at the outer surface of the cylindrical particle, ''$i$'' is the initial value at $t=0$, $k_{mA}$ the mass transfer coefficient of $A$ across the fluid film around the unreacted-core, and $k_s$ the rate constant for the reaction between $A$ and $B$ as shown in Eq. (3). It is important to note here that

$$R=R(r_c). \tag{14}$$

The energy balance is given by

$$\frac{\partial^2 T}{\partial r^2}+\frac{1}{r}\frac{\partial T}{\partial r}=\frac{C_{pe}}{k_e}\frac{\partial T}{\partial t} \quad \forall \quad r_c<r<R \tag{15}$$

with the following boundary condition for accumulation:

$$@r=R \quad -k_e\frac{\partial T}{\partial r}\bigg|_{r=R}=h_c(T_s-T_o). \tag{16}$$

where $C_{pe}$ is the volumetric heat capacity of the reaction product layer, $k_e$ is the effective thermal conductivity of the reaction product layer, and $h_c$ the heat transfer coefficient. The accumulation term now is given by

$$@r=r_c$$

$$(2\pi r_c)ak_{s(T_c)}C_{so}C_{Ac}(-\Delta H)+2\pi r_ck_e\frac{\partial T}{\partial r}\bigg|_{r=r_c}=\pi r_c^2\rho_c c_{pc}\frac{dT_c}{dt} \tag{17}$$

with the initial condition for temperature being

$$@t=0 \quad T=T_i=T_c \tag{18}$$

where $\Delta H$ is the heat of reaction, ''$a$'' is the stoichiometric coefficient of ''$A$,'' $c_{pc}$ is the heat capacity of the unreacted core, and $r_c$ is the density of unreacted core. Obtaining the nondimensional form of the equations, the mole balance reduces to

$$\frac{d^2\omega}{d\xi^2}+\frac{1}{\xi}\frac{d\omega}{d\xi}=0 \quad \forall \quad \xi_c<\xi<\xi_o \tag{19}$$

where $\omega_i$ is the mole fraction of the $i$th reactant, $\xi$ being $r/R_o$, with the boundary conditions changing at the outer surface of the cylinder due to volume expansion, given by

$$@\xi=\xi_o \quad \frac{d\omega}{d\xi}\bigg|_{\xi=\xi_o}=N_{\mathrm{Sh}}(1-\omega_s) \tag{20}$$

where $\xi_o$ denotes the outer moving radius, $N_{\mathrm{Sh}}$ the modified Sherwood number which is given by $k_{mA(T_o)}R_o/D_{eA(T_o)}$, and

$$@\xi=\xi_c \quad \frac{d\omega}{d\xi}\bigg|_{\xi=\xi_c}=\phi_s\frac{\omega_c}{U_c}\exp\left[\frac{E}{\Re T_o}\left(1-\frac{1}{U_c}\right)\right] \tag{21}$$

$$@\xi=\xi_c \quad \frac{d\omega}{d\xi}\bigg|_{\xi=\xi_c}=-\phi_s\frac{d\xi_c}{d\theta} \tag{22}$$

$$@\xi=\xi_c \quad \omega=\omega_c \tag{23}$$

with the initial condition for the reaction front and the outer moving front given by

$$@\theta=0 \quad \xi_c=\xi_o=1 \tag{24}$$

where $\phi_s$ is the modified Thiele modulus given by a $k_{s(T_o)}C_{So}R/D_{eA(T_o)}$, $\theta$ a nondimensional time$=k_{s(T_o)}C_{Ao}t/R_o$, $U_i$ the nondimensional temperature$=T_i/T_o$, $\Re$ the universal gas constant, and $E$ the activation energy of reaction.

The nondimensionalization factor is selected same as before (given below):

$$\text{Nondimensionalizing factor}=ak_{s(T_o)}C_{Ao}C_{so}(-\Delta H)\xi_c \tag{25}$$

The effectiveness factor is defined as

$$\eta_s=\frac{2\pi r_cak_{s(T_c)}C_{so}C_{Ac}}{2\pi r_cak_{s(T_o)}C_{so}C_{Ao}}=\frac{k_{s(T_c)}C_{Ac}}{k_{s(T_o)}C_{Ao}}$$

$$=\frac{\omega_c}{U_c}\exp\left[\frac{E}{\Re T_o}\left(1-\frac{1}{U_c}\right)\right]=-\frac{d\xi_c}{d\theta} \tag{26}$$

The derivation of the five equations for this model for cylinders is detailed in Refs. [38–40].

Summarizing the equations to be solved, we have

$$\omega_c = \omega_s + N_{\mathrm{Sh}}(1-\omega_s)\xi_o \ln\left(\frac{\xi_c}{\xi_o}\right) \tag{27}$$

$$\frac{N_{\mathrm{Sh}}(1-\omega_s)\xi_o}{\phi_s \xi_c} = \frac{\omega_c}{U_c}\exp\left[\frac{E}{\Re T_o}\left(1-\frac{1}{U_c}\right)\right] = \eta_s \tag{28}$$

$$\frac{N_{\mathrm{Sh}}(1-\omega_s)\xi_o}{\phi_s \xi_c} = -\frac{d\xi_c}{d\theta} = \eta_s \tag{29}$$

$$U_s = U_c + \xi_o \ln\left(\frac{\xi_c}{\xi_o}\right) N_{\mathrm{Nu}(c)}(U_s - 1) \tag{30}$$

and

$$\left\{\frac{G\xi_c}{2} - S\xi_c^2 + S\xi_o^2 + \frac{SQ}{1-W} - SQ\right\}\frac{dU_c}{d\theta}$$
$$= \frac{\omega_c}{U_c}\exp\left[\frac{E}{\Re T_o}\left(1-\frac{1}{U_c}\right)\right] - \frac{N_{\mathrm{Nu}(c)}(U_s-1)}{\beta\phi_s\left(\frac{E}{\Re T_o}\right)}\frac{\xi_o}{\xi_c}$$
$$+ SU_c\eta_s(2\xi_o R_{r_c}) - \left(\frac{SQ}{1-W}\right)V(U_s-1) - SP(U_s-U_c)$$
$$- \frac{A}{\beta\left(\frac{E}{\Re T_o}\right)}\left(\frac{\xi_o R_{r_c} U_s \eta_s}{\xi_c}\right) \tag{31}$$

where

$$S = \frac{A}{2\beta\left(\frac{E}{\Re T_o}\right)\xi_c} \tag{32}$$

$$Q = \xi_o^2 + \frac{\xi_o^2 - \xi_c^2}{2\ln\left(\frac{\xi_c}{\xi_o}\right)} \tag{33}$$

$$P = \frac{dQ}{d\theta} = -\eta_s\left[2\xi_o R_{r_c} + \frac{\xi_c\xi_o \ln(\xi_c/\xi_o)(2\xi_o R_{r_c} - 2\xi_c) - (\xi_o^2 - \xi_c^2)(\xi_o - \xi_c R_{r_c})}{2\xi_c\xi_o[\ln(\xi_c/\xi_o)]^2}\right] \tag{34}$$

$$W = \xi_o \ln\left(\frac{\xi_c}{\xi_o}\right) N_{\mathrm{Nu}(c)} \tag{35}$$

$$V = \frac{dW}{d\theta} = -\eta_s N_{\mathrm{Nu}(c)}\left[\frac{\xi_o - \xi_c R_{r_c}}{\xi_c} + R_{r_c}\ln\left(\frac{\xi_c}{\xi_o}\right)\right]. \tag{36}$$

In the above equations, $N_{\mathrm{Nu}(c)} = h_c R_o / k_e$, the modified Nusselt number for convective heat transfer. For a cylinder, we have, from Eq. (8), that the outer moving radius given by

$$\xi_o^2 = \alpha + (1-\alpha)\xi_c^2. \tag{37}$$

Equation (37) has to be solved in combination with Eqs. (27) through (31) for solutions of nonisothermal URCS processes for cylindrical fibers. The solutions were obtained using a simple Euler finite difference scheme to solve for the six essential unknowns ($\omega_c$, $\omega_s$, $U_c$, $U_s$, $\xi_c$, and $\xi_o$) with the help of a *C* program written for this purpose.

**3.2 Macro Modeling: Permeability of Hexagonally Packed Array of Cylinders.** Since permeability depends mainly on the geometry of the medium, i.e., the pore geometry, it is important to determine the variation of permeability as a function of the porosity of the medium. For this purpose the following approach was adopted. Finite element models were developed for RVEs of hexagonally packed carbon-coated SiC fibers with porosity values ranging from 95 percent to 15 percent, corresponding to a volume fraction range of 5 percent to 85 percent, respectively. The maximum capacity of a hexagonally packed RVE is a volume fraction of 90.69 percent. A pressure gradient across the RVE was specified ([38]) and the Navier-Stokes equations for steady flow through the RVE were solved. Once the steady-state solutions were obtained for the velocities in the RVE, the inlet velocities were then averaged with respect to the element lengths at the inlet as shown in Fig. 2, i.e., the average velocity in the RVE is estimated since this is needed in Darcy's law. The permeabilities were estimated from Darcy's law given by Eq. (1), identical to the procedure adopted by Gebart [20]. The pressure gradients for each volume fraction were calculated from the relation given by Messner and Chiang [24] for capillary pressure as

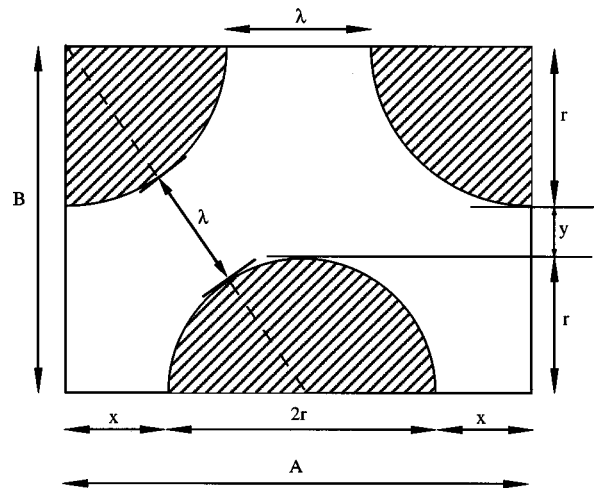



**Fig. 2 Representative volume element with dimensions shown for calculation of permeability**

$$\Delta P_{\mathrm{cap}} = (\gamma_{SV} - \gamma_{SL})\frac{S_P}{V_P} = \left(\frac{S_P}{V_P}\right)\gamma_{LV}\cos\theta \tag{38}$$

where $\gamma_{SV}$, $\gamma_{SL}$, and $\gamma_{LV}$ are the solid/vapor, solid/liquid, and liquid/vapor interfacial energies; for the Si/C system these correspond to the carbon/air, carbon/liquid Si, and liquid Si/air interfacial energies, respectively. $S_p$ and $V_p$ stand for the surface area and volume, respectively.

*Porosity and Permeability as Functions of Time: Integration.* Once we have the growing radii of the cylinders as a function of time in the RVE, the next step is to obtain the porosity function. We determine the porosity from the simple areal fraction as follows (since the geometry is cylindrical):

$$\text{proposity} = 1 - V_f = 1 - \frac{\pi r^2}{AB} = 1 - R_o^2\frac{\pi\xi_o^2}{AB} \tag{39}$$

where $V_f$ is the volume fraction, and $A$ and $B$ denote the dimensions of the RVE. Once we have the porosity functions, we refer back to observe the variation of the permeability with porosity. The remaining task is to substitute for the porosity variation with time into the permeability relation. Finally, substitution of the permeability in Darcy's law describes the motion of the infiltration front. The results from these studies are discussed next.

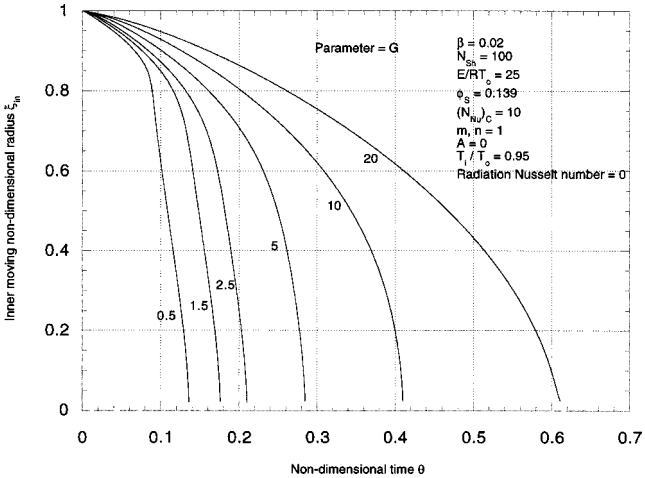



**Fig. 3 Inner moving radii (cylinders) for different values of *G***

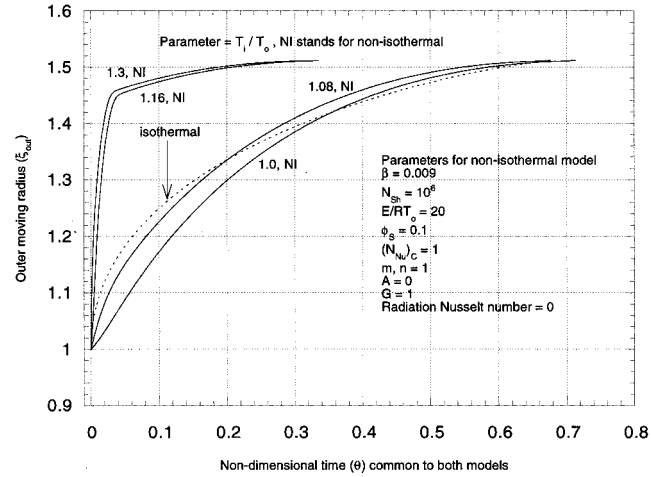



**Fig. 5 Comparison of isothermal and nonisothermal unreacted-core shrinking models and the effect of initial temperature of carbon**

## 4 Results and Discussion

**4.1 Nonisothermal URCS Models: Porosity and Transient Permeability.** Let us first consider Figs. 3, 4 and 5. A significant effect in the outer radii is observed due to slight changes in the values of $G$ and initial temperatures, respectively. Thus these two parameters (initial temperature $T_i$ and $G$) were considered the most significant factors contributing to drastic changes in permeability with time. The depletion of the unreacted core was measured from Fig. 3 for different values of $G$. Corresponding to these curves, the growth of the reaction product layer was obtained from Eq. (37) and were plotted on Fig. 4. Porosity functions were obtained for Figs. 4 and 5 from Eq. (39) for the RVE shown in Fig. 2. These functions of time are shown in Figs. 6 and 7, respectively. The permeabilities were then determined as a function of time for each curve by calculating the permeability value as a composite function of the porosity values for the RVE. The methodology to determine the permeability was as follows. Schematically, if $K$ was the permeability, $\theta$ was time and $\phi$ the porosity, we find $\phi(\theta)$ from Figs. 6 and 7. Then from Fig. 8, we get $K(\phi)$ and thus, $K[\phi(\theta)]$. However, $\phi(\theta)$ depends on the outer radius $R(\theta)$ given by each curve in Figs. 4 and 5. Each curve for $R(\theta)$ in turn corresponds to a set of parameter values in the nonisothermal URCS models with volume expansion. This is the way micro-macro integration is accomplished.

From observations of Figs. 4 and 5, the $T_i$ and $G$ values seemed to influence the porosity functions the most and hence were selected to generate the permeabilities in Figs. 9–12. In Figs. 9 and 10, we observe the steep fall in permeability for higher initial temperatures of the solid reactant due to the reaction rate being very fast compared to diffusion, thus causing a steep increase in $R(\theta)$, in turn leading to a steep decrease in porosity. This effect is less significant when the initial temperature nearly matches the bulk fluid temperature. This effect would play an important role in deciding preform preheat temperatures while carrying out reactive infiltration for CMCs. In Fig. 10, we observe the same effect as in Fig. 9, except that the values are lower due to the initial volume fraction being higher for Fig. 10.

The effect of having different systems with different $G$ values on permeability are depicted in Figs. 11 and 12 for two initial volume fractions for the RVE shown in Fig. 2. Higher values of $G$ give rise to thermally stable systems and lower values result in instability ([38]). The higher permeabilities for high $G$ values are attributed to lower heats of reaction compared to the heat capacity of solid reactant (unreacted core). Lower heats of reaction correspond to lower exothermicity of reactions and thus slower reaction rates, thus giving slower rates of decrease of permeability.

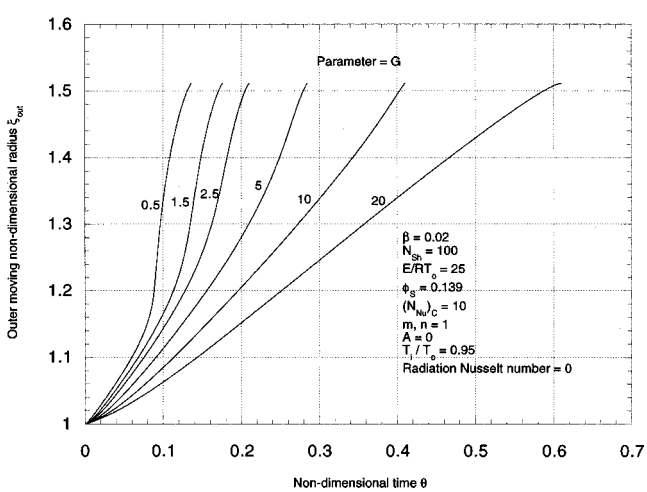



**Fig. 4 Outer moving radii (cylinders) for different values of *G***

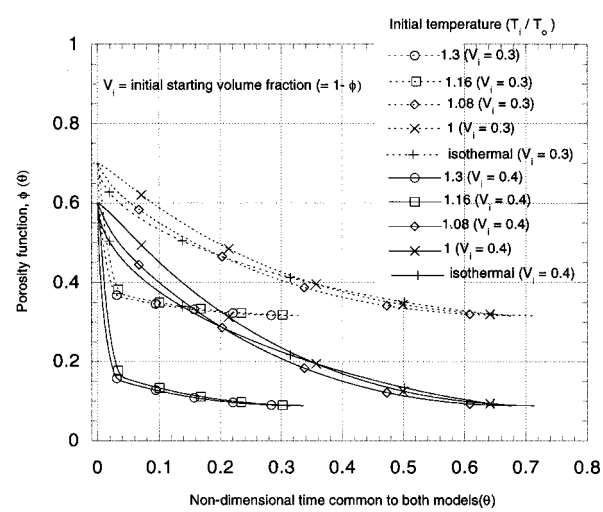



**Fig. 6 Porosity as a function of time and the effect of initial temperature (hexagonal packing of cylinders) in the Si/C system**

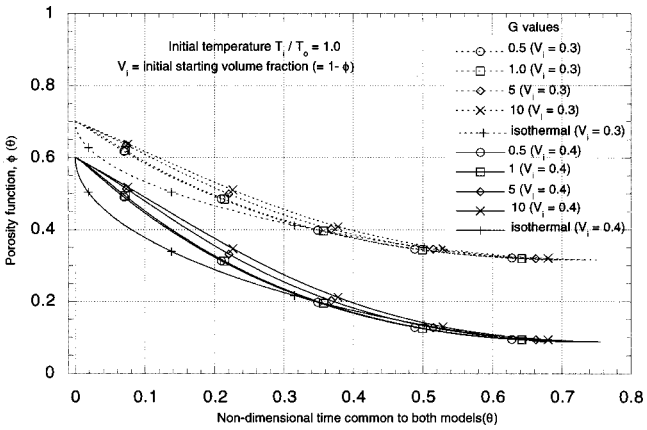



**Fig. 7 Porosity as a function of time and the effect of *G* values (hexagonal packing of cylinders) in the Si/C system**

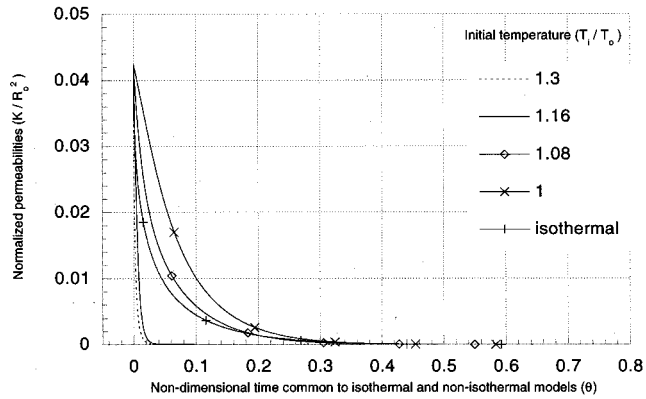



**Fig. 10 Permeability as a function of time for different initial temperatures, for a starting volume fraction of 0.4 (hexagonal packing of cylinders) in the Si/C system**

For the silicon/carbon system, since the diffusivity of silicon carbide is extremely low to liquid silicon (refer Table 1), we would require the permeability to decrease at a relatively low rate. Thus for the silicon/carbon system, for a slow rate of decrease of permeability, we need to start with a low preheat carbon-coated SiC temperature. It has also been observed that the heat of reaction in this system differs significantly for different kinds of carbon ([11,12,44,45]). Thus we need to select a carbon type such that we get a high heat capacity to heat of reaction ratio ($G$) values. This would ensure slower rates of decrease of permeabilities in an array of hexagonally packed carbon-coated SiC cylinders.

Finally, an estimate of the three-dimensional permeability tensor for a hexagonally packed array of cylinders can be given as

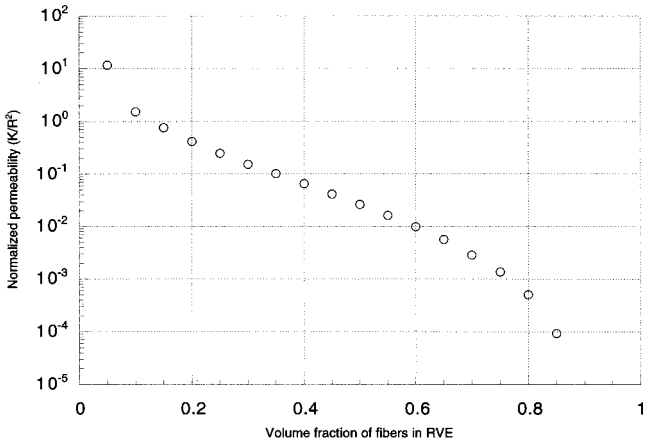



**Fig. 8 Normalized transverse permeability values used in this study (hexagonal packing of cylinders; fixed representative volume element size)**

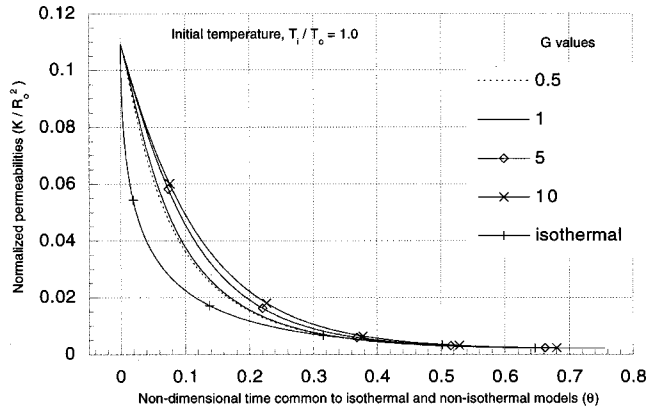



**Fig. 11 Permeability as a function of time for different *G* values, for a starting volume fraction of 0.3 (hexagonal packing of cylinders) in the Si/C system**

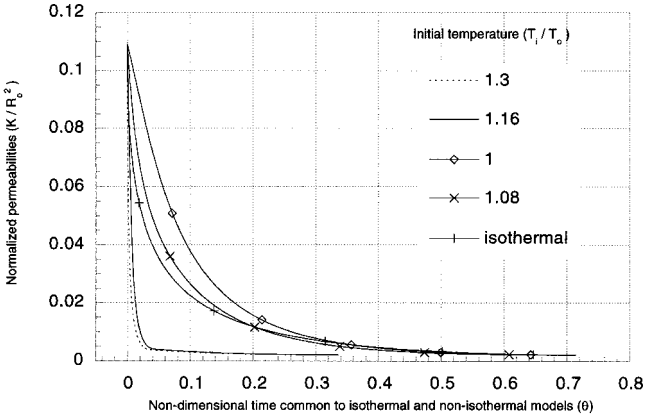



**Fig. 9 Permeability as a function of time for different initial temperatures, for a starting volume fraction of 0.3 (hexagonal locking of cylinders) in the Si/C system**

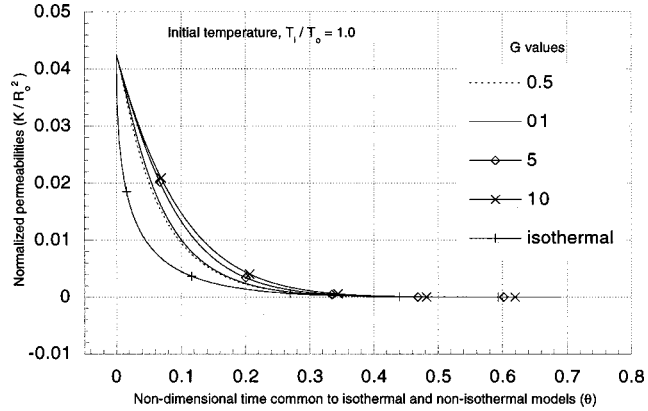



**Fig. 12 Permeability as a function of time for different *G* values, for a starting volume fraction of 0.4 (hexagonal packing of cylinders) in the Si/C system**

**Table 1 Properties used for the Si/C system (properties at 1700 K have been used)**

| Material | Property | Value | Units |
|---|---|---|---|
| **Silicon** | Density | 2470.0 | $kg/m^3$ |
| | Viscosity | $6.2 \times 10^{-4}$ | $Ns/m^2$ |
| | Specific heat | 1046.7 | J/kg/K |
| | Conductivity | 22.0 | W/m/K |
| | Diffusivity | $9. \times 10^{-11}$ | $m^2/s$ |
| | | | |
| **Carbon** | Density | 1100.0 | $kg/m^3$ |
| | Specific heat | 2093.4 | J/kg/K |
| | Average fiber dia | 10.0 | μm |
| | Volume fraction for some sections | 30.0 % | - |
| | Pre-exponential factor of Si/C reaction | $\approx 10^{25}$ | $mol / m^3s$ |
| | | | |
| **Silicon carbide** | Specific heat | 1256.0 | J/kg/K |
| | Activation energy of formation | 60.0 | kJ/mol |
| | Heat of reaction when formed | $115.0 \times 10^3$ | J/mol |

$$\tilde{K}=\begin{pmatrix} K_{\parallel\ \mathrm{hex}} & 0 & 0 \\ 0 & K_{\perp\ \mathrm{hex}} & 0 \\ 0 & 0 & K_{\perp\ \mathrm{hex}} \end{pmatrix}$$

where $\tilde{K}$ denotes the permeability tensor, given in terms of the principal axes' directions for a hexagonally packed array of cylinders. The quantity $K_{\parallel\ \mathrm{hex}}$ is the permeability of the hexagonal RVE to longitudinal flow, which was estimated to a good degree of accuracy in terms of the hydraulic radius by Messner and Chinag [18]. The second quantity $K_{\perp\ \mathrm{hex}}$ is the permeability to transverse flow obtained from the present work in terms of several quantities, namely, the initial temperature of the solid reactant, heat of reaction, heat capacity of the reaction product, density of product and initial concentration of solid reactant. These values are given quantitatively for the Si/C system in the form of curves on Figs. 9 through 12 for two initial volume fractions in the RVE of Fig. 2. *It is important to note that some parameters (for instance, the pre-exponential factor for the rate constant) for the Si/C system need to be accurately determined and the values used here (as shown in Table 1) are just estimates.* Once the permeability tensor is obtained, we can now modify Darcy's law in Eq. (1) and write it as follows:

$$\mathbf{u}=-\frac{\tilde{K}}{\mu}\nabla P \tag{40}$$

where $\mathbf{u}$ is the macroscopic velocity vector in three dimensions, $m$ the coefficient of viscosity of the infiltrant, and $\nabla P$ the pressure gradient in the principal directions. The permeability tensor in Eq. (40) is now written as

$$\tilde{K}=\tilde{K}\{\phi[(\xi_o(\theta))]\}. \tag{41}$$

Equation (41) completes the macroscopic (global) description of reactive infiltration in terms of the microscopic (local) phenomena accounted for in the transient permeability tensor $\tilde{K}$.

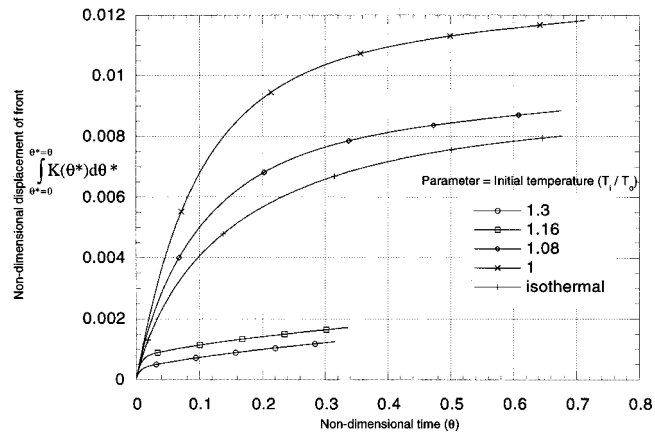



**Fig. 13 Infiltration front displacement for different initial temperature values (initial volume fraction=30 percent)**

**4.2 Tracking the Infiltration Front.** Once the permeability has been found as a function of time, we can write a nondimensional form of Darcy's law as follows:

$$\frac{d\chi}{d\theta}=U=u\left(\frac{\mu}{R^2\nabla P}\right)=K(\theta) \tag{42}$$

where here, $K$ is the nondimensional permeability normalized by $R^2$ as shown in Fig. 8. Here $\chi$ is the nondimensional ''displacement'' or the position of the infiltration front and $U$ is the nondimensional velocity of the front. Then the position of the front for flow transverse to the cylindrical fibers can be obtained as

$$\chi(\theta)=\int_0^{\theta}K(\theta^*)d\theta^*. \tag{43}$$

(*Important assumptions here are that the pressure gradient is constant with time, and that the permeability is spatially constant. In reality, the pressure gradient and the spatio-temporal permeability have to be solved for along with the material balance equation in addition to Eq. (42). Relaxation of these assumptions is planned for future work.)*

The nondimensional infiltration front positions have been plotted in Figs. 13 through 16 for the corresponding permeability variations already discussed in Figs. 9–12. The integrations were performed numerically from the permeability values. We observe that the front advance is very fast during the initial stages of reaction when the rate of permeability decrease is the largest. As

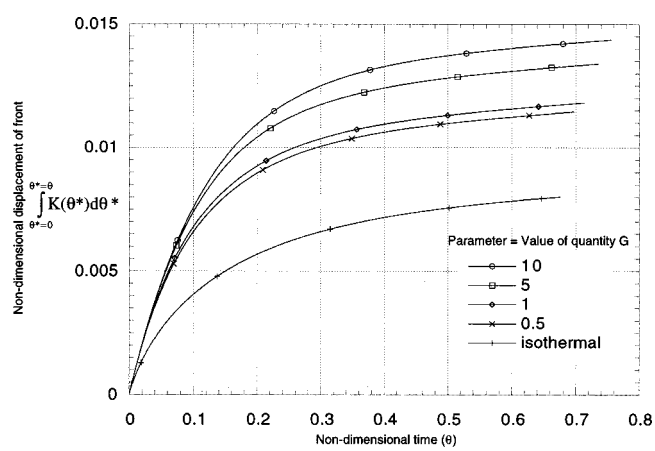



**Fig. 14 Infiltration front displacement for different *G* values (initial volume fraction=30 percent)**

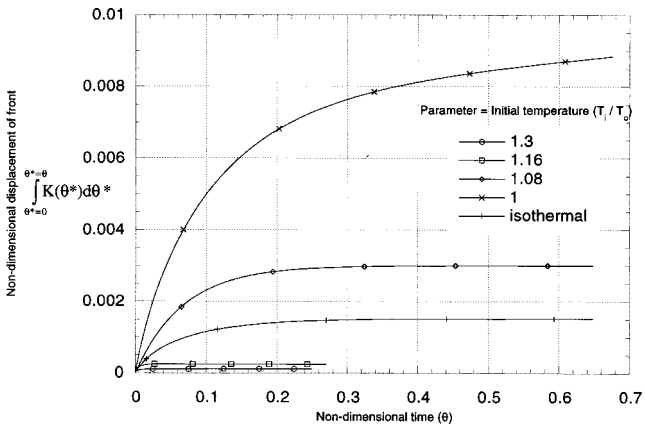



**Fig. 15 Infiltration front displacement for different initial temperature values (initial volume fraction=40 percent)**

the rate of permeability decrease is slower and as the porosity starts to diminish, the advance of the front becomes slower. The effect of high initial temperatures giving very low infiltration is clearly observed. This is due to the extremely high rate of decrease of permeability. For a higher initial volume fraction (40 percent), we observe a tendency of the front to come to a complete stop (a ''flattening'' of the curves) faster than lower initial volume fractions, which is expected. Moreover, we find that higher the value of $G$, we have increased infiltration. Thus lower heats of reaction and higher heat capacities of the solid reactant facilitate in maximizing infiltration. The faster rate of pore ''closure'' or flattening of the infiltration curves for higher initial volume fractions is attributed to the porosity diminishing sooner.

## 5 Conclusions

A numerical/computational model was developed for reactive melt infiltration incorporating micro processes (heat and mass transfer with chemical reaction and volume expansion) and macro processes (fluid flow through porous preforms). This model predicts permeability as a function of time by an approach wherein six independent equations in six unknowns (from the nonisothermal unreacted-core shrinking model) have to be solved numerically at the micro level. The three-dimensional permeability tensor for hexagonally packed array of reacting cylinders is now written as

$$\tilde{K}=\tilde{K}\{\phi[(\xi_o(\theta))]\}$$

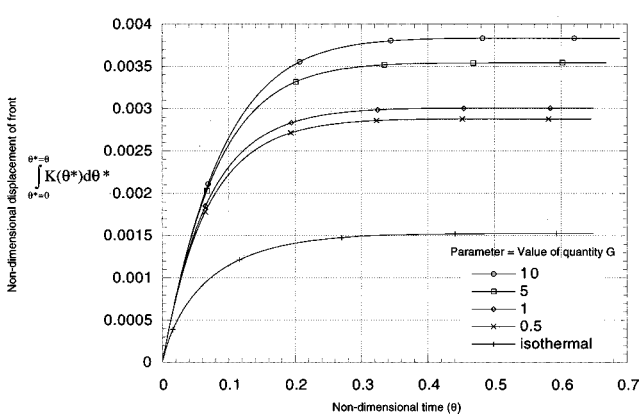



**Fig. 16 Infiltration front displacement for different *G* values (initial volume fraction=40 percent)**

where $\tilde{K}$ is the permeability and $\phi$ the porosity in the RVE written as a function of the growing outer radius of the reaction product layer $\xi_o$, which in turn is a function of time $\theta$ (and also depends on various process parameters). In general for any reacting system of the first order, it is found that low preheat temperatures ensure lower rates of decrease of permeability. Lower heats of reaction and solid reactants with higher heat capacities facilitate lower rates of decrease of permeability (which is advantageous for the silicon/carbon system); in turn this increases the amount of infiltration before choking off occurs. Thus a suitable combination of solid/fluid reactants has to be used to ensure lower heats of reaction. The micro and macro models developed here can be applied to a wide variety of systems just by changing the input parameters and the material properties. By tailoring the initial conditions prior to the onset of infiltration, we can adjust the rate of decrease of permeability to suitably carry out infiltration and arrive at fully dense products (fully dense being the theoretical density of the material with zero porosity).

## References


[1] Hillig, W. B., Mehan, R. L., Morelock, C. R., DeCarlo, V. J., and Laskow, W., 1974, ''Silicon/Silicon Carbide Composites,'' Am. Ceram. Soc. Bull., **54**, No. 12, pp. 1054–1056.

[2] Hillig, W. B., 1985, ''Ceramic Composites by Infiltration,'' Ceram. Eng. Sci. Proc., **6**, pp. 674–683.

[3] Hillig, W. B., 1988, ''Melt Infiltration Approach to Ceramic Matrix Composites,'' J. Am. Ceram. Soc., **71**, pp. C96–C99.

[4] Hillig, W. B., 1990, ''Melt Infiltration Process for Making Ceramic-Matrix Composites,'' *Fiber Reinforced Ceramic Composites*, K. S. Mazdiyasni, ed., Noyes, Park Ridge, NJ, pp. 260–277.

[5] Ness, J. N., and Page, T. F., 1986, ''Microstructural Evolution in Reaction-Bonded Silicon Carbide,'' J. Mater. Sci., **21**, pp. 1377–1397.

[6] Popper, P., 1960, ''The Preparation of Dense Self-Bonded Silicon Carbide,'' *Special Ceramics*, Heywood, London, p. 209.

[7] Forrest, C. W., Kennedy, P., and Shennan, J. V., 1972, ''The Fabrication and Properties of Self-Bonded Silicon Carbide Bodies,'' *Special Ceramics*, Vol. 5, British Ceramic Research Association, Stoke-on-Trent, UK, pp. 99–102.

[8] Tomas Åström, B., Byron Pipes, R., and Advani, S. G., 1992, ''On Flow Through Aligned Fiber Beds and Its Application to Composites Processing,'' J. Compos. Mater., **26**, No. 9, pp. 1351–1373.

[9] Edwards, G. R., and Olson, D. L., 1987, ''The Infiltration Kinetics of Aluminum in Silicon Carbide Compacts,'' Defense Technical Information Center (DTIC) Technical Report, July.

[10] Martins, G. P., Olson, D. L., and Edwards, G. R., 1988, ''Modeling of Infiltration Kinetics for Liquid Metal Processing of Composites,'' Metall. Trans. B, **19B**, Feb., pp. 95–101.

[11] Fitzer, E., and Gadow, R., 1986, ''Fiber Reinforced Silicon Carbide,'' Am. Ceram. Soc. Bull., **65**, No. 2, pp. 326–335.

[12] Fitzer, E., and Gadow, R., 1984, ''Investigations of the Reactivity of Different Carbons With Liquid Silicon,'' *Proceedings of the International Symposium on Ceramic Components for Engines*, Ceramic Society of Japan, Hyakunicho, Japan, pp. 561–572.

[13] Sotirchos, S. V., and Yu, Huei-Chung, 1985, ''Mathematical Modeling of Gas-Solid Reactions With Solid Product,'' Chem. Eng. Sci., **40**, No. 11, pp. 2039–2052.

[14] Bhatia, S. K., and Perlmutter, D. D., 1980, ''A Random Pore Model for Fluid-Solid Reactions: I. Isothermal, Kinetic Control,'' AIChE J., **26**, No. 3, p. 379.

[15] Burganos, V. N., and Sotirchos, S. V., 1987, ''Diffusion in Pore Networks: Effective Medium Theory and Smooth Field Approximation,'' AIChE J., **33**, No. 10, pp. 1678–1689.

[16] Gavalas, G. R., 1980, ''A Random Capillary Model With Application to Char Gasification at Chemically Controlled Rates,'' AIChE J., **26**, No. 4, pp. 577–585.

[17] Chiang, Y.-M., Messner, R. P., Terwilliger, C. D., and Behrendt, D. R., 1991, ''Reaction Formed Silicon Carbide,'' Mater. Sci. Eng., **A144**, pp. 63–74.

[18] Messner, R. P., and Chiang, Y.-M., 1988, ''Processing of Reaction-Bonded Silicon Carbide Without Residual Silicon Phase,'' Ceram. Eng. Sci. Proc., **9**, No. 7–8, pp. 1052–1060.

[19] Brittin, W. E., 1946, ''Liquid Rise in a Capillary Tube,'' J. Appl. Phys., **17**, pp. 37–44.

[20] Gebart, B. R., 1992, ''Permeability of Unidirectional Reinforcements for RTM,'' J. Compos. Mater., **26**, No. 8, p. 1100.

[21] Cai, Zhong, 1992 ''Analysis of Mold Filling in RTM Process,'' J. Compos. Mater., **26**, No. 9, p. 1310.

[22] Cai, Zhong, 1993, ''A Generalized Model for Flow of Polymer Fluids Through Fibrous Media,'' J. Adv. Mater., **25**, No. 1, pp. 58–63.

[23] Chan, A. W., Larive, D. E., and Morgan, R. J., 1993, ''Anisotropic Permeability of Fiber Preforms: Constant Flow Rate Measurement,'' J. Compos. Mater., **27**, No. 10, p. 996.

[24] Messner, R. P., and Chiang, Y.-M., 1990, ''Liquid-Phase Reaction Bonding of

Silicon Carbide Using Alloyed Si-Mo Melts,'' J. Am. Ceram. Soc., pp. 73–1193.
[25] Watson, K. K., 1963, ''The Permeability of an Idealized Two-Dimensional Porous Medium Using the Navier-Stokes Equations,'' Fourth Australia-New Zealand Conference on *Soil Mechanics and Foundation Engineering*, pp. 37–40.
[26] Westhuizen, J. Vd., and Prieur Du Pleiss, J., 1994, ''Quantification of Unidirectional Fiber Bed Permeability,'' J. Compos. Mater., **28**, No. 7, pp. 619–637.
[27] Ishida, M., and Wen, Y., 1965, ''Comparison of Kinetic and Diffusional Models for Solid-Gas Reactions,'' AIChE J., **14**, No. 2, pp. 311–317.
[28] Levenspiel, O., 1972, *Chemical Reaction Engineering*, 2nd Ed., John Wiley and Sons, New York, pp. 349–408.
[29] Froment and Bischoff, *Advanced Chemical Reactor Design*, 1st Ed.,
[30] Fogler, H. S., 1992, *Elements of Chemical Reaction Engineering*, 2nd Ed. (Prentice-Hall International Series in the Physical and Chemical Engineering Sciences), Prentice-Hall, Englewood Cliffs, NJ.
[31] Bird, R. B., Stewart, W. E., and Lightfoot, E. N., 1960, *Transport Phenomena*, John Wiley and Sons, New York.
[32] Yagi, S., and Kunii, D., 1955, *Fifth Symposium (International) on Combustion* Van Nostrand Reinhold, New York, p. 231.
[33] Ausman, J. M., and Watson, C. C., 1962, Chem. Eng. Sci., **17**, p. 323.
[34] Wen, C. Y., and Wang, S. C., 1970, ''Thermal and Diffusional Effects in Noncatalytic Solid Gas Reactions,'' Ind. Eng. Chem., **62**, No. 8, pp. 30–51.
[35] Reint de Boer, 1996, ''Highlights in the Historical Development of the Porous Media Theory: Toward a Consistent Macroscopic Theory,'' Appl. Mech. Rev., **49**, pp. 201–262.
[36] Adler, P. M., 1992, *Porous Media: Geometry and Transports* (Butterworth-Heineman Series in Chemical Engineering), Butterworths, London.
[37] Rege, S. D., and Fogler, H. S., 1989, ''Competition Among Flow, Dissolution, and Precipitation in Porous Media,'' AIChE J., **35**, No. 7, pp. 1177–1185.
[38] Gopalaswamy, R., 1997, Ph.D. thesis, Engineering Science and Mechanics, The Pennsylvania State University, University Park, PA.
[39] Rajesh, G., and Bhagat, R. B., 1998, ''Modeling Micro Level Volume Expansion During Reactive Melt Infiltration Using Non-Isothermal Unreacted-Core Shrinking Models,'' Modell. Simul. Mater. Sci. Eng., **6**, pp. 771–786.
[40] Rajesh, G., and Bhagat, R. B., 1997, ''Micro Modeling of Reactive Melt Infiltration for SiC/SiC Ceramic-Matrix Composites,'' J. Sci. Eng. Compos. Mater., **6**, No. 3, pp. 169–184.
[41] Rajesh, G., and Bhagat, R. B., 1999, ''Infiltration of Liquid Metals in Porous Compacts: Modeling of Permeabilities During Reactive Melt Infiltration,'' Transp. Porous Media, **36**, No. 1, pp. 43–68.
[42] Mazet, N., 1992, ''Modeling of Gas-Solid Reactions. 1. Nonporous Solids,'' Int. Chem. Eng., **32**, No. 2, pp. 271–284.
[43] Mazet, N., and Spinner, B., 1992, ''Modeling of Gas-Solid Reactions. 2. Porous Solids,'' Int. Chem. Eng., **32**, No. 3, pp. 395–408.
[44] Schönberger, E., and Heinrich, J., 1993, ''Differential Thermoanalytical Investigation of the Reaction of Carbon With Silicon to Silicon Carbide,'' Cer. For. Int., **70**, No. 4, pp. 161–164.
[45] Pampuch, R., Walasek, E., and Bialoskorski, J., 1986, ''Reaction Mechanism in Carbon-Liquid Silicon Systems at Elevated Temperatures,'' Ceram. Int., **12**, pp. 99–106.

**Z. Zong**[1]
e-mail: zongzhi@ihpc.nus.edu.sg

**K. Y. Lam**
Mem. ASME

Institute of High Performance Computing,
89C Science Park Drive,
#02-11/12, The Rutherford,
Singapore Science Park I,
Singapore 118261

# Dynamic Plastic Response of a Submarine Oil Pipeline to an Underwater Explosion Bubble

*Dynamic plastic response of a submarine oil pipeline to an underwater explosion bubble is studied in this paper. A detailed fluid-structure analysis is given to find the fluid forces acting on the pipeline. Assuming the pipeline is made from rigid, perfectly plastic material, we obtain the hydroplastic equation governing the dynamic plastic deflection of such a pipeline. The equation is analytically solved. Three examples are given to discuss the features of the dynamic plastic response and the applicability range of the present analysis. The relation between the half pipe length and maximum permanent plastic deflection is also given.* [S0021-8936(00)01004-7]

## Introduction

The present study is motivated by the recent finding of World War II mines near submarine oil pipelines in the Southeast Asia waters ([1]). To estimate the potential damages the mines might result in on the pipelines, a dynamic analysis of such a submarine pipeline subjected to an underwater explosion is the prerequisite for any further actions (like neutralization) to take.

Some of the free spans of the submarine pipelines near the mines are quite long, which are over 10–100 times their diameters, or even longer. They may undergo large lateral deflections and experience high bending stresses when subjected to underwater explosion loading. Therefore, there exists the risk that the pipelines fail due to large bending deflections and high bending stresses. This damage mode, known as global damage, is analyzed in this paper.

An underwater explosion interacts with the surrounding fluid in two different phases. The first is a transient shock wave, which causes a rapid rise in the fluid velocity, and large inertial loading. The peak pressure of this phase is very high, but its duration is extremely short. The second phase in the explosion is a radial pulsation of the gas bubble. Compared with the shock wave in the first phase, this phase is characterized by lower peak pressure and much longer duration. This relation between shock wave and bubble pulse pressures and duration is shown in the sketch of Fig. 1 of a continuous pressure-time record at a point 18 meters from a 135 kg TNT charge fired 15 meters below the water surface. The pulsation is a result of the imbalance of internal explosion product pressure and the external fluid hydrodynamic pressure. In an underwater explosion, we must consider these two types of loads: shock wave and bubble pulsation. Because of their quite different time scales, they are usually separately analyzed. The elastic response of a submarine pipeline to shock loading was analyzed by Zong et al. [1] and the elastic response of a submarine pipeline to bubble pulsation loading was analyzed by Zong and Lam [2].

The elastic deflections of the pipelines induced by the bubble pulsation are quite significant ([1]). It is quite likely for the pipelines to exceed elastic limit and plastically yield. The present paper continues the study of such a pipeline from dynamic elastic deformation to dynamic plastic deformation. We use hydroplasticity to indicate such inelastic fluid-structure interaction problems. To the best of our knowledge, this problem is studied in detail here for the first time, so the purposes of this paper are (1) to provide a simple approximate solution with enough accuracy to the problem so as to save computational efforts greatly and provide a reference for more complicated numerical analysis and (2) to develop some understanding of the features in the dynamic plastic response of a submarine pipeline to a pulsating bubble.

Many papers have been written on bubble dynamics ([3,4]) and some on bubble-structure interaction analysis ([5,6]), among which the simplest model is a Rayleigh bubble. Many papers have also been written on dynamic plastic deformations of structures ([7]), among which a model suitable for the present problem is rigid-plastic simplification. In this paper, the Rayleigh bubble is used to model explosion bubble dynamics, the rigid-plastic material is used to model the pipeline plastic deflections, and a detailed fluid-structure interaction analysis is given to obtain the hydroplastic equation governing the dynamic plastic deformation of a pipeline. This equation is analytically solved. Numerical examples are used to demonstrate the response features and to obtain the relation between the half pipe length and the maximum plastic deflection.

## Problem Formulation

Consider a circular steel pipe covered with a layer of reinforced concrete subjected to a pulsating bubble as shown in Fig. 2. The thickness of the steel pipe is $t_s$ and that of the concrete layer is $t_c$. The densities of steel and concrete are $\rho_s$ and $\rho_c$, respectively. The outer radiuses of the steel pipe and the concrete layer are $R_s$ and $R_c$. The elastic modulus for steel and concrete are $E_s$ and $E_c$, respectively. The inertia moment of the steel pipe is $I_s$ and that of the concrete layer is $I_c$. The pipeline length is $2L$. Because the free span of the pipeline under consideration is much longer than its radius, the pipeline is simplified as an Euler beam fixed at both

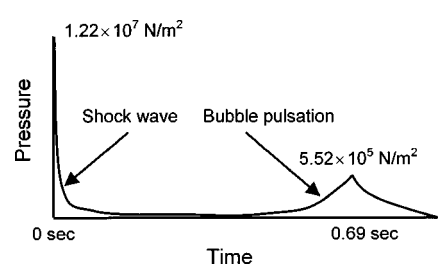



**Fig. 1 Pressure-time record at 18 meters from a 135 kg TNT charge**

[1]To whom correspondence should be addressed.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Sept. 29, 1999; final revision, April 5, 1999. Associated Technical Editor: K. T. Ramish. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

 **Copyright © 2000 by ASME** 

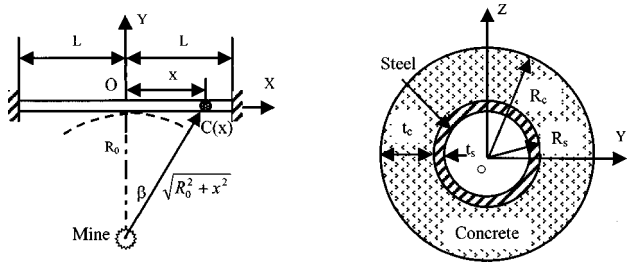



**Fig. 2 A submarine pipeline subjected to an underwater explosion**

ends. Two coordinate systems $Oxyz$ and $O'x'y'z'$ are used as shown in Fig. 3, with the former origin located at the pipe center and the latter origin located at the bubble center. The deflection of the pipeline is described by $W(x,t)$, and the fluid force acting on the pipe per unit length is $f(x,t)$. It is our purpose in the next sections to find both $f(x,t)$ and $W(x,t)$.

It is assumed the fluid inviscid and incompressible. Then there exists a potential $\Phi$ satisfying the Laplacian equation. The second assumption is that the maximum pipe radius is much smaller than the distance from the bubble center to the pipeline center (stand-off distance), that is, $R_c/R_0 \ll 1$, where $R_0$ is the stand-off distance. Then the fluid domain is decomposed into two: one is $D_b$ near the bubble and far away from the pipeline, and the other is $D_p$ near the pipeline and far away from the bubble. We may write $\Phi=\varphi_b+\varphi_p$, where $\varphi_b$ denotes the potential purely produced by the bubble, and $\varphi_p$ denotes all other effects due to the presence of the pipeline. The main disturbance in the fluid is produced by the bubble, from which comes our third assumption: In $D_b$, $\varphi_b$ is much greater than $\varphi_p$, i.e., $\varphi_b \gg \varphi_p$, and in $D_p$, $\varphi_b$ is of same order as $\varphi_p$, i.e., $O(\varphi_b)=O(\varphi_p)$. The solution to $\Phi$ can then be found through solving the two potentials $\varphi_b$ and $\varphi_p$, respectively. This method is somewhat like the matched asymptotic method in asymptotic analysis. The fourth assumption is that the pipeline is slender, i.e., $\varepsilon=R_c/L \ll 1$. By using the fourth assumption, a three-dimensional flow is locally approximated by a two-dimensional flow. The fifth assumption is that the maximum plastic deflection is small, which will be explained in detail later.

**(a) Bubble Dynamics (in $D_b$).** In the domain $D_b$, consider an underwater explosion bubble, which radius and pressure at any time are $\zeta(t)$ and $P(t)$. The initial radius $\zeta_0$ and pressure $P_0$ are given. In this domain, $\Phi \approx \varphi_b(x',y',z';t)$ from our third assumption, and $\varphi_b$ satisfies the Laplacian equation and the boundary conditions on the bubble surface

$$\nabla^2\varphi_b=\frac{\partial^2\varphi_b}{\partial x'^2}+\frac{\partial^2\varphi_b}{\partial y'^2}+\frac{\partial^2\varphi_b}{\partial z'^2}=0, \tag{1}$$

$$\frac{\partial\varphi_b}{\partial t}=-\frac{P_g}{\rho_0}-\frac{1}{2}|\nabla\varphi_b|^2-gd_0 \quad \text{at } r'=\zeta \tag{2}$$

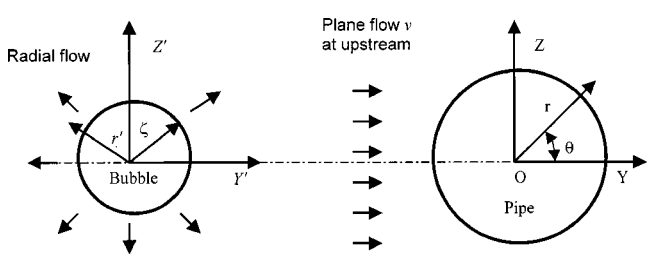



**Fig. 3 Interaction between a bubble and a pipeline**

$$\frac{d\varsigma}{dt}=\frac{\partial\varphi_b}{\partial r'}, \quad \text{at} \quad r'=\zeta \tag{3}$$

$$|\nabla\varphi_b|\to 0, \quad \text{at infinity} \tag{4}$$

where $P_g$ is the pressure inside the bubble, $d_0$ is charge depth, $g$ is gravity acceleration, and $\rho_0$ is water density. Inside the bubble, the gas is assumed ideal, and the pressure is uniform. Then we have

$$\frac{P_g}{P_0}=\left(\frac{4/3\pi\varsigma_0^3}{4/3\pi\zeta^3}\right)^\gamma=\left(\frac{\varsigma_0}{\varsigma}\right)^{3\gamma}, \quad \gamma=1.4. \tag{5}$$

The solution to Eqs. (1)–(4) is a point source with time-dependent strength $q(t)$ located at the center of the bubble of the form

$$\varphi_b=\frac{q(t)}{r'}. \tag{6}$$

Substituting Eqs. (5) and (6) into Eqs. (2) and (3), we obtain

$$\frac{dq}{dt}=-\frac{\zeta}{\rho_0}\left(\frac{\varsigma_0}{\varsigma}\right)^{3\gamma}-\frac{q^2}{2\varsigma^2}-g\zeta d_0, \tag{7}$$

$$\frac{d\zeta}{dt}=-\frac{q}{\zeta^2}, \tag{8}$$

$$q(t)=-\zeta^2\dot{\zeta}. \tag{9}$$

Equations (7) and (8) are nonlinear, and can be numerically integrated using the Runge-Kutta method once the initial conditions are given. Its solutions are two time-dependent functions $q(t)$ and $\varsigma(t)$.

**(b) Fluid Forces on the Pipeline (in $D_p$).** Introducing the following nondimensional quantities in domain $D_p$,

$$\bar{x}=\frac{x}{L}, \quad \bar{y}=\frac{y}{R_c}, \quad \bar{z}=\frac{z}{R_c}$$

we have

$$\frac{\partial\Phi}{\partial x}=\frac{1}{L}\frac{\partial\Phi}{\partial\bar{x}}, \quad \frac{\partial\Phi}{\partial y}=\frac{1}{R_c}\frac{\partial\Phi}{\partial\bar{y}}, \quad \frac{\partial\Phi}{\partial z}=\frac{1}{R_c}\frac{\partial\Phi}{\partial\bar{z}}.$$

Because $\varepsilon=R_c/L \ll 1$ and the derivatives with respect to the nondimensional quantities are of same order, the above equations indicate that

$$\frac{\partial\Phi}{\partial x} \ll \left(\frac{\partial\Phi}{\partial y},\frac{\partial\Phi}{\partial z}\right), \quad \frac{\partial^2\Phi}{\partial x^2} \ll \left(\frac{\partial^2\Phi}{\partial y^2},\frac{\partial^2\Phi}{\partial z^2}\right). \tag{10}$$

Thus, in the fluid domain near the pipeline, the three-dimensional flow is locally approximated by a two-dimensional one as sketched in Fig. 3. Moreover, from the fourth assumption, we conclude that $(y,z) \ll R_0$ in domain $D_p$. Then $\varphi_b$ takes the following asymptotic form in $D_p$,

$$\varphi_b=\frac{q}{\sqrt{(y+R_0)^2+z^2+x^2}}\approx\frac{q}{\sqrt{R_0^2+x^2}}\left(1-\frac{yR_0}{R_0^2+x^2}\right),$$
$$\frac{\partial\varphi_b}{\partial y}\approx-\frac{qR_0}{(R_0^2+x^2)^{3/2}}=v. \tag{11}$$

With such approximations, we obtain

$$\nabla^2\Phi=\frac{\partial^2\Phi}{\partial y^2}+\frac{\partial^2\Phi}{\partial z^2}=0, \quad \text{in } D_p$$

$$\frac{\partial\Phi}{\partial r}=\frac{\partial W}{\partial t}\cos\theta, \quad \text{at } r=R_c, \tag{12}$$

$$\frac{\partial\Phi}{\partial r}=0, \quad \text{at infinity}$$

where $\theta$ is defined in Fig. 3. Because $\varphi_b$ is known, the above equations uniquely determine $\varphi_p$. We may further write $\Phi$ in the form of

$$\Phi=\varphi_b+\varphi_D+\varphi_R,$$

where $\varphi_D$ is the diffraction flow filed induced by the pipeline free of any deformations, satisfying

$$\nabla^2\varphi_D=0,$$

$$\frac{\partial\varphi_D}{\partial r}=-\frac{\partial\varphi_b}{\partial r}=-v\cos\theta,\quad \text{at}\ \ r=R_c, \tag{13}$$

$$|\nabla\varphi_D|\to 0\quad \text{at infinity.}$$

The solution to the above problem is well known of the form

$$\varphi_D=v\frac{R_c^2}{r}\cos\theta. \tag{14}$$

$\varphi_R$ is the radiation potential produced purely by the deflection of the pipeline, satisfying

$$\nabla^2\varphi_R=0,$$

$$\frac{\partial\varphi_R}{\partial r}=\frac{\partial W}{\partial t}\cos\theta,\quad \text{at}\quad r=R_c, \tag{15}$$

$$|\nabla\varphi_R|\to 0,\quad \text{at}\quad \text{infinity.}$$

Similarly, the solution is

$$\varphi_R=-\frac{\partial W(x,t)}{\partial t}\frac{R_c^2}{r}\cos\theta. \tag{16}$$

Because $R_c/R_0\ll 1$, the linearized Bernoulli's equation $P=-\rho_0\partial\Phi/\partial t$ is used to find the pressure on the pipeline. The fluid force per unit length is then obtained by integrating pressure $P$ on the pipeline surface,

$$f(x,t)=\oint_{C(x)}P(x,y,z;t)n_y dl(x)$$
$$=-\rho_0\oint_{C(x)}\left[\frac{\partial\varphi_b}{\partial t}+\frac{\partial\varphi_D}{\partial t}+\frac{\partial\varphi_R}{\partial t}\right]n_y dl(x), \tag{17}$$

where $\int_{C(x)}(\cdot)n_y dl(x)$ is the contour integral at section $C(x)$. Substituting Eqs. (11), (14), and (16) into Eq. (17), we obtain

$$f(x,t)=\pi\rho_0R_c^2\dot{v}+\pi\rho_0R_c^2\dot{v}-\pi\rho_0R_c^2\frac{\partial^2W}{\partial t^2}=2m_a\dot{v}-m_a\frac{\partial^2W}{\partial t^2}, \tag{18}$$

where $m_a=\pi\rho_0R_c^2$ is the added mass of a circular section in water. The first term is the bubble-induced force without diffraction effect considered. The second term is caused by the diffraction effect. They are identical here. In general cases, however, they are not necessarily identical. The third term is radiation force caused by the pipeline motion.

**(c) Fluid-Pipe Interaction Equation.** Consider the pipeline of the section as shown in Fig. 2, with the outer pipe being reinforced concrete and the inter pipe being steel. The equivalent bending stiffness $EI$ and equivalent mass $m_e$ per unit length are given by

$$EI=E_cI_c+E_sI_s,\quad m_e=\pi\rho_s[R_s^2-(R_s-t_s)^2]+\pi\rho_c[R_c^2-R_s^2]. \tag{19}$$

Suppose the pipeline is made from the rigid-plastic material as shown in Fig. 4 and is subjected to a pure moment $M$. The limit cross-sectional moment $M_0$ of the pipeline is obtained by integration

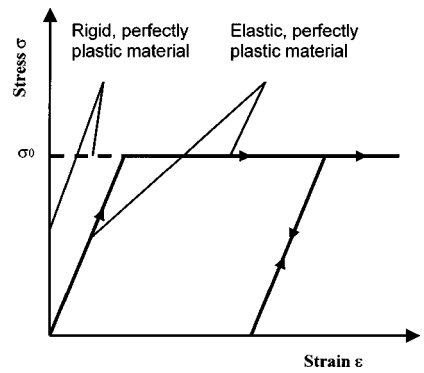



**Fig. 4 Elastic, perfectly plastic and rigid, perfectly plastic uniaxial stress-strain idealizations**

$$M_0=2\sigma_c^0\int_{R_s}^{R_c}r^2dr\int_0^{\pi}\sin\theta d\theta+2\sigma_s^0\int_{R_s-t_s}^{R_s}r^2dr\int_0^{\pi}\sin\theta d\theta$$
$$=\frac{4\sigma_c^0}{3}(R_c^3-R_s^3)+\frac{4\sigma_s^0}{3}[R_s^3-(R_s-t_s)^3] \tag{20}$$

where $\sigma_c^0$ and $\sigma_s^0$ are yield stresses of concrete and steel, respectively. It has been observed that the influence of the lateral shear forces on the magnitude of the plastic limit moment of a cross section may be disregarded for most beam structures ([8]). By assuming that the deflection is small, the dynamic behavior of the pipeline is governed by the equations

$$Q=\frac{\partial M}{\partial x}, \tag{21}$$

$$\frac{\partial Q}{\partial x}=-f(x,t)+m_e\frac{\partial^2W}{\partial t^2} \tag{22}$$

where $Q$ is shear force at a section. Substituting Eq. (21) into Eq. (22) gives

$$\frac{\partial^2M}{\partial x^2}=-f(x,t)+m_e\frac{\partial^2W}{\partial t^2},$$

which, when using Eq. (18), becomes

$$\frac{\partial^2M}{\partial x^2}=(m_a+m_e)\frac{d^2W}{dt^2}-2m_a\dot{v}=(m_a+m_e)\frac{\partial^2W}{\partial t^2}-p(x,t) \tag{23}$$

where

$$p(x,t)=2m_a\dot{v},\quad \dot{v}=\frac{(2\zeta\dot{\zeta}^2+\zeta^2\ddot{\zeta})R_0}{\sqrt{(R_0^2+x^2)^3}} \tag{24}$$

obtained from Eqs. (9) and (11). A uniform dynamic loading $p(x,t)=p(t)$ may cause a plastic deflection if $p(t)>f_c=4M_0/L^2$, where $f_c$ is the limit static loading ([4]). Beyond $f_c$, there are two types of plastic deformations: If $f_c<p(t)\leqslant 3f_c$, there is one stationary plastic hinge in the middle of the span. If $p(t)>3f_c$, the deformation diagram is characterized by two symmetrically located nonstationary plastic hinges. In the portion between the hinges, the bending moment is $M_0$ and the portion constitutes a plastic section. The end sections behave as absolutely rigid parts. The boundary between the rigid and plastic section changes its position during deformation: Bending waves propagate along the pipeline. In this paper, we only consider the former case $f_c<p(t)\leqslant 3f_c$, so the deflection profile is defined by (stationary plastic hinge in the middle)

$$W(x,t)=w(t)(1-x/L),\quad 0\leqslant x\leqslant L. \tag{25}$$

The pipeline is, therefore, idealised as two rigid arms (i.e., $-M_0<M<M_0$) which are connected by the central plastic hinge ($M=M_0$). Substituting Eq. (25) into Eq. (23), Eq. (23) may be integrated spatially,

$$M=(m_a+m_e)\left(\frac{x^2}{2}-\frac{x^3}{6L}\right)\frac{d^2w}{dt^2}$$
$$-2m_a(2\zeta\dot{\zeta}^2+\zeta^2\ddot{\zeta})\sqrt{1+\left(\frac{x}{R_0}\right)^2}+M_0, \qquad (26)$$

where the arbitrary constants of integration have been determined from the requirements that $M=M_0$ and $Q=\partial M/\partial x=0$ at $x=0$. Using $M=-M_0$ at $x=L$, we obtain

$$\frac{d^2w}{dt^2}=\frac{6M_0}{(m_e+m_a)L^2}\left[\frac{m_a}{M_0}(2\zeta\dot{\zeta}^2+\zeta^2\ddot{\zeta})\sqrt{1+\left(\frac{L}{R_0}\right)^2}-1\right] \qquad (27)$$

subject to the initial conditions $w=\dot{w}=0$. This is the equation governing the plastic deflections with time. The pipeline does not commence to plastically deform until $p(x,t)\geqslant f_c$. If Eq. (27) is integrated with respect to time, then

$$\frac{dw(t)}{dt}=\frac{6M_0}{(m_e+m_a)L^2}\left[\frac{m_a}{M_0}(\zeta^2\dot{\zeta})\sqrt{1+\left(\frac{L}{R_0}\right)^2}-t\right] \qquad (28)$$

$$w(t)=\frac{6M_0}{(m_e+m_a)L^2}\left[\frac{m_a}{3M_0}\sqrt{1+\left(\frac{L}{R_0}\right)^2}(\zeta^3-\zeta_0^3)-\frac{t^2}{2}\right]. \qquad (29)$$

The pipeline reaches its permanent position when $\dot{w}=0$.

## Numerical Results and Discussions

As a first example, consider a bubble of initial radius $\varsigma_0=0.75$ m and initial pressure $P_0=4.5\times10^7$ N/m$^2$. The charge depth is 45 m. The results obtained from Eqs. (7) and (8) are shown in Fig. 5 with solid lines, and the results given by Vermon [9] are shown with dashed lines. Both are in good agreement.

To derive the dynamic plastic response Eq. (27), we assume that the deflection is small. To check the applicability of this assumption, we will compare it with more accurate theories, that is, finite-deflection theories developed by Jones [10]. Consider a fully clamped beam which is subjected to uniformly distributed dynamic pressure pulse without fluid-structure interaction taken into account. The pressure pulse is of rectangular shape with respect to time with peak pressure $p_0$ and duration $\tau$. The results obtained from finite-deflection theories and the present method are shown in Fig. 6. In the figure, the lower and upper bounds refer to the bounds on the yield surface, and are obtained using a square yield curve. The parameters in the figure are beam thickness $H$, impact density $I=p_0\tau/(m_eHf_c)$. $M_0$ is the limit cross-sectional moment. From the figure, we observe that when the maximum deflections are of the order of one-half beam thickness or smaller, the permanent deflections predicted by the finite-deflection theories and the small-deflection theory are quite close. Therefore, the small-deflection theory used in this paper is expected to yield good results when the maximum deflections are of the order of one-half pipeline thickness, or when the maximum deflections are of the order of the pipeline radius.

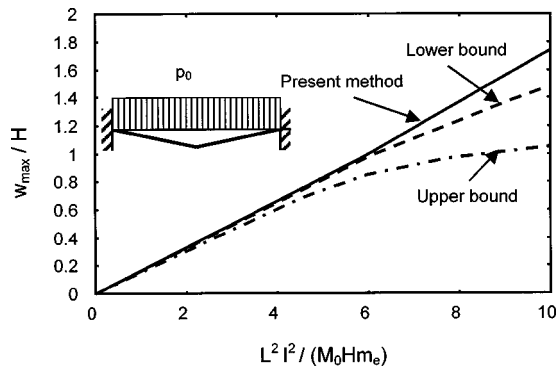



**Fig. 6 Maximum permanent transverse deflection versus a nondimensionalized impulse parameter for a fully clamped beam which is subjected to a uniformly distributed dynamic pressure**

In the third example, a practical case that submarine oil pipeline is subjected to an explosion bubble is considered. Material parameters are: $\sigma_s^0=2.5\times10^8$ N/m$^2$, $t_s=0.011176$ m, $R_s=0.5588$ m, $\rho_s=7800$ kg/m$^3$, $\sigma_c^0=2.0\times10^7$ N/m$^2$, $R_c=0.6985$ m, $\rho_c=2400$ kg/m$^3$. The initial conditions for the bubble are $\zeta_0=0.5$ m and $P_0=1.8\times10^8$ N/m$^2$. Charge depth is 45 m.

The time history of the plastic deflection at the midspan for $R_0=50$ m obtained from Eq. (29) is shown in Fig. 7, with $L=20$ m, 30 m, and 40 m, respectively. There are two observations from this figure. First, short span pipelines reach their permanent positions faster than long span pipelines for the same bubble. For example, the pipeline $L=20$ m reaches its permanent position at $t=0.09$ seconds, and the pipeline $L=40$ m reaches its permanent position at $t=0.15$ seconds. Secondly, the maximum deflections of short pipelines are larger than those of long pipelines for the same bubble. As shown in Fig. 7, the maximum deflection of the pipeline $L=20$ m is 0.226 m, and that of the pipeline $L=40$ m is 0.147 m.

Figure 8 shows the relation between the maximum deflection and half pipe length. Three standoff distances are considered, $R_0=40$, 50, and 70 m, respectively. For each curve in the figure, there is a jump. The limit moment is proportional to $L^2$, and thus a shorter pipeline has higher limit moment. There exists a critical

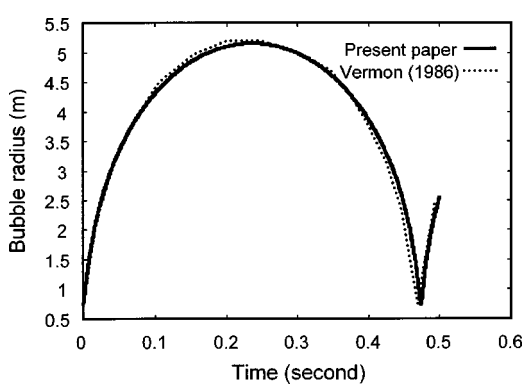



**Fig. 5 Comparison of the bubble radius calculated from present paper and Vermon [6]**

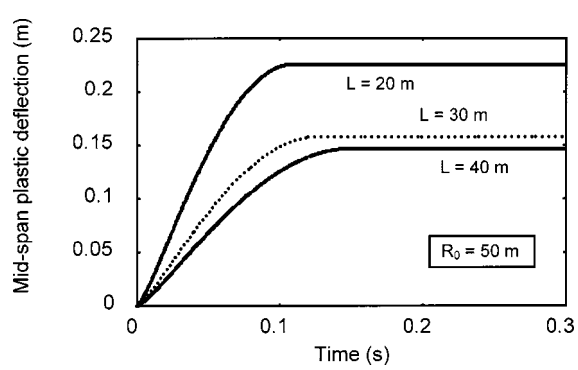



**Fig. 7 Time history of the plastic deflections of midspan for $R_0=50$ m**

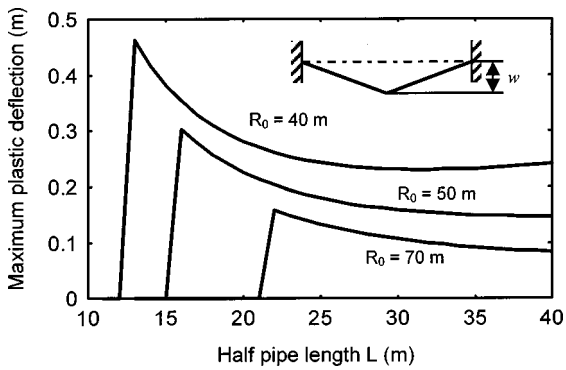



**Fig. 8 Relation between half pipe length and maximum plastic deflection**

half pipeline length, below which the pipeline does not plastically deforms, and above which the pipeline yields with plastic hinge formed in the middle. Once plastic hinge is formed, the deflection is proportional to the reciprocal of the squared half pipe length $L^2$, and thus a shorter pipeline has larger plastic deflections. These two reasons lead to the jump in each curve in Fig. 8. The point at which the jump occurs is the critical half pipeline length.

The above analysis is valid for $f_c < p(t) \leqslant 3f_c$. In all the above computations, this condition is checked and satisfied. The case for $p(t) > 3f_c$ will be considered in our future work.

## Conclusions

In this paper, a detailed fluid-structure interaction analysis was given to obtain the equation governing the dynamic plastic response of a pipeline subjected to the bubble. The equation was analytically solved.

Numerical examples show that there exists a critical half pipe length, below which the pipeline does not deform plastically. Once the pipeline yields, the maximum plastic deflection is proportional to the reverse of squared half pipe length.

The present analysis applies to distant explosion and the maximum plastic deflection is smaller than the pipe radius.

## References


[1] Zong, Z., Lam, K. Y., and Liu, G. R., 1999, ''Probabilistic Risk Prediction of Submarine Pipelines Subjected to Underwater Shock,'' ASME J. Offshore Mech. Arct. Eng., **121**, pp. 251–254.

[2] Zong, Z., and Lam, K. Y., 2000, ''The Flexural Response of a Submarine Pipeline to an Underwater Explosion Bubble,'' ASME J. Offshore Mech. Arct. Eng., **122**, pp. 194–199.

[3] Prosperetti, A., 1993, ''Bubble Dynamics: Some Things We Did Not Know 10 Years Ago,'' *Bubble Dynamics and Interface Phenomena*, J. B. Blake, J. M. Boulton-Stone, and N. H. Thomas, eds., Kluwer, Dordrecht, The Netherlands, pp. 3–16.

[4] Cole, R. H., 1948, *Underwater Explosion*, Princeton University Press, Princeton, NJ, pp. 228–269.

[5] Duncan, J. H., and Milligan, C. D., 1993, ''Numerical Simulation of the Interaction of an Explosion Bubble With a Submerged Spherical Pressure Vessel,'' *Bubble Dynamics and Interface Phenomena*, J. B. Blake, J. M. Boulton-Stone, and N. H. Thomas, eds., Kluwer, Dordrecht, The Netherlands, pp. 437–444.

[6] Hicks, A. N., 1986, ''Explosion Induced Hull Whipping,'' *Advances in Marine Structures*, C. S. Smith and J. D. Clarke, eds., Elsevier, New York, pp. 390–410.

[7] Jones, N., 1989, *Structural Impact*, Cambridge University Press, Cambridge, UK, pp. 1–109.

[8] Hodge, P. G., 1959, *Plastic Analysis of Structures*, McGraw-Hill, New York, pp. 50–100.

[9] Vermon, T. A., 1986, ''Whipping Response of Ship Hulls From Underwater Explosion Bubble Loading,'' Technical Memorandum 86/225, Defense Research Establishment Atlantic, pp. 1–41.

[10] Jones, N., 1971, ''A Theoretical Study of the Dynamic Plastic Behavior of Beams and Plates With Finite Deflections,'' Int. J. Solids Struct., **7**, pp. 1007–1029.

# Energy-Based Stochastic Estimation for Nonlinear Oscillators With Random Excitation

**J. B. Roberts**
**M. Vasta**

School of Engineering,
University of Sussex,
Falmer, Brighton,
East Sussex BN1 9QT, UK

*A physically based averaging procedure is applied to a general form of a nonlinear single-degree-of-freedom equation, with nonwhite random excitation, leading to a one-dimensional continuous Markov model for the energy envelope. It is demonstrated that, in combination with an energy-based technique for estimating the potential energy function, the Markov model can be used as the basis of a stochastic identification method for estimating the spectrum of the excitation, the static nonlinear restoring characteristic, and the nonlinear damping function, from measurements of the response alone. Moreover it is shown that, by combining results for two levels of stochastic excitation, it is possible to obtain good estimates of damping and stiffness parameters, both linear and nonlinear.* [S0021-8936(00)02304-7]

## 1 Introduction

The problem of predicting the dynamic response of flexible structures to random environmental loading occurs in a number of engineering fields. Examples include the wave-induced motion of offshore structures and the vibration of tall buildings in a turbulent wind. Although the general features of the appropriate equations of motion are usually known there are often fairly major uncertainties regarding the magnitude, or even the form, of various terms. This is particularly true in cases of fluid-induced loading where, due to the complexity of the flow processes involved, it is frequently difficult to quantify the damping and excitation terms. Thus there is considerable interest in testing the validity of existing design methods through an analysis of experimental data. Such an analysis can be applied to either full-scale or model test data.

In cases where measurements of both the excitation forces and the corresponding system response are available then it is possible to estimate unknowns in the system equations through an application of deterministic system identification techniques (e.g., see Masri and Caughey [1], Chassiakos et al. [2], and Yasuda et al. [3]). A wide variety of these are now available (e.g., see Unbehauen and Rao [4]). However, in many cases in engineering practice it is impractical, or actually impossible, to measure the excitation process. As an example, in the case of a tower structure responding to turbulent wind excitation, the actual motion of the tower can be measured but the wind excitation forces experienced by the structure are unmeasurable, even in controlled model tests.

To deal with such problems it is necessary to model the excitation as a stochastic process and to derive estimates of the unknown quantities, such as damping, from measurements of the system response alone. If the system is linear this type of stochastic system identification problem can be treated through application of spectral identification methods and the random decrement technique (e.g., see Feng et al. [5]). Thus, for example, if the excitation is wide-band and the damping is light, the resonant peaks in the response spectra provide information on the stiffness and damping properties of the system. However, if the system is nonlinear, with respect to damping, and possibly also stiffness, then the estimation problem is much more difficult.

The principal author, with co-workers, has proposed a number of possible techniques for addressing this nonlinear stochastic identification problem over the last few years (Roberts et al. [6–9] and Battaini and Roberts [10]). In this work the excitation was assumed to be a stationary Gaussian white noise, but more recently (Vasta and Roberts [11] and Roberts and Vasta [12]) a method of dealing with non-Gaussian excitation, with a specified parametric form, has been developed, based on the use of higher-order spectra. All the methods proposed so far have two basic limitations. First, they do not allow the damping term(s) to be estimated directly, when these are nonlinear, with reasonable accuracy. Thus, for example, in previous attempts to apply Markov process theory (Roberts et al. [6,9]) only the ratio of the damping to the excitation could be determined and their separation required the additional estimation of an equivalent linear damping. Secondly, they do not provide a nonparametric means of estimating the spectrum of the excitation: this is important because in many applications even the parametric form of the excitation power spectrum is unknown.

In this paper an alternative approach is proposed which, for the case of a nonlinear oscillator responding to stationary random excitation, overcomes these two limitations. The method relies on the fact that, when the damping is light, the energy envelope of the response is approximately a one-dimensional Markov process, even when the excitation is nonwhite (e.g., see Krenk and Roberts [13]). Estimates of the drift and diffusion coefficients of the energy envelope process are estimated from the response data and compared with the corresponding theoretical expressions for these coefficients. From the diffusion coefficient estimate an estimate of the power spectrum of the excitation is directly obtainable. Hence, using the estimate of the drift coefficient the nonlinear damping can be estimated.

In the first part of the paper the theoretical basis for the Markov model of the energy envelope is derived in a compact and simplified manner. It is then shown how the estimation method proposed here, based on the energy envelope, can be combined with a technique proposed earlier by Roberts et al. [6] for estimating the linear and nonlinear stiffness parameters. Although the latter was validated earlier for while noise excitation only it is also appli-

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, June 23, 1999; final revision, Aug. 14, 2000. Associate Technical Editor: A. A. Ferri. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

 **Copyright © 2000 by ASME** 

cable when the excitation is nonwhite, as will be demonstrated in this paper. The combination of these methods allows a complete determination of all the damping and restoring terms in the equation of motion, both linear and nonlinear, together with an estimation of the input power spectrum, from measurements of the response displacement alone. A method of significantly improving the accuracy of estimation for the nonlinear damping and stiffness parameters is then proposed, based on an analysis of the response process at two different excitation levels.

The authors have discussed an estimation scheme based on the drift and diffusion coefficients in an earlier study, of much more limited scope, specific to ship roll motion in random seas (Roberts and Vasta [14]). In this work it was assumed that the stiffness parameters of the system are known a priori and that the restoring characteristic was of the softening kind. The latter assumption, which is relevant to ship roll, means that the response cannot reach stationary conditions. This imposes a severe restriction on the extent to which the method can be validated through application to simulated data. Moreover, the method of improving the accuracy of the parameter estimates by using response data at two different levels, referred to above, was not considered.

To validate the proposed methods they are applied to some simulated data relating to a Duffing oscillator with linear-plus quadratic damping and nonwhite random excitation.

## 2 The Equations of Motion

An equation of motion of a single-degree-of-freedom system, of the following general form

$$\ddot{x}+\varepsilon^2\dot{x}h(x,\dot{x})+g(x)=\varepsilon y(t)=z(t) \tag{1}$$

is considered here. Here $\dot{x}h(x,\dot{x})$ is the nonlinear damping force and $g(x)$ is a nonlinear restoring force. $h(x,\dot{x})$ is assumed to be a positive function of $x$ and $\dot{x}$ while $g(x)$ is taken to be an antisymmetric function of $x$, of the nonsoftening kind.

The excitation $y(t)$ will here be modeled as a stationary random process, with zero mean and a power spectrum $S_y(\omega)$, defined by the relation

$$S_y(\omega)=\frac{1}{2\pi}\int_{-\infty}^{\infty}w_y(\tau)\cos\omega\tau d\tau \tag{2}$$

where

$$w_y(\tau)=E\{y(t)y(t+\tau)\} \tag{3}$$

is the covariance function of the excitation and $E\{\}$ is the expectation, or ensemble averaging operator.

The scaling parameter $\varepsilon$ is introduced into Eq. (1) to help clarify the order of magnitude of the damping and excitation terms in the subsequent analysis. It will be assumed throughout that the damping is small: Hence $\varepsilon$ is small. The scaling of the excitation given in Eq. (1) ensures that the variance of the stationary response remains finite as $\varepsilon\to 0$: i.e., is of order $\varepsilon^0$, with respect to $\varepsilon$ (see Roberts and Spanos [15]). There is no loss of generality implied by scaling the excitation in this way: It simply reflects the fact that, as the response reaches stationarity, the contribution of the damping and excitation forces becomes relatively weaker during a typical cycle in the response, compared to the total energy (kinetic plus potential) in that cycle.

The total energy envelope process $E(t)$ associated with the response is as the sum of kinetic and potential energies: Thus

$$E(t)=\frac{\dot{x}^2}{2}+V(x) \tag{4}$$

where

$$V(x)=\int_0^x g(\xi)d\xi. \tag{5}$$

It is possible to rewrite the equation of motion in terms of $E(t)$ and an associated phase process, $\Phi$, defined by the relationships

$$\mathrm{sgn}(x)\sqrt{V(x)}=\sqrt{E}\cos\Phi, \tag{6}$$

$$\dot{x}=-\sqrt{2E}\sin\Phi \tag{7}$$

The result is two first-order equations, as follows:

$$\begin{bmatrix}\dot{E}\\ \dot{\Phi}\end{bmatrix}=\begin{bmatrix}-\varepsilon^2\alpha_1\\ -\varepsilon^2\alpha_2+\gamma\end{bmatrix}+\begin{bmatrix}\beta_1\\ \beta_2\end{bmatrix}\varepsilon y(t) \tag{8}$$

where

$$\alpha_1=2Eh\sin^2\Phi, \tag{9}$$

$$\alpha_2=-h\sin\Phi\cos\Phi, \tag{10}$$

$$\gamma=\frac{|g(x)|}{V(x)} \tag{11}$$

$$\beta_1=-\sqrt{2E}\sin\Phi, \tag{12}$$

$$\beta_2=-\frac{\cos\Phi}{\sqrt{2E}}. \tag{13}$$

## 3 Markov Modeling

If the damping is light then it is possible to approximate the energy envelope process $E(t)$ as a one-dimensional, continuous Markov process. In the theory leading to this approximation, as developed earlier ([16]), use was made of the Stratonovitch-Khasminskii (SK) limit theorem ([15]). Recently an alternative approach to the derivation of a Markov model has been developed ([17]), based on physical reasoning, which gives much greater insight into the nature of the approximations involved. This analysis shows that a small correction is needed to the result obtained by the SK limit theorem: However, a comparison with simulation results has revealed that this correction is negligible in nearly all circumstances ([17]). Here a brief and simplified account of this physically based approach is given.

**3.1 The Markov Process Approximation.** A fundamental assumption, in developing a Markov model, is that it is possible to find an interval of time $\Delta t$ such that $\Delta t>\tau_{\mathrm{cor}}$, where $\tau_{\mathrm{cor}}$ is the correlation time scale of the excitation, but small enough that the change of energy $\Delta E=E(t+\Delta t)-E(t)$, in that interval, is relatively small. This requirement can be met if the damping is sufficiently light: As the bandwidth of the excitation increases $\tau_{\mathrm{cor}}$ becomes smaller, and the restriction on the damping level reduces.

If the coefficients defined by

$$H_n(E)=\frac{1}{\Delta t}E\{\Delta E^n\}\quad n=1,2,\ldots \tag{14}$$

are such that $H_n(E)=0$, for $n>2$, and both

$$m(E)\equiv H_1(E), \tag{15}$$

$$D(E)\equiv H_2(E) \tag{16}$$

approach a limiting value, for small $\Delta t$ (subject to $\Delta t>\tau_{\mathrm{cor}}$), then $E(t)$ can be modeled, approximately, as a Markov process (e.g., see Gardiner [18]). The coefficients $m(E)$ and $D(E)$ are known, respectively, as the drift and diffusion coefficients.

This Markov process is governed by the following stochastic Itô equation

$$dE=mdt+D^{1/2}dW \tag{17}$$

where $W$ is a unit Wiener (or Brownian) process.

**3.2 Averaging the Dissipation Terms.** If $s$ is some time in the interval $t$ to $t+\Delta t$ then one can write

$$E(s)=E_0+e(s) \tag{18}$$

$$\Phi(s)=\Phi_0(s)+\theta(s) \tag{19}$$

where $E_0$, $\Phi_0$ are the free undamped solutions ($\varepsilon=0$). From Eq. (8)

$$E_0=a \quad \text{constant} \tag{20}$$

and

$$\Phi_0=\int \gamma_0 dt \tag{21}$$

where $\gamma_0$ is the variation of $\gamma$ in the case of free undamped motion. As a first step in developing the Markov model the dissipation terms, $\alpha_1$ and $\alpha_2$ in Eqs. (8), can be averaged over the period of free undamped oscillation, $T(E)$, treating $E$ as a constant during this period and setting $\Phi=\Phi_0$. The equations of motion then become

$$\begin{bmatrix}\dot{E}\\ \dot{\Phi}\end{bmatrix}=\begin{bmatrix}-\varepsilon^2\Lambda_1\\ -\varepsilon^2\Lambda_2+\gamma\end{bmatrix}+\begin{bmatrix}\beta_1\\ \beta_2\end{bmatrix}\varepsilon y(t) \tag{22}$$

where

$$\Lambda_1(E)=\frac{2E}{T(E)}\oint h(E,\Phi_0)\sin^2\Phi_0 dt, \tag{23}$$

$$\Lambda_2(E)=-\frac{1}{T(E)}\oint h(E,\Phi_0)\sin\Phi_0\cos\Phi_0\, dt. \tag{24}$$

In the subsequent analysis only the first of these averaged damping functions $\Lambda_1(E)$ is required. This function can be expressed as

$$\Lambda_1(E)=\frac{2E}{T(E)}\sum_{n=1,3,\ldots}^{\infty} s_n^2\oint h(E,\Phi_0)\sin^2 n\omega(E)t dt \tag{25}$$

using a Fourier expansion of $\sin\Phi_0$, as follows:

$$\sin\Phi_0(t)=\sum_{n=1,3,\ldots}^{\infty} s_n\sin n\omega(E)t. \tag{26}$$

The period of free undamped oscillation is given by

$$T(E)=\frac{2\pi}{\omega(E)}=2\sqrt{2}\int_0^b\frac{d\xi}{\sqrt{[E-V(\xi)]}} \tag{27}$$

where $b$ is such that $V(b)=E$ and $\omega(E)$ is the corresponding frequency.

**3.3 Evaluation of the Drift and Diffusion Coefficients.** On integrating the first row of Eq. (22), over the interval $t$ to $t+\Delta t$, and using the definition of the drift coefficient given by Eqs. (14) and (15), one obtains the expression

$$m(E)=-\varepsilon^2\Lambda_1(E)+\frac{\varepsilon}{\Delta t}\int_t^{t+\Delta t}E\{\beta_1 y\}du. \tag{28}$$

The integral term is nonzero due to the correlation between $\beta_1$ and $x$.

To evaluate the integral, $\beta_1$ can be expanded about its deterministic evolution, corresponding to $\varepsilon=0$, denoted $\beta_{1,0}=\beta(E_0,\Phi_0)$. Thus, for $t<u<t+\Delta t$,

$$\beta_1(u)=\beta_{1,0}(u)+\left(\frac{\partial\beta_1}{\partial E}\right)_0^u e(u)+\left(\frac{\partial\beta_1}{\partial\Phi}\right)_0^u\theta(u). \tag{29}$$

Here the notation $(\,)_0^u$ means that the quantity in the bracket relates to free undamped oscillation and is evaluated at time $u$. The terms of higher order than unity, with respect to $e$ and $\theta$, have neglected. The expansion is thus correct to order $\varepsilon$. On substituting this expansion into Eq. (28) and using the fact that the contribution of the first term is zero (since $E\{y(t)\}=0$), one obtains

$$m(E)=-\varepsilon^2\Lambda_1(E)+\frac{\varepsilon}{\Delta t}\int_t^{t+\Delta t}\left(\frac{\partial\beta_1}{\partial E}\right)_0^u E\{e(u)y(u)\}du$$
$$+\frac{\varepsilon}{\Delta t}\int_t^{t+\Delta t}\left(\frac{\partial\beta_1}{\partial\Phi}\right)_0^u E\{\theta(u)y(u)\}du. \tag{30}$$

An expression for $e(u)$ can be obtained by integrating the energy equation (first row of Eq. (22)). To evaluate the second integral term it is necessary to obtain an expression for $\theta(u)$. As a first step the term $\gamma$ (see Eq. (11)) can be expanded about its value when excitation and damping is absent in a similar way to the method used for $\beta_1$. Thus

$$\gamma(t)=\gamma_0(t)+\left(\frac{\partial\gamma}{\partial E}\right)_0^t e(t)+\left(\frac{\partial\gamma}{\partial\Phi}\right)_0^t\theta(t). \tag{31}$$

Again this expansion is correct to order $\varepsilon$. A detailed examination has shown ([17]) that the terms second and third terms on the right-hand side have a negligible contribution. Hence on simplifying and integrating the phase equation to obtain an expression for $\theta(u)$, and using the expression for one $e(u)$, one finds from equation (30) that correct to order $\varepsilon^2$,

$$m(E)=-\varepsilon^2\Lambda_1(E)+\frac{\varepsilon^2}{\Delta t}\int_t^{t+\Delta t}\int_t^u[\sin\Phi_0(u)\sin\Phi_0(v)$$
$$+\cos\Phi_0(u)\cos\Phi_0(v)]w_y(u-v)dv du$$
$$=-d(E)+\frac{\pi}{2}\sum_{n=1,3,\ldots}^{\infty}(s_n^2+c_n^2)S_z[n\omega(E)] \tag{32}$$

where $c_n$ are Fourier coefficients in the expansion of $\cos\Phi_0(t)$ (c.f. Eq. (25)) and the last step relies on the assumption that $\Delta t>\tau_{\text{cor}}$. In Eq. (32) $d(E)=\varepsilon^2\Lambda_1(E)$ is the average of the actual damping force and $S_z(\omega)=\varepsilon^2 S_y(\omega)$ is the power spectrum of $z(t)=\varepsilon y(t)$, the actual excitation process.

In the special case where the excitation can be modeled as a white noise then this result reduces to a result has been obtained earlier by Stratonovitch [19] using a different argument.

The diffusion coefficient is much easier to calculate than the drift coefficient. From Eq. (26) one finds immediately that, correct to order $\varepsilon^2$,

$$D(E)=\frac{\varepsilon^2}{\Delta t}\int_t^{t+\Delta t}\int_t^{t+\Delta t}\beta_{1,0}(u)\beta_{1,0}(v)w_y(u-v)du dv$$
$$=\frac{\varepsilon^2 2E}{\Delta t}\int_t^{t+\Delta t}\int_t^{t+\Delta t}\sin\Phi_0(u)\sin\Phi_0(v)w_y(u-v)du dv$$
$$=2\pi E\sum_{n=1,3,\ldots}^{\infty}s_n^2 S_z[n\omega(E)]. \tag{33}$$

Again, the last step here follows from the assumption that $\Delta t>\tau_{\text{cor}}$. In the particular case of white noise excitation, Eq. (33) reduces to a result obtained earlier by Stratonovitch [19].

Numerical studies have shown that, in nearly all cases, the first terms in the summations in Eqs. (25), (32), and (33) are completely dominant. Hence, subsequently in this paper, the summations are assumed to be truncated at $n=1$ for computational purposes.

## 4 Stochastic System Identification

The theory given earlier in this paper can also be used for stochastic estimation purposes. The stochastic identification problem of concern here can be stated as follows: From a single measured time history of the displacement response, $x(t_i)\equiv x(i\delta t)$ $(i=1,2,\ldots N)$, where $\delta t$ is the sampling interval, together with a stationary stochastic process model of the excitation, how can one

generate estimates of the static restoring characteristic, the damping function, and the power spectrum of the excitation?

In this paper a new method of completely solving this stochastic estimation problem is proposed. It is partly based on the theory developed in Section 2 and, in particular, on the approximate theoretical expressions obtained there for the drift coefficient (see Eq. (32)) and the diffusion coefficient (see Eq. (33)). Both these coefficients can be estimated directly by suitably processing the measured response data $x(t_i)$ using the definitions given by Eqs. (14) to (16). Prior to application of this method, however, it is necessary to estimate the potential energy function $V(x)$ and the static restoring function $g(x)$ from the displacement response data.

**4.1 Estimation of the Potential Energy Function and the Static Restoring Characteristic.** For light damping the displacement response is oscillatory, with well-defined peaks. At the zero-crossing times, denoted $t_{zi}(i=1,2,\ldots)$, the total energy $E$ is equal to $\dot{x}^2(t_{zi})/2$. Hence estimates of $E$, at these times, can be obtained from estimates of the response velocity. By fitting a curve through these estimates it is possible to estimate $E$ at the intermediate times, $t_{pi}(i=1,2,\ldots)$ where the sample function has maxima or minima. At the peak values the response velocity is zero: Hence $E(t_{pi})=V[x(t_{pi})]$ and the interpolated values are actually estimates of $V$. These estimates may be plotted against the corresponding peak amplitudes, $x(t_{pi})$, the number of such estimates equalling the number of peaks in the sample function: Thus an estimate of the potential energy function is obtainable.

The restoring function $g(x)$ can be derived from the estimated potential energy function using $g(x)=dV/dx$. However, a more robust way of estimating $g(x)$ is to base the analysis on a linear-in-the-parameters representation of $g(x)$ and $V(x)$: e.g.,

$$g(x)=\sum_{i=1}^{M} \alpha_i G_i(x), \tag{34}$$

$$V(x)=\sum_{i=1}^{M} \alpha_i W_i(x), \tag{35}$$

where

$$W_i(x)=\int_0^x G_i(\xi)d\xi \tag{36}$$

and $\alpha_i$ are stiffness parameters. Thus, the stiffness parameters, and hence $g(x)$, can be estimated by fitting the above parametric form for $V(x)$ to the estimate of this function. This can be done using simple least-square fitting (linear regression).

**4.2 Estimation of the Damping Function and the Input Power Spectrum.** The displacement response data can be converted to energy samples $E(t_i)$ using the definition of $E$ given by Eqs. (4) and (5), and the estimated potential energy function found by using the method outlined in the previous section.

Suppose that an interval of time $\Delta t=m\delta t$ is chosen and the increments

$$\Delta E_i=E(t_i)-E(t_{i-m}) \quad i=m+1,\ldots,N \tag{37}$$

are computed. Then the energy range is divided up into a number of contiguous ''slots,'' each of width $\delta E$ such that for the $j$th slot covers the energy range $(j-1)\delta E$ to $j\delta E(j=1,2,\ldots)$. Each $\Delta E_i$ is assigned to the slot for which $j\delta E<E(t_{i-m})<(j+1)\delta E$. Estimates of the drift and diffusion coefficients may now be generated by averaging the $\Delta E_i$ values, and their square, within each slot: Thus

$$\hat{m}(E_j)=\frac{1}{n_j\Delta t}\sum_{(j-1)\delta E<E(t_{i-m})<j\delta E}\Delta E_i \tag{38}$$

$$\hat{D}(E_j)=\frac{1}{n_j\Delta t}\sum_{(j-1)\delta E<E(t_{i-m})<j\delta E}(\Delta E_i)^2 \tag{39}$$

where $E_j$ is the mid value of $E$, in the $j$th slot and $n_j$ is the number of $\Delta E_i$ values obtained for $j$th slot.

For identification purposes a correspondence between these estimates and the theoretical values, given by Eqs. (32) and (33), is required: This will clearly depend on the choice of $\Delta t$. If $\Delta t$ is too small then $\Delta t<\tau_{\text{cor}}$, contravening one of the basic assumptions in the theory. On the other hand if $\Delta t$ is too large then the $\Delta E_i$ will become large, whereas the theory assumes small changes in energy level over the interval $\Delta t$. For lightly damped systems, at least, one can expect the estimates of the coefficients, given by Eqs. (38) and (39), to be almost independent of $\Delta t$ over a significant ''stable'' range of $\Delta t$ values and to correspond well with the theoretical expressions within this range. The correlation time scale of the excitation $\tau_{\text{cor}}$ will not be known in practice but one can evaluate the drift and diffusion coefficients for a range of $\Delta t$ and choose a $\Delta t$ value within the stable range.

Using Eq. (33) it is possible to generate an estimate of the power spectrum of the excitation process $y(t)$ directly from $\hat{D}(E_j)$. Thus

$$\hat{S}_y(\omega_j)=\frac{\hat{D}(E_j)}{2\pi E_j s_1^2} \tag{40}$$

where

$$\omega_j=\omega(E_j) \tag{41}$$

To evaluate $s_1^2$ a knowledge of the restoring characteristic is required. This can be obtained directly from the data, using the method described in Section 4.1.

Once $\hat{S}_y(\omega_j)$ is determined then the averaged damping function can be estimated by using Eq. (32): Thus

$$\hat{d}(E_j)=-\hat{m}(E_j)+\frac{\pi}{2}(s_1^2+c_1^2)\hat{S}_y(\omega_j). \tag{42}$$

To evaluate the Fourier coefficients in this expression a knowledge of the restoring characteristic is again needed, in general.

**4.3 Damping Parameter Estimation.** If a linear-in-the-parameters form of the damping function $\varepsilon^2\dot{x}h(x,\dot{x})$ in Eq. (1) is assumed, i.e.,

$$\varepsilon^2\dot{x}h(x,\dot{x})=\sum_{i=1}^{p}\beta_i\dot{x}h_i(x,\dot{x}) \tag{43}$$

where $\beta_i$ are damping parameters, then from Eq. (25)

$$d(E)=\sum_{i=1}^{P}\beta_i H_i(E) \tag{44}$$

where

$$H_i(E)=\frac{2E}{T(E)}\oint h_i(E,\Phi_0)\sin^2\Phi_0 dt. \tag{45}$$

The damping parameters can be estimated by fitting the parametric form for $d(E)$ to the estimate of this function, obtained from Eq. (42), using simple linear regression.

## 5 An Example

To illustrate the method, and to validate it, through application to some simulation data, a particular form of Eq. (1) is chosen where

$$\ddot{x}+A\dot{x}+B\dot{x}|\dot{x}|+k_1x+k_2x^3=Z(t). \tag{46}$$

This represents a Duffing oscillator with nonlinear damping of the linear-plus quadratic type. On nondimensionalizing time with re-

spect to the linear undamped natural frequency, $\omega_0=\sqrt{k_1}$, and the response displacement with respect to $X=\sqrt{k_1/k_2}$, Eq. (46) can be recast as

$$\ddot{x}+a\dot{x}+b\dot{x}|\dot{x}|+x+x^3=z(t). \tag{47}$$

**5.1 Theoretical Calculations.** As an initial step in the application of the theory it is useful to obtain a solution to the equation of motion for free undamped oscillation (i.e., Eq. (47) with $a=b=z(t)=0$). This solution can be expressed as

$$x(t)=x_0 cn(\tilde{\omega}t|m) \tag{48}$$

where $cn$ is the Jacobian elliptic function, $x_0$ is the amplitude of oscillation, and

$$\tilde{\omega}=(1+x_0^2)^{1/2}, \tag{49}$$

$$m=\frac{x_0^2}{2(1+x_0^2)}. \tag{50}$$

The period of oscillation is given

$$T(E)\tilde{\omega}=4K(m) \tag{51}$$

where $K(m)$ is the complete elliptic integral.

To evaluate the drift and diffusion coefficients, according to Eqs. (32) and (33) it is necessary to find expressions for $\sin\Phi_0$ and $\cos\Phi_0$. The former is given by

$$\sin\Phi_0(t)=\frac{1}{(1-m)^{1/2}}sn(\tilde{\omega}t|m)dn(\tilde{\omega}t|m) \tag{52}$$

where $sn$ and $dn$ are further Jacobian functions. Hence one can obtain $\cos\Phi_0$. As $m\rightarrow 0$ $\sin\Phi_0\rightarrow\sin t$ and $\cos\Phi_0\rightarrow\cos t$, which is the linear result. With increasing $m$ deviations from the harmonic motion become more pronounced with a maximum distortion at $m=0.5$ (the highest possible value for $m$).

It is also necessary to calculate the coefficients $s_1$ and $c_1$ which relate, respectively, to the Fourier expansions of $\sin\Phi_0$ and $\cos\Phi_0$. An expression for $s_1$ can be found analytically: The result is

$$s_1=\frac{\pi^2 q^{1/2}}{K^2(m)[m(1-m)]^{1/2}(1+q)}, \tag{53}$$

where

$$q=\exp\left\{-\frac{\pi K(1-m)}{K(m)}\right\}. \tag{54}$$

It does not seem possible to obtain an analytical expression for $c_1$: However, this coefficient is easy to compute numerically.

The damping term, $d(E)$, is also required in the calculation of the drift coefficient. This can be evaluated numerically from Eq. (25). If the Fourier expansion is truncated at the first term one obtains

$$d(E)=as_1^2E+1.20bs_1^2E^{3/2}. \tag{55}$$

**5.2 Simulation Method.** For the purpose of testing the estimation method it is convenient to choose a band-limited excitation process, having a power spectrum with a single, well-defined peak. A suitable process can be obtained by filtering white noise, $n(t)$, through two identical first-order linear filters, placed in series. The output, $\eta(t)$, is obtained from

$$\dot{\eta}+\beta\eta=\beta\lambda(t), \tag{56}$$

$$\dot{\lambda}+\beta\lambda=\beta n(t). \tag{57}$$

If two such filtering operation are applied to independent white noises the resulting outputs, $\eta_1(t)$ and $\eta_2(t)$, may be combined to form the input process

$$y(t)=h|\eta_1(t)\cos\omega_p t+\eta_2(t)\sin\omega_p t|. \tag{58}$$

The power spectrum of $y(t)$, so defined, is given by

$$S_y(\omega)=K\left[\left\{\frac{1}{\beta^2+(\omega-\omega_p)^2}\right\}^2+\left\{\frac{1}{\beta^2+(\omega+\omega_p)^2}\right\}^2\right] \tag{59}$$

where $K$ is a scaling constant. For small $\beta$ this spectrum has a single peak in the neighborhood of the frequency $\omega_p$. The parameter $\beta$ controls the bandwidth of the excitation while $\omega_p$ determines the position of the peak, relative to the linear undamped natural frequency ($\omega=1$).

The equation of motion was solved numerically, using the fourth-order Runge Kutta algorithm, to generate sample functions of the displacement response.

**5.3 Simulation Results: Single Excitation Level.** A dynamical system governed by Eq. (46), for which both the stiffness and damping terms were nonlinear ($A=0.01$, $B=0.01$ and $k_1=1$, $k_2=1$), was chosen for study. A number of sample functions of the response $x(t_i)$ (referred to here as blocks of data) were computed where, for each block, $t_i=i\,\delta t$, $i=1,2,\ldots,100{,}000$ and $\delta t=0.05$. Thus each block was of length 5000 and contained about 1000 cycles. The input spectrum parameters were chosen to be $\beta=1$ and $\omega_p=2$.

The first stage in the estimation process is to obtain estimates of the stiffness coefficients, $k_1$ and $k_2$, using the method described in Section 4.1. Figure 1(*a*) shows the variation of estimated energy level with one half of the square of the peak amplitudes ($x^2(t_{pi})/2$), obtained from a single block of data. For comparison the corresponding theoretical variation, calculated from the exact values of the parameters ($k_1=k_2=1$), is also shown. The nonlinear behavior in stiffness is well represented at high energy levels with a consistent departure from the linear variation which would occur in the case of linear stiffness only.

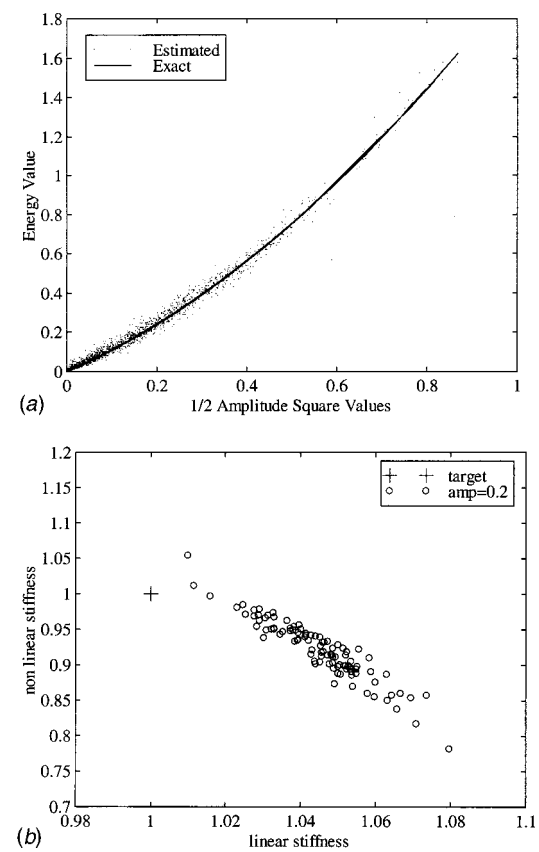



**Fig. 1 (*a*) Comparison of estimated energy values with the exact values, (*b*) scatter plot of the estimates of the linear and nonlinear stiffness parameters**

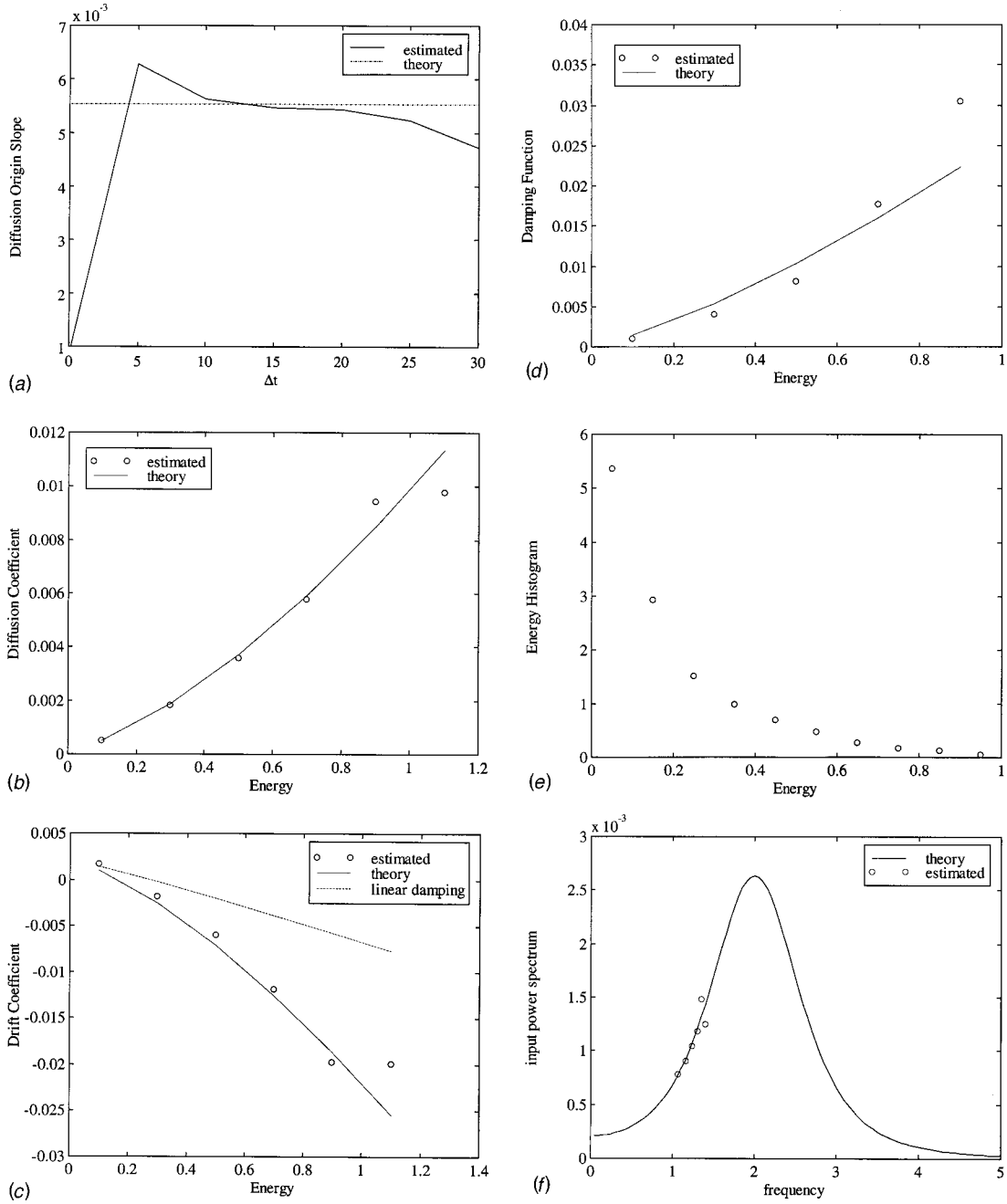



**Fig. 2** (*a*) **Variation of the slope of the diffusion coefficient with $\Delta t$,** (*b*) **comparison between the estimated and true variation of the diffusion coefficient with energy level,** (*c*) **comparison between the estimated variation and true variation of the drift coefficient with energy level,** (*d*) **comparison between the estimated and true damping function variation with energy level,** (*e*) **estimated histogram of the energy,** (*f*) **comparison between estimates of the input power spectrum and the true spectrum,** (*g*) **variation with frequency of the standard deviation of the estimates of the input power spectrum,** (*h*) **variation of the frequency of free undamped oscillation with energy level,** (*i*) **scatter plot for the estimates of the linear and nonlinear damping parameters**

Stiffness parameter estimates are easily obtained by fitting the parametric form for the potential energy (see Eqs. (34) and (35)) to estimated energy values. Figure 1(*b*) shows a plot of pairs of stiffness estimates ($\hat{k}_1$ versus $\hat{k}_2$) obtained from 100 blocks of data, for a value of the input amplitude equal to $K=0.2$ with $K$ as defined in Eq. (59). The scatter in the estimates is relatively small but there is a clear negative correlation and a close to linear dependency (see Section 5.4 for a further discussion of this feature). Averaging these estimates gives the estimates $\hat{k}_1=1.04$, $\hat{k}_2=0.95$: thus the error is about five percent.

Using the estimated stiffness coefficients the data was rescaled, as indicated in Eq. (47), prior to implementation of the estimation method based on the drift and diffusion coefficients. It is straightforward to rescale estimates obtained in this way so that they relate to the original equation of motion (Eq. (46)).

To apply the proposed estimation technique the range of $\Delta t$ values for which the estimates of the drift and diffusion coefficients, as obtained from Eqs. (38) and (39), are stable must be established. The diffusion coefficient estimate, $\hat{D}(E)$, was found to be the most sensitive to the choice of $\Delta t$. For small values of

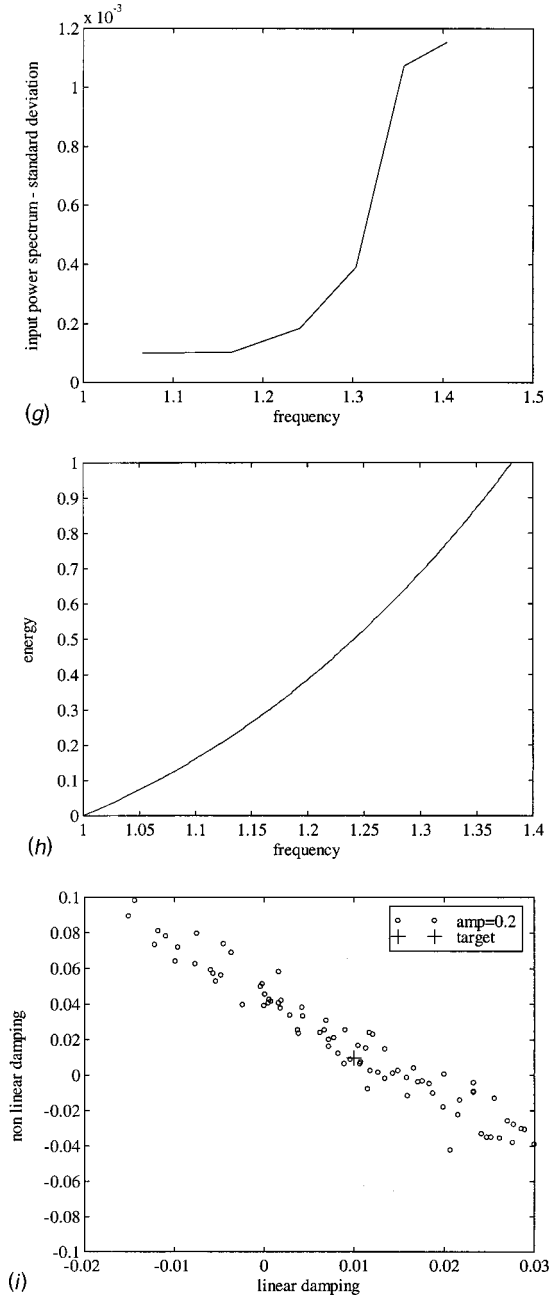



**Fig. 2** (***Continued***).

$E$, $\hat{D}(E)$ becomes proportional to $E$ since the restoring moment is nearly linear at low amplitudes: Thus $s_1 \rightarrow 1$ and $S_y[\omega(E)] \rightarrow S_y(1)$ as $E \rightarrow 0$ (see Eq. (47)). Figure 2($a$) shows the variation of the slope, $\hat{D}(E)/E$, at low $E$ values, with $\Delta t$ and compares this with the theoretical value of $2\pi S_y(1)$. As expected from the theory there is a range of $\Delta t$ values ($5 < \Delta t < 20$) over which the slope is reasonably stable. On the basis of this result a value of $\Delta t$ of 10 was chosen for calculating the drift and diffusion coefficient estimates.

Figures 2($b$), 2($c$) and 2($d$) show, respectively, the variation of the estimates of the diffusion and drift coefficients and damping function, $d(E)$, with energy level. These were obtained by averaging 100 blocks of data. The corresponding histogram of the energy values is shown in Fig. 2($e$). The histogram indicates that $E < 1$ for most of the data. There is considerable scatter in the estimates of the two coefficients but at low amplitudes, where there is most data, the agreement between the estimates and the theory is very good.

Figure 2($f$) shows estimates of the input power spectrum, derived from the estimates of the diffusion coefficient, according to Eq. (40) (again averaging the results from 100 blocks). This shows that the theory correctly accounts for the shape of the input spectrum. Figure 2($g$) shows the standard deviation of the spectral estimates, over the energy range of interest. The standard deviation is low, at low frequencies but increases significantly at higher frequencies. This is a reflection of the fact that as the frequency increases the energy level increases and the amount of data available to form the estimates decreases. Thus, a comparison of the energy histogram with a plot of the variation of the free oscillation frequency, $\omega(E)$, with $E$ (see Fig. 2($h$)), shows that reliable estimates of the spectrum can only be achieved over a fairly limited frequency range (say $1 < \omega < 1.4$).

It is noted that, for a single set of data, the segment of the excitation spectrum which can be estimated is limited to the frequency range corresponding to the energy range in the data. Different segments of the spectrum can, in principle, be estimated by varying the system stiffness since this changes the energy-frequency relationship.

Figure 2($i$) shows a scatter diagram for estimates of the linear and nonlinear damping parameters ($\hat{A}$ and $\hat{B}$) obtained from 100 blocks with an input amplitude of $K = 0.2$. Here the estimates were derived by the method given in Section 4.3: Thus, for each block, Eq. (55) was fitted to the estimated damping function. Again, there is a negative correlation between these estimates and a significant degree of scatter around the ''target point,'' corresponding to the true damping parameter values (see Section 5.4 for a further discussion of this feature). Averaging these estimates gives the values $\hat{A} = 0.013$ and $\hat{B} = -0.0085$: These are significantly different from the true values.

In the next section it will be shown that, by using results for two different amplitude levels, an individual estimates from a number of blocks, the accuracy of the parameter estimates can be significantly improved.

**5.4 Simulation Results: Two Different Amplitude Levels.** Earlier it has been shown that, for both the stiffness and damping, the linear parameter estimates are negatively correlated with the corresponding nonlinear parameter estimates.

In the case of damping this characteristic arises from the nature of the damping function $d(E)$, as given by Eq. (55). Consider the case of very weak excitation first, where the energy levels are very low. In this case the linear term is dominant. The linear damping parameter $a$ will then be accurately estimated whereas the nonlinear parameter $b$ will be poorly estimated, reflecting the fact that the nonlinear damping contribution is negligible. Thus, in the scatter plot for the estimates, in the $a$, $b$ plane ($a$ horizontal), the estimates will cluster around a vertical straight line: This corresponds to a small dispersion of $\hat{a}$ and a large dispersion of $\hat{b}$. At the other extreme, when there is a high level of excitation, with correspondingly high energy levels, the nonlinear term is dominant. The parameter estimates in the $a$, $b$ plane will then scatter about a horizontal straight line, corresponding to low dispersion of $\hat{b}$ and high dispersion of $\hat{a}$. The intercept of the lines obtained a linear regression applied to the estimates, for the two extreme cases, and can be expected to be close to the point corresponding to the true values of $a$ and $b$. The accuracy of the intercept point is directly related to the accuracy of the damping function estimate.

For intermediate levels of excitation earlier results show that damping estimates still cluster close to a linear variation (see Fig. 2($i$)): The slope depends on the amplitude of the input. The strong negative correlation effect can be explained by considering the difference between the exact damping function given by Eq. (55) and an estimate given by

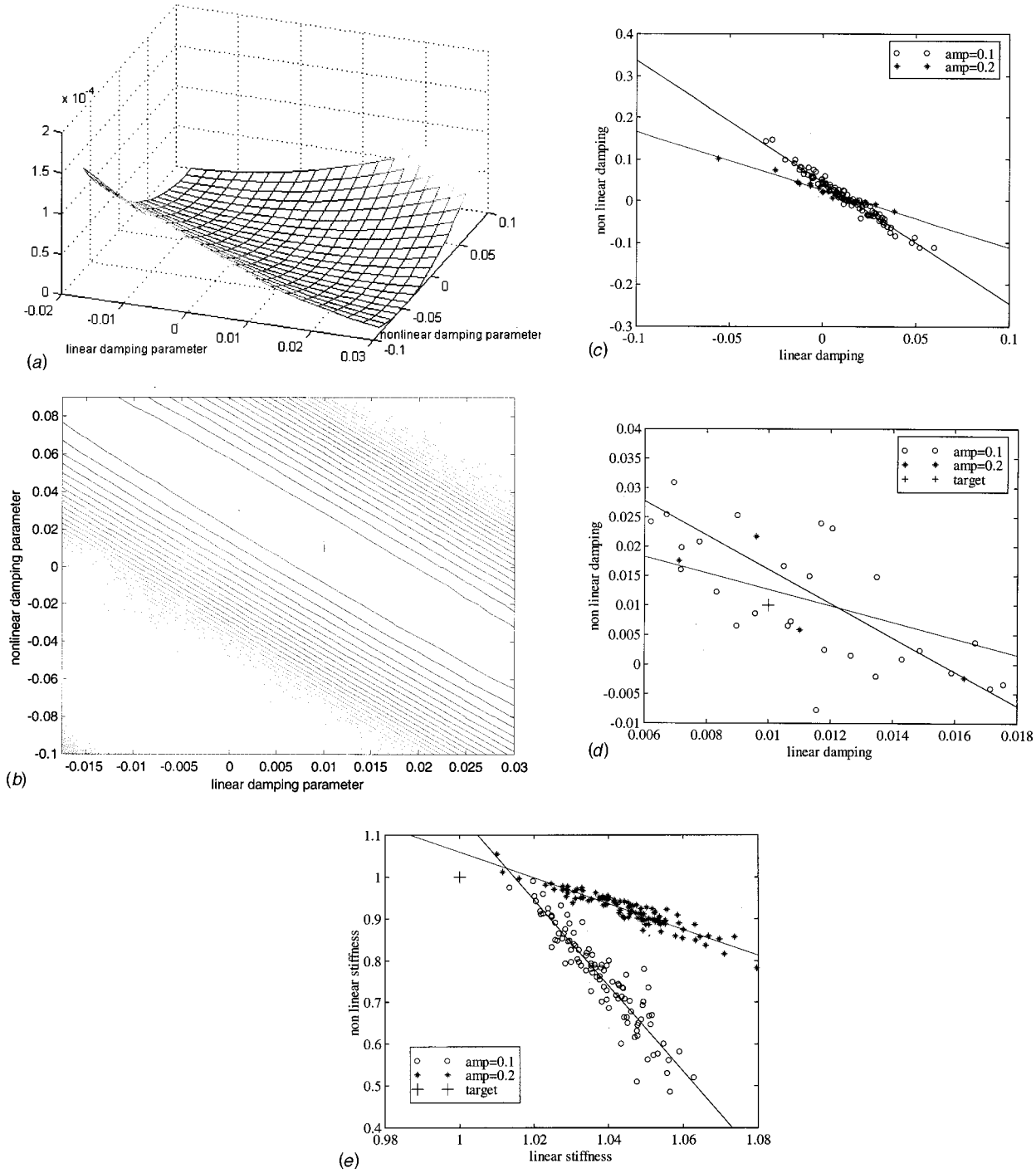



**Fig. 3 (*a*) Three-dimensional plot of $J_d$ versus $\hat{a}$ and $\hat{b}$; (*b*) contour plot of $J_d$ versus $\hat{a}$ and $\hat{b}$; (*c*) scatter plots of the estimates of the linear and nonlinear damping parameters, for two different excitation levels, showing the intersection of the regression lines; (*d*) enlarged plot of part of Fig. 3(*b*); (*e*) scatter plots of the estimates of the linear and nonlinear stiffness parameters, for two different excitation levels, showing the intersection of the regression lines**

$$\hat{d}(E_i)=\hat{a}s_1^2E_i+1.20\hat{b}s_1^2E_i^{3/2} \tag{60}$$

where $E_i$ are discrete values of $E$. A measure of the overall mean square error between the estimated function and the true function is as follows:

$$J_d=\sum_i\ [d(E)-(\hat{A}s_1^2E_i+\hat{B}s_1^2E_i^{3/2})]^2 \tag{61}$$

where $d(E)$ is given by Eq. (55). Figure 3(*a*) is a three-dimensional plot of $J_d$ versus $\hat{a}$ and $\hat{b}$, for the case where $a=0.01$, $b=0.01$, and $E_i=0.1i$ $(i=1,2,\ldots,10)$: The corresponding contour plot is shown in Fig. 3(*b*) (the latter also shows the target point). These figures show that the minimum of $J_d$ is a long and virtually flat valley. Thus, if the true damping function is known, the damping estimates will be at the bottom of the valley; i.e., they will tend to scatter along a straight line which passes through the target point. In practice the true damping function is unknown and must be estimated from the data, in the manner

described in Section 4.2. Errors associated with damping function estimation will be reflected in a small variation in the location of the valley and hence in a small additional scatter of the damping estimates.

These considerations lead to the conclusion that a straight line fitted to the damping estimates in the $a$, $b$ plane, using linear regression, will pass close to the target point: The distance between this line and the target point is dependent on the accuracy of the estimation of the damping function and will reduce if the damping function estimation is improved. Thus if results are obtained at *two different excitation levels*, then *two straight lines* can be obtained, with *different slopes*. Since both lines will pass close to the target point the intercept will be in the vicinity of this point; thus the intercept can be expected to yield improved estimates of $a$ and $b$. These statements apply equally well to the estimates of $A$ and $B$, since these are simply scaled versions of $a$ and $b$.

Figure 3($c$) (and Fig. 3($d$), which is an enlarged view) shows damping parameter estimates, $\hat{A}$ and $\hat{B}$, obtained from the two levels of excitation $K=0.1$ and $K=0.2$. As deduced from the argument above, the intersection of the two linear regression lines, shown in the figures, provides significantly improved damping estimates; $\hat{A}=0.012$ and $\hat{B}=0.0095$.

A similar improvement is possible in the case of stiffness estimation, if results for two different excitation levels are available. Again, as the results in Fig. 1($b$) indicate, the estimates in the $k_1,k_2$ plane scatter about a linear variation. This can be explained by considering the difference between the exact potential energy function, $V(x)$, and an estimate given by

$$\hat{V}(x_i)=\hat{k}_1\left(\frac{x^2}{2}\right)+\hat{k}_2\left(\frac{x^4}{4}\right). \tag{62}$$

The overall mean square error $J_s$, between the estimated function and the true function, may be defined in a similar way to the approach used to define $J_d$ (see Eq. (61)). Numerical studies show that a three-dimensional plot of $J_s$ against $k_1$ and $k_2$ has similar characteristics to the corresponding plot for $J_d$; i.e., a long and virtually flat valley.

Thus, as for damping estimation, one can expect that improved estimation of the stiffness parameters is obtainable from the intercept of two linear regression lines, fitted to estimates from a number of blocks, at two different excitation levels. Figure 3($e$) shows scatter diagrams for the $k_1,k_2$ estimates obtained from excitation at two different levels ($K=0.1$ and 0.2). In both cases 100 blocks of data were processed. The intersection of the regression lines yields the estimates $\hat{k}_1=1.01$, $\hat{k}_2=1.02$: The error here is thus very small and significantly smaller than that obtainable from results at a single input level.

## 6 Conclusions

The principal conclusions are summarized as follows:

1 For a lightly damped nonlinear oscillator, driven by a stationary stochastic process it is possible to obtain estimates the restoring characteristic from measurements of the response process only, in the general case where the excitation is nonwhite.

2 This method can be combined with a technique based on a Markov model of the response, to enable estimates of the damping function and the spectrum of the excitation to be generated. Again, only measurements of the response are required.

3 By combining results from tests at two different levels of excitation, the accuracy of estimation for both the stiffness and damping parameters can be significantly improved.


## Acknowledgments

Dr. Vasta was funded by an EU Marie Curie Fellowship. The authors gratefully acknowledge this support.


## References


[1] Masri, S. F., and Caughey, T. K., 1979, ''A Nonparametric Identification Technique for Nonlinear Dynamic Problems,'' ASME J. Appl. Mech., **46**, pp. 433–447.

[2] Chassiakos, A. G., Masri, S. F., Smyth, A. W., and Caughey, T. K., 1998, ''On-Line Identification of Hysteretic Systems,'' ASME J. Appl. Mech., **65**, pp. 194–203.

[3] Yasuda, K., Kamiya, K., and Komakine, M., 1997, ''Experimental Identification Technique of Vibrating Structures With Geometrical Nonlinearity,'' ASME J. Appl. Mech., **64**, pp. 275–280.

[4] Unbehauen, H., and Rao, G. P., 1990, ''Continuous Time Approaches to System Identification—A Survey,'' Automatica, **26**, pp. 23–35.

[5] Feng, M. Q., Kim, J.-M., and Xue, H., 1998, ''Identification of a Dynamic System Using Ambient Vibration Measurements,'' ASME J. Appl. Mech., **65**, pp. 1010–1021.

[6] Roberts, J. B., Dunne, J. F., and Debonos, A., 1992, ''Estimation of Ship Roll Parameters in Random Waves,'' ASME J. Offshore Mech. Arct. Eng., **114**, pp. 114–121.

[7] Roberts, J. B., Dunne, J. F., and Debonos, A., 1994, ''Stochastic Estimation Methods for Nonlinear Ship Roll Motion,'' Prob. Eng. Mech., **9**, pp. 83–93.

[8] Roberts, J. B., Dunne, J. F., and Debonos, A., 1995, ''A Spectral Method for Estimation of Nonlinear System Parameters for Measured Response,'' Prob. Eng. Mech., **10**, pp. 199–207.

[9] Roberts, J. B., Dunne, J. F., and Debonos, A., 1996, ''Parameter Estimation for Randomly Excited Non-linear Systems: A Method Based on Moment Equations and Measured Response Histories,'' *Proceedings of the IUTAM Symposium on Advances in Non-linear Stochastic Mechanics*, A. Naess and S. Krenk, eds., Trondheim, Norway, July 1995, Kluwer, Dordrecht, The Netherlands, pp. 361–372.

[10] Battaini, M., and Roberts, J. B., 1977, ''Moment and Spectral Methods for Stochastic Parameter Estimation of Multi-Degree of Freedom Systems,'' *Proceedings of the ESREL '97 International Conference on Safety and Reliability*, Vol. 2, G. Guedes Soares, ed., Lisbon, Portugal, Pergamon, Oxford, UK, pp. 1323–1330.

[11] Vasta, M., and Roberts, J. B., 1998, ''Stochastic Parameter Estimation of Non-linear Systems Using Higher-Order Spectra,'' J. Sound Vib., **213**, pp. 201–221.

[12] Roberts, J. B., and Vasta, M., 2000, ''Parametric Identification of Systems With Non-Gaussian Excitation Using Measured Response Spectra,'' Prob. Eng. Mech., **15**, pp. 59–71.

[13] Krenk, S., and Roberts, J. B., 1999, ''Local Similarity in Nonlinear Random Vibration,'' ASME J. Appl. Mech., **66**, pp. 225–235.

[14] Roberts, J. B., and Vasta, M., 2000, ''Markov Modelling and Stochastic Estimation for Nonlinear Ship Rolling in Irregular Seas,'' Phil. Trans. R. Soc. Lond. A., **358**, pp. 1917–1941.

[15] Roberts, J. B., and Spanos, P. D., 1986, ''Stochastic Averaging: An Approximate Method of Solving Random Vibration Problems,'' Int. J. Non-Linear Mech., **21**, pp. 111–134.

[16] Roberts, J. B., 1982, ''A Stochastic Theory for Nonlinear Ship Rolling in Irregular Seas,'' J. Ship Res., **26**, pp. 229–245.

[17] Roberts, J. B., and Vasta, M., 1999c, ''Response of Nonlinear Oscillators to Nonwhite Random Excitation Using an Energy Based Method,'' *Proceedings of the IUTAM Symposium on Nonlinearity and Stochastic Structural Dynamics*, Chennai, India, to be published by Kluwer Academic Publishers.

[18] Gardiner, C. W., 1985, *Handbook of Stochastic Methods*, Springer-Verlag, Berlin.

[19] Stratonovitch, R. L., 1963, *Topics in the Theory of Random Noise*, Vol. 1 and 2, Gordon and Breach, New York.

# Vibrations of a Taut Cable With an External Damper

**S. Krenk**
Department of Structural Engineering
and Materials,
Technical University of Denmark,
DK-2800 Lyngby, Denmark

*A solution is presented of the problem of vibrations of a taut cable equipped with a concentrated viscous damper. The solution is expressed in terms of damped complex-valued modes, leading to a transcendental equation for the complex eigenfrequencies. A simple iterative solution of the frequency equation for all complex eigenfrequencies is proposed. The damping ratio of the vibration modes, determined from the argument of the complex eigenfrequency, are typically determined to within one percent in two iterations. An accurate asymptotic approximation of the damping ratio of the lower modes is obtained. This formula permits explicit determination of the optimal location of the viscous damper, depending on its damping parameter.* [S0021-8936(00)00404-9]

## 1 Introduction

Cables constitute an efficient means of supporting static loads and find increasing use as stay cables for, e.g., bridges and towers ([1]). However, the slenderness of the cables makes them prone to vibrations, either induced by direct loads on the cable from wind or a combination of wind and rain, or via motion of the supported structure (Watson and Stafford [2]). The cables are usually made of strands of high strength steel, and this leads to extremely low material damping, typically in the order of 0.001 of critical Yamaguchi and Fujino [3]. Experiments and analysis of vibrating cables indicate that the material damping can be largely associated with hysteretic energy loss due to stretching of the cable (Yamaguchi and Jayawardena [4]).

The light damping of the cable itself leads to a need to provide external damping, and for this transverse dampers attached at a distance of two to four percent of the span from one of the supports have been used in several cases ([5]). The motion of the damper is limited by its closeness to the support, and it is of importance to find the optimal combination of damper location and external damping magnitude. If the external damping is too large, the cable will vibrate as if it was supported at the location of the damper. Thus, in that case the effect of the damper would mainly be a slight increase in the vibration frequency, without any appreciable damping effect on the vibrations.

The problem of the optimal damping constant of a viscous damper located close to one end of a taut cable was studied by Kovacs [6] who used a semi-empirical interpolation between the solutions of undamped vibrations of a cable without damper and a cable supported at the location of the damper. The results suggested that the maximum modal damping ratio that could be obtained by a concentrated viscous damper, would be about half the relative distance of the damper from the support. This result was confirmed in a numerical study by Pacheco et al. [7], who also found an approximate value of the optimal external damping constant for the lower modes of vibration, when the damper is close to a support. Their method of analysis consisted in representing the displacement in terms of a sine function expansion, and the concentrated action of the damper made the use of 200–300 terms necessary for good accuracy. A later analysis by Baker et al. [8] aiming at active control of the cable included the static deflection in the representation of the transverse motion, and thereby reduced the necessary number of terms considerably.

In the present paper the problem of damping of a vibrating string by a concentrated viscous damper is formulated and solved by use of complex-valued modes. The problem formulation results in a fairly simple transcendental equation for the complex eigenfrequencies, and the modal damping ratios are identified directly from the imaginary parts of the eigenfrequencies. As it turns out an explicit asymptotic result can be obtained, yielding a simple analytical formula for the numerical results of Pachero et al. [7]. Furthermore, a simple iterative scheme is derived that yields accurate complex eigenfrequencies within two iterations also for higher modes and general location of the damper. The complex modes are illustrated, and it is found that optimal external damping corresponds to a phase lag of $1/4\pi$ of the combined cable and damper force acting on the long part of the cable at the location of the damper.

The equation of motion of the taut string is similar to that of axial vibrations of a beam. This problem has been studied by use of complex modes for the special case of a damper at an otherwise free end ([9,10]). However, in that problem the boundary condition is fairly simple, and the main issue is the representation of forced vibrations in terms of the complex modes. Analogous beam problems with viscous dampers at the ends have been treated by Oliveto et al. [11] and Yang and Wu [12]. It is clear from these analyses that the response can be represented in terms of the complex modes by an extension of classical modal analysis, in which the inertial term is represented as the time derivative of a separate momentum representation, much in the spirit of Hamiltonian mechanics. The general conclusion of relevance for the present work is that the damping ratio, determined from the complex eigenvalue, plays the same role as in classical modal analysis.

## 2 Problem Formulation

The cable and the concentrated viscous damper are shown in Fig. 1. The length of the cable is $l$, and the damper is located at the distance $a$ from the left end. The cable force is $T$, and the mass per unit length is $m$. The coefficient of the concentrated damper is denoted $c$, and the cable itself is assumed undamped. In the following it is convenient to introduce the complementary length $a'=l-a$ and the complementary coordinate $x'=l-x$.

Linear oscillations of the cable under the assumption of virtually unchanged cable force are described by the partial differential equation

$$T\frac{\partial^2 v}{\partial x^2}-m\frac{\partial^2 v}{\partial t^2}=c\frac{\partial v}{\partial t}\,\delta(x-a) \tag{1}$$

where $v(x,t)$ is the transverse displacement, and $\delta(x-a)$ is the delta function, specifying the location of the damper force at $x=a$.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, August 12, 1999; final revision, May 2, 2000. Associate Technical Editor: A. A. Ferri. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

 **Copyright © 2000 by ASME** 

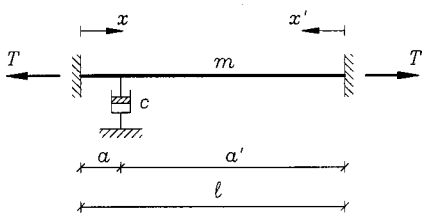



**Fig. 1 Taut cable with external viscous damper**

The partial differential Eq. (1) is to be solved with the boundary conditions

$$v(0,t)=0, \quad v(l,t)=0 \tag{2}$$

specifying fixed ends. At the location of the damper there is a discontinuity in the inclination of the cable, providing a transverse force matching the force in the damper. This relation has the form

$$T\left(\left.\frac{\partial v}{\partial x}\right|_{a+}-\left.\frac{\partial v}{\partial x}\right|_{a-}\right)=c\,\frac{\partial v}{\partial t}. \tag{3}$$

It is seen that this equation implies a linear relation between the displacement space and time derivatives at $x=a$, and this coupling will lead to complex mode shapes and frequencies of free vibrations.

The transverse displacement of free damped vibrations of the string can be expressed in the form

$$v(x,t)=\mathrm{Re}[\tilde{v}(x)\exp(i\omega t)] \tag{4}$$

where $\omega$ is a complex natural eigenfrequency, and $\tilde{v}(x)$ is the corresponding complex mode shape. The complex mode shape $\tilde{v}(x)$ satisfies the ordinary differential equation

$$\frac{d^2\tilde{v}}{dx^2}+\beta^2\tilde{v}=0, \quad \begin{cases}0<x<a\\ 0<x'<a'\end{cases} \tag{5}$$

where the wave number $\beta$ has been introduced as

$$\beta=\omega\sqrt{\frac{m}{T}}. \tag{6}$$

The differential Eq. (5) is homogeneous in each of the two intervals between the ends and the damper. The solution is conveniently expressed in the form

$$\tilde{v}(x)=\begin{cases}A\sin(\beta x), & 0\leqslant x\leqslant a\\ A'\sin(\beta x'), & 0\leqslant x'\leqslant a'\end{cases} \tag{7}$$

where $A$ and $A'$ are arbitrary constants. Continuity of the shape function at $x=a$ requires

$$A\sin(\beta a)=A'\sin(\beta a')=\tilde{v}_a \tag{8}$$

where $\tilde{v}_a=\tilde{v}(a)$. Hereby the solution (7) takes the form

$$\tilde{v}(x)=\begin{cases}\tilde{v}_a\dfrac{\sin(\beta x)}{\sin(\beta a)}, & 0\leqslant x\leqslant a\\[2ex] \tilde{v}_a\dfrac{\sin(\beta x')}{\sin(\beta a')}, & 0\leqslant x'\leqslant a'.\end{cases} \tag{9}$$

The amplitude $\tilde{v}_a$ is undetermined in a free vibration.

Substitution of the mode shape representation (9) into the force balance relation (3) provides an equation for determination of the complex eigenfrequencies.

$$\cot(\beta a)+\cot(\beta a')=-i\,\frac{c}{\sqrt{Tm}} \tag{10}$$

The solution of this frequency equation is treated in the following section.

## 3 Complex Eigenfrequencies

The wave number solutions to the frequency Eq. (10) are denoted $\beta_n$, $n=1,2,\ldots$, and the corresponding eigenfrequencies $\omega_n$. The eigenfrequencies of the damped modes are complex, with the imaginary part representing the attenuation due to damping. The magnitude of the imaginary part is represented via the modal damping ratio $\zeta_n$, obtained from the representation

$$\omega_n=|\omega_n|(\sqrt{1-\zeta_n^2}+i\zeta_n). \tag{11}$$

The modulus $|\omega_n|$ describes the magnitude of the angular frequency, and for small damping ratio the frequency of the oscillatory part of the solution.

The coefficient $c$ of the concentrated damper enters the solution via the nondimensional damping coefficient

$$\eta=\frac{c}{\sqrt{Tm}} \tag{12}$$

where the damping is normalized by the square root of the cable force and the cable mass per unit length. This is similar to the normalization with respect to stiffness and mass in an oscillator. However, in the present continuous system, the effectiveness of the damper depends on the location of the damper.

In terms of the nondimensional damping coefficient $\eta$ the frequency Eq. (10) is

$$\cot(\beta a)+\cot(\beta a')=-i\,\eta. \tag{13}$$

In this equation $\cot(\beta a')$ is expressed in terms of $\cot(\beta l)$ and $\cot(\beta a)$ by use of the trigonometric relation

$$\cot(\beta a')=\cot(\beta l-\beta a)=\frac{1+\cot(\beta l)\cot(\beta a)}{\cot(\beta a)-\cot(\beta l)}. \tag{14}$$

This gives the equation

$$\frac{1+\cot^2(\beta a)}{\cot(\beta l)-\cot(\beta a)}=i\,\eta. \tag{15}$$

In this equation the term $\tan(\beta l)$ is isolated on the left side,

$$\tan(\beta l)=\frac{i\,\eta\sin^2(\beta a)}{1+i\,\eta\cos(\beta a)\sin(\beta a)}. \tag{16}$$

This form of the frequency equation is suitable for solution, either in asymptotic form or numerically by iteration.

**3.1 Asymptotic Solution.** It is convenient first to obtain an asymptotic solution of the frequency Eq. (16), valid for $a/l\ll 1$. The asymptotic solution identifies suitable parameters and illustrates the main features of the general solution.

The asymptotic solution is obtained by considering a small perturbation on the solution without damping. In the absence of damping the wave numbers of the simply supported string are

$$\beta_n^0=n\,\frac{\pi}{l}, \quad n=1,2,\ldots. \tag{17}$$

Assuming the wave numbers $\beta_n$ of the damped problem to be small perturbations from $\beta_n^0$, the tangent can be approximated by

$$\tan(\beta_n l)\simeq\beta_n l-n\pi. \tag{18}$$

The assumption $a/l\ll 1$ leads to the following representations:

$$\sin(\beta_n a)\simeq n\pi\frac{a}{l}, \quad \cos(\beta_n a)\sin(\beta_n a)=\frac{1}{2}\sin(2\beta_n a)\simeq n\pi\frac{a}{l}. \tag{19}$$

These representations are useful for $\beta_n a\lesssim 1/4\pi$, although perhaps not very accurate around the upper limit. With the estimate $\beta_n l\simeq n\pi$ this corresponds to $na/l\lesssim 1/4$.

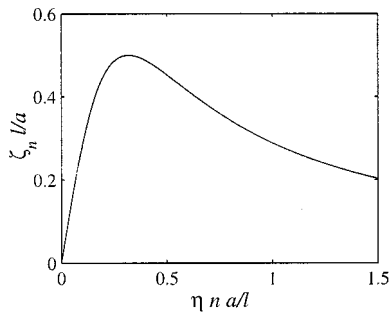



**Fig. 2 The asymptotic modal damping relation (23)**

Introduction of the asymptotic representations (18) and (19) into the frequency Eq. (16) leads to the following explicit asymptotic formula for the wave number $\beta_n$:

$$\beta_n l \simeq n\pi + \frac{i\eta(n\pi a/l)^2}{1+i\eta(n\pi a/l)}. \tag{20}$$

The imaginary part is evaluated from the last term as

$$\mathrm{Im}[\beta_n l] \simeq \frac{\eta(n\pi a/l)^2}{1+\eta^2(n\pi a/l)^2} \tag{21}$$

while the real part is asymptotically equal to $n\pi$. Thus an asymptotic approximation of the damping ratio is obtained as

$$\zeta_n \simeq \frac{\mathrm{Im}[\beta_n l]}{n\pi} \simeq \frac{(\eta n\pi a/l)}{1+(\eta n\pi a/l)^2}\frac{a}{l}. \tag{22}$$

When the factor $a/l$ is transferred to the left side of the equation, the result takes the form

$$\frac{\zeta_n}{a/l} \simeq \frac{(\eta n\pi a/l)}{1+(\eta n\pi a/l)^2}. \tag{23}$$

In this form the relation gives the modal damping in terms of the parameter $\zeta_n/(a/l)$ as a simple rational function of the parameter $\eta n\pi a/l$, combining the parameter $\eta$ of the external damper, the mode number $n$ and the location $a/l$ of the damper. These two parameters were identified from numerical computations using a large number of real *undamped* mode-shapes by Pacheco et al. [7].

The asymptotic relation (23) is shown in Fig. 2. It is seen that initially the modal damping increases with increasing external damping $\eta$, it then reaches a maximum and decreases upon further increase of $\eta$. For a given location $a/l$ of the damper the optimal value of the external damping coefficient $\eta$ is obtained from the condition

$$\left.\frac{\partial\zeta_n}{\partial\eta}\right|_a = 0 \tag{24}$$

leading to the value

$$\eta_{\mathrm{opt}}\pi n a/l = 1 \quad \text{or} \quad \eta_{\mathrm{opt}} = \frac{1}{\pi n a/l}. \tag{25}$$

The asymptotic estimate of the maximum damping follows from substitution of this value into (22), leading to

$$\frac{\zeta_{n,\mathrm{opt}}}{a/l} \simeq 0.5 \quad \text{or} \quad \zeta_{n,\mathrm{opt}} \simeq \frac{a}{2l}. \tag{26}$$

The solution is asymptotically valid for small $na/l$. By (25) this implies large values of the optimal external damping parameter, e.g., by the previous estimate $na/l \lesssim 1/4$ and $\eta_{\mathrm{opt}} \gtrsim 4/\pi$.

The asymptotic relation (22) and the optimum values (25) and (26) constitute an analytical solution of the ''universal curve'' determined from a numerical solution of the cable problem by Pacheco et al. [7]. In their paper the external damping was normalized via the first undamped eigenfrequency $\omega_1 = \beta_1^0\sqrt{T/m} = (\pi/l)\sqrt{T/m}$ using the parameter $ca/(\omega_1 m l^2) = \eta/\pi$. According to (25) the optimal value of this parameter is $\pi^{-2} \simeq 0.10$ as also found in the numerical study of Pacheco et al. [7]. The approximate procedure of Kovacs [6] leads to the estimate $ca/(\omega_1 m l^2) \simeq 1/(2\pi) = 0.16$ for the optimal external damping parameter, thus overestimating its value by approximately 60 percent.

**3.2 Numerical Solution by Iteration.** The form (16) of the frequency equation is also suitable for direct iterative solution, substituting current estimate $\beta_n^j$ on the right side and evaluating a new estimate $\beta_n^{j+1}$ by inversion of the tangent function. Starting from the undamped wave numbers $\beta_n^0$ from (17), the iterative scheme is given by

$$\beta_n^{j+1} l = n\pi + \arctan\left(\frac{i\eta\sin^2(\beta_n^j a)}{1+i\eta\cos(\beta_n^j a)\sin(\beta_n^j a)}\right), \quad j=0,1,\ldots. \tag{27}$$

Results are presented in Figs. 3(*a*) and 3(*b*) for $a/l = 0.02$ and $a/l = 0.05$, respectively. The fully drawn graph corresponds to the asymptotic relation (23). It is seen that the deviations from the asymptotic curve (23) are negligible for $a/l = 0.02$, while for $a/l = 0.05$ the optimal damping increases slightly and corresponds to a value of the external damping parameter $\eta$ slightly larger than $(\pi n a/l)^{-1}$.

The iterative scheme (27) converges rapidly as illustrated in Figs. 4(*a*) and 4(*b*). The figures show the error $|\zeta_n^j - \zeta_n| l/a$ at $\eta n a/l = 1/\pi$ for the cases of Figs. 3(*a*) and 3(*b*), respectively. Index $j=0$ corresponds to the undamped wave number (17), for which the damping ratio vanishes. Thus, the error for $j=0$ is

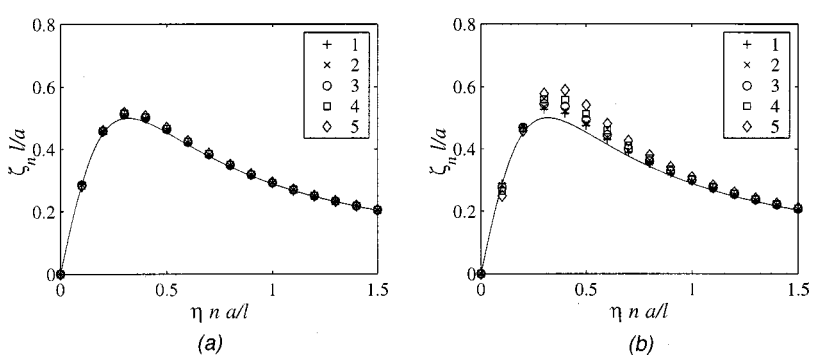



**Fig. 3 Damping ratio of the first five modes; (*a*) $a/l = 0.02$, (*b*) $a/l = 0.05$**

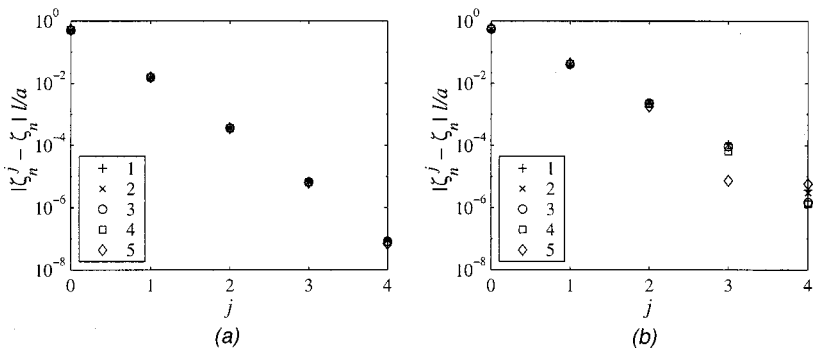



**Fig. 4 Iteration convergence of first five modes; (*a*) $a/l=0.02$, (*b*) $a/l=0.05$.**

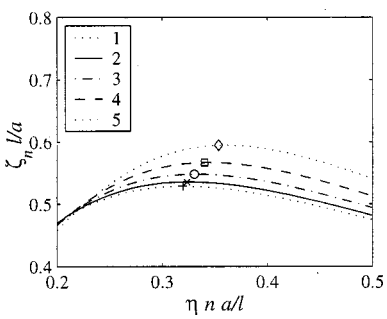



**Fig. 5 Optimal modal damping for $a/l=0.05$. Curves from Fig. 3(*b*) and estimates from (29).**

equal to the final damping ratio. Convergence is slightly faster for $a/l=0.02$ than for $a/l=0.05$, but in both cases the error is less than one percent after two iterations.

The first iterate $\beta_n^1$ from (27) does not provide a systematic improvement over the asymptotic result (20). However, the damping ratio determined from $\beta_n^1$ correctly represents the *location* of the maximal modal damping. Thus, a good approximation of the optimal value of the parameter $\eta n a/l$ can be determined from the damping ratio from $\beta_n^1$. The optimal value is estimated from the condition

$$\frac{d\,\mathrm{Im}[\beta_n^1 l]}{d\eta}=\mathrm{Im}\left[\frac{d(\beta_n^1 l)}{d\eta}\right]=0. \tag{28}$$

This condition leads to a replacement of the asymptotic result (25) with

$$\eta_{\mathrm{opt}}\sin(\pi n a/l)=1 \quad \text{or} \quad \eta_{\mathrm{opt}}=\frac{1}{\sin(\pi n a/l)}. \tag{29}$$

This result based on the first iterate contains the asymptotic result (25) as the limit for $a/l \ll 1$. For finite values of $a/l$ it correctly represents an increase in the optimal external damping parameter $\eta_{\mathrm{opt}}$. This is illustrated in Fig. 5, showing the central part of the curves from Fig. 3(*b*) for $a/l=0.05$ together with the maximum values estimated from $\eta_{\mathrm{opt}}$ from (29). The estimate (29) of the optimal external damping parameter $\eta_{\mathrm{opt}}$ is seen to be quite accurate. However, the modal damping obtained by using the optimal external damping must be calculated by the iteration procedure, using at least two iterations.

## 4 Complex Mode Shapes

The complex mode shapes are given by (9) in terms of the displacement amplitude $\tilde{v}_a$ at $x=a$ and the complex wave number $\beta$. When the displacement $v(a,t)=\mathrm{Re}[\tilde{v}_a \exp(i\omega t)]$ attains its maximum value, the velocity $\partial v(a,t)/\partial t=\mathrm{Re}[i\omega\tilde{v}_a \exp(i\omega t)]$ vanishes. Thus, at this instant in time there is no damping force and the tangent changes continuously across $x=a$. Conversely, a quarter period later the displacement $v(a,t)$ vanishes, while the velocity $\partial v(a,t)/\partial t$ reaches its numerically largest value. Figure 6(*a*) illustrates the development of the mode shape in this quarter period, starting with a smooth curve and positive displacements, and gradually developing a slope discontinuity at $x=a$, reaching its maximum value a quarter period later.

The slope discontinuity at the location of the damper follows from multiplication of (13) by $\beta\tilde{v}_a$,

$$\left(\left.\frac{d\tilde{v}(x)}{dx}\right|_{a+}-\left.\frac{d\tilde{v}(x)}{dx}\right|_{a-}\right)=-\beta(\cot(\beta a)+\cot(\beta a'))\tilde{v}_a=i\beta\eta\tilde{v}_a. \tag{30}$$

When the optimal external damping is estimated from (29), and $\beta$ is approximated by its undamped value (17), the maximum value of the slope discontinuity, attained at the end of the quarter period, is

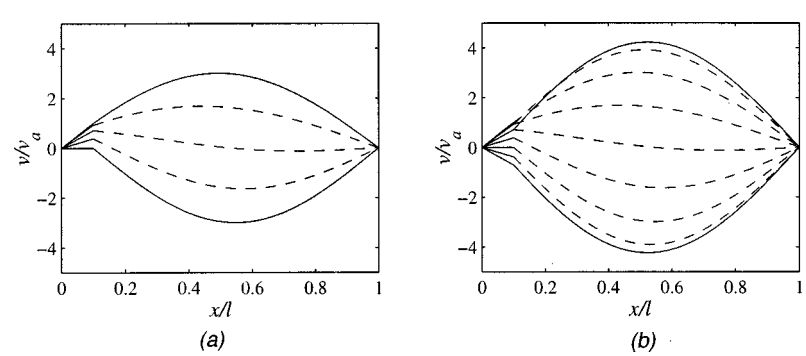



**Fig. 6 First complex mode $a/l=0.1$ and $\eta=\eta_{\mathrm{opt}}$; (*a*) quarter period from max $v(a)$, (*b*) half period from max $v(1/2a')$.**

$$\max_t\left(\left.\frac{\partial v}{\partial x}\right|_{a+}-\left.\frac{\partial v}{\partial x}\right|_{a-}\right)\simeq\beta_n^0\eta_{\mathrm{opt}}\tilde{v}_a\simeq\frac{\pi n/l}{\sin(\pi na/l)}\tilde{v}_a\simeq\frac{\tilde{v}_a}{a}. \tag{31}$$

When the damper is close to the end of the cable, the part of cable between the damper and the closest support remains approximately straight, and thus the slope discontinuity (31) at the end of the quarter period is approximately equal to the slope at the support at the beginning of the quarter period, as clearly seen in Fig. 6(*a*).

At the beginning and the end of the quarter period the mode shape looks quite similar, but with a length that is reduced from $l$ to $l-a$. However, the maximum displacement within the main span is attained before the displacement of the damper $v(a,t)$ reaches its maximum. Figure 6(*b*) shows the development of the mode shape over a half period, starting at the time of the maximum displacement at $x'=1/2a'$.

The asymptotic optimal value $\eta_{\mathrm{opt}}$ for a damper located close to a support can be given an interpretation in terms of the transverse forces on the cable at the location of the damper. The part of the cable extending from the damper to the remote support experiences two transverse force components at the location of the damper: a transverse component $F_T(t)$ of the cable tension $T$ due to the inclination of the part of the cable between the damper and the closest support, and the force $F_c(t)$ from the damper. In the lower modes of vibration the part of the cable between the damper and the closest support is approximately straight, and thus $F_T\simeq T\tilde{v}_a/a$. The force from the damper $F_c=i\omega_n c\tilde{v}_a$ is $1/2\pi$ out of phase. The ratio between the magnitude of these two force components is

$$\frac{|F_c|}{|F_T|}=\frac{\omega_n ca}{T}\simeq\pi n\frac{a}{l}\eta \tag{32}$$

where the approximation $\omega_n\simeq(n\pi/l)\sqrt{T/m}$ has been introduced. When the asymptotic optimal damping parameter $\eta_{\mathrm{opt}}$ is inserted from (25), optimal damping is seen to correspond to $|F_c|\simeq|F_T|$. Thus, at optimal damping the two force components are equal, and the force acting on the longer part of the cable lags $1/4\pi$ behind the displacement.

## 5 Conclusions

The damping of a taut cable by an isolated viscous damper has been analyzed in terms of complex-valued modes. An asymptotic solution has been obtained for the damping ratio of the lower modes, valid for damper location close to one support. The asymptotic solution defines the ''natural'' damping parameters as $\zeta_n/(a/l)$ for modal damping and $n(a/l)\sqrt{c/(Tm)}$ for the external damping, and identifies maximal modal damping to be $\zeta_n\simeq 1/2a/l$, attained for $c_{\mathrm{opt}}\simeq\sqrt{Tm}/(\pi na/l)$. For $a/l\gtrsim 0.05$ differences from the asymptotic results become noticeable, and a slightly larger external damping becomes optimal. An accurate approximation has been obtained for the optimal external damping, reflecting the increase with increasing $na/l$, and a general iterative procedure has been given for the modal damping. Typically one or two iterations yield the modal damping ratio within a relative error of less than one percent.


## Acknowledgment

This work has been supported by the Danish Technical Research Council and a visiting appointment in July 1999 at University of Notre Dame.


## References


[1] Irvine, H. M., 1981, *Cable Structures*, M.I.T. Press, Cambridge, MA.
[2] Watson, S. C., and Stafford, D., 1988, ''Cables in Trouble,'' Civil Engineering ASCE, **58**, No. 4, pp. 38–41.
[3] Yamaguchi, H., and Fujino, Y., 1998, ''Stayed Cable Dynamics and Its Vibration Control,'' *Bridge Aerodynamics*, A. Larsen and S. Esdahl, eds., Balkma, Rotterdam, pp. 235–253.
[4] Yamaguchi, H., and Jayawardena, L., 1992, ''Analytical Estimation of Structural Damping in Cable Structures,'' Journal of Wind Engineering and Industrial Aerodynamics, **43**, pp. 1961–1972.
[5] Gimsing, N. J., 1997, *Cable Supported Bridges*, 2nd Ed., Wiley, Chichester,
[6] Kovacs, I., 1982, ''Zur Frage der Seilschwigungen und der Seildampfung,'' Die Bautechnik, **59**, No. 10, pp. 325–332.
[7] Pacheco, B. N., Fujino, Y., and Sulekh, A., 1993, ''Estimation Curve for Modal Damping in Stay Cables With Viscous Damper,'' J. Struct. Eng., **119**, pp. 1961–1979.
[8] Baker, G. A., Johnson, E. A., Spencer, B. F., and Fujino, Y. 1999, ''Modeling and Semiactive Damping of Stay Cables,'' 13'th ASCE Engineering Mechanics Division Conference, Johns Hopkins University, Baltimore MD, June 13–16.
[9] Singh, R., Lyons, W. M., and Prater, G., 1989, ''Complex Eigensolution for Longitudinally Vibrating Bars With a Viscously Damped Boundary,'' J. Sound Vib., **133**, pp. 364–367.
[10] Hull, A. J., 1994, ''A Closed Form Solution of a Longitudinal Bar With a Viscous Boundary Layer,'' J. Sound Vib., **169**, pp. 19–28.
[11] Oliveto, G., Santini, A., and Tripodi, E., 1997, ''Complex Modal Analysis of a Flexural Vibrating Beam With Viscous End Conditions,'' J. Sound Vib., **200**, pp. 327–345.
[12] Yang, B., and Wu, X., 1997, ''Transient Response of One-Dimensional Distributed Systems: A Closed Form Eigenfunction Expansion Realization,'' J. Sound Vib., **208**, pp. 763–776. UK.

# Initial Post-Buckling and Growth of a Circular Delamination Bridged by Nonlinear Fibers

**S. Li**

**J. Nie**

**J. Qian**

Department of Civil Engineering,
Tsinghua University,
Beijing 100084, P. R. China

**Y. Huang**

**Y. Hu**

Department of Mechanics,
Huazhong University of
Science and Technology,
Wuhan 430074, P. R. China

*Axisymmetric buckling, initial post-buckling and growth of a circular delamination bridged by nonlinear fibers in three-dimensional composites are studied by a perturbation method. The through-thickness fibers are assumed to provide nonlinear restoring traction resisting the deflection of the delaminated layer. A closed-form solution for the central deflection of the delamination due to on applied compressive stress during initial post-buckling is obtained. In addition, some simple formulas for calculating the strain energy release rate and the mixed mode stress intensity ratio (i.e., Mode II versus Mode I) at the delamination crack tip are also established. Some interesting conclusions arising directly from the perturbation solutions are drawn. These include: (1) initial post-buckling behavior of a circular delamination is unstable for a softening bridging model; this may result in initial delamination growth for some materials with lower fracture toughness when the delamination buckles rather than post-buckles. However, stable growth is obtained for a hardening bridging model; (2) with an increase of the nonlinear fiber bridging parameter ($\beta$), the residual stiffness of a three-dimensional composite structure with a circular delamination increases gradually; (3) bridging force changes the catastrophic nature of the delamination growth and increases the stability of the delamination. The range and the dynamic effect of the unstable delamination growth diminish or disappear as the bridging parameters increase; (4) for the bridged delamination, the higher the material fracture toughness, the higher the stability of the delamination growth, and the smaller the range and dynamic effect of its unstable growth.* [S0021-8936(00)03203-7]

## 1 Introduction

Owing to the imperfection of production techniques and accidental impact by foreign objects, composite structures will inevitably develop interface cracks and delaminations, which will seriously degrade the structural integrity and thus the compressive stiffness and buckling strength. In addition, delaminations will degrade other structural properties of composite structures, such as energy-absorbing ability ([1]).

Researchers began to pay attention to the buckling and growth of delaminations in composites structures loaded by compressive forces in the early 1980s (see [2]). Much experimental and theoretical research has been conducted on this subject (e.g., [3–12]). It was found that post-buckling produces delamination growth, thereby rendering composites of little practical use. So, the delamination problem can be considered a plate post-buckling problem coupled with interfacial fracture.

Presently, three-dimensional composites are used more and more widely owing to the development of better material production technique and needs of practical engineering. Since the lateral fibers acting on the delamination can increase the residual strength of the structure during post-buckling, it is important, from theoretical and practical points of view, to discuss delamination buckling and growth in the presence of bridging by fibers. In recent years there have been some investigations on this topic (a partial list includes investigations by Shu and Mai [13], Cox [14], Cox and Marshall [15], Li et al. [16,17], Hu et al. [18], and Massabo and Cox [19]). The buckling problem of a one-dimensional delamination bridged by fibers was discussed by Cox [14]. Bridging fibers were assumed to produce a continuously distributed traction which was proportional to the delamination deflection when the delamination buckles. In fact, this is a linear bridging model. Based on the linear theory of a beam on elastic foundation, some key features of the delamination buckling were discovered by Cox [14]. In our recent paper ([20]), the energy variational method was used to solve the problem of a circular delamination bridged by linear fibers. Under the boundary condition that the ends of the delaminated layer are clamped and with deflections

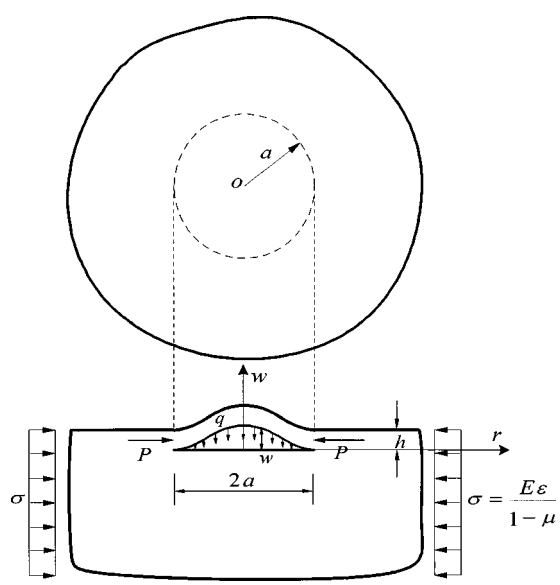



**Fig. 1 A circular thin-film delamination with fiber bridging**

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Jan. 5, 1999; final revision, Apr. 7, 2000. Associate Technical Editor: M.-J. Pindera. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

 **Copyright © 2000 by ASME** 

permitted in one direction only, there exists a characteristic length $a_0$ for buckling: if the radius, $a$, of a circular delamination crack exceeds $a_0$, then, when buckling occurs, it will be limited to the subregion $r<a_0$, and will usually not span the whole delamination; if $a$ does not exceed $a_0$, the bucking will span the whole region ($r\leqslant a$) (see Appendix A). We further discussed the problem of the post-buckling and growth of a circular delamination bridged by linear fibers in Li et al. [16,17,21] and Hu et al. [18]. In the present paper, we will study the interesting problem of buckling, post-buckling, and growth of circular delamination with a small radius ($a\leqslant a_0$) bridged by nonlinear fibers. For mathematical convenience, the composite structure is assumed to be macroscopically isotropic in the $r-\theta$ plane, Fig. 1 (e.g., quasi-isotropic laminates). The previous work on the linear bridging model is a special case of this paper.

## 2 Governing Equations of Post-Buckling for a Circular Delamination Bridged by Nonlinear Fibers

A circular delamination subjected to a uniformly distributed compression $P$ along its edge is shown in Fig. 1. Let the deflection of the delamination during axisymmetric buckling be $w(r)$. The delamination is restrained by lateral fibers which produce lateral restoring traction $q(r)$ when buckling occurs. For nonlinear fibers, we assume

$$q(r)=\alpha w(r)+\beta[w(r)]^2>0 \tag{1}$$

where $\alpha$, $\beta$ denote, respectively, linear and nonlinear parameters of the lateral restoring traction. Substituting Eq. (1) into von Karman's nonlinear plate equations ([22]) leads to

$$\left.\begin{aligned} &D\frac{1}{r}\frac{d}{dr}\left\{r\frac{d}{dr}\left[\frac{1}{r}\frac{d}{dr}\left(r\frac{dw}{dr}\right)\right]\right\}+\frac{1}{r}\frac{d}{dr}\left(rP\frac{dw}{dr}\right)+\alpha w+\beta w^2=0\\ &r^2\frac{d^2P}{dr^2}+3r\frac{dP}{dr}+\frac{Eh}{2}\left(\frac{dw}{dr}\right)^2=0\end{aligned}\right\}. \tag{2}$$

The boundary conditions are

$$w(r)|_{r=a}=0, \tag{3a}$$

$$\left.\frac{dw}{dr}\right|_{r=0,a}=0, \tag{3b}$$

$$\left.\frac{dP}{dr}\right|_{r=0}=0. \tag{3c}$$

where $P(>0)$ denotes radial pressure acting on the edge, $D=Eh^3/[12(1-\mu^2)]$ denotes the bending stiffness of the delaminated plate, $E$ and $\mu$ denote, respectively, Young's modulus and Poisson's ratio of the material, and $h$ denotes the thickness of the delamination.

## 3 Perturbation Solution

Letting

$$\xi=\frac{r}{a},\quad p(\xi)=\frac{a^2}{D}P(r),\quad \overline{w}(\xi)=\frac{\sqrt{6(1-\mu^2)}}{h}w(r),$$
$$\overline{\alpha}=\frac{a^4}{D}\alpha,\quad \overline{\beta}=\frac{a^4h}{D\sqrt{6(1-\mu^2)}}\beta. \tag{4}$$

Eqs. (2) and (3) assume the following nondimensionalized forms:

$$\left.\begin{aligned}&\xi^{-1}\{\xi[\xi^{-1}(\xi\overline{w}')']'\}'+\xi^{-1}(\xi p\overline{w}')'+\overline{\alpha}\overline{w}+\overline{\beta}\overline{w}^2=0\\ &\xi^{-1}(\xi^3p')'-\overline{w}'^2=0\end{aligned}\right\} \tag{5}$$

$$\overline{w}(1)=0, \tag{6a}$$

$$\overline{w}'(0)=\overline{w}'(1)=0, \tag{6b}$$

$$p'(0)=0. \tag{6c}$$

where the prime indicates differentiation with respect to $\xi$.

Choosing $\overline{W}_m$ (the deflection of the delamination at $r=0$) as a perturbation parameter, $\overline{w}(\xi)$ and $p(\xi)$ are expanded into series of $\overline{W}_m$ as follows:

$$\overline{w}(\xi)=\sum_{i=1}^{n}\overline{w}_i(\xi)\overline{W}_m^i, \tag{7a}$$

$$p(\xi)=\sum_{i=0}^{n}p_i(\xi)\overline{W}_m^i. \tag{7b}$$

Substituting (7) into (5) and (6) and setting the coefficients of each order of $\overline{W}_m$ to zero yields, after some manipulation,

$$\overline{w}_i^{(4)}+\frac{2}{\xi}\overline{w}_i^{(3)}+\left(p_0-\frac{1}{\xi^2}\right)\overline{w}_i''+\left(\frac{p_0}{\xi}+\frac{1}{\xi^3}\right)\overline{w}_i'+\overline{\alpha}\overline{w}_i$$
$$=f_i(\xi)\quad(i=1,2,3,\ldots), \tag{8a}$$

$$\xi^{-1}(\xi^3p_i')'=\sum_{n=1}^{i}\overline{w}_n'\overline{w}_{i-n}', \tag{8b}$$

$$\xi^2p_0''=0\quad(i=1,2,3,\ldots). \tag{8c}$$

$$\overline{w}_i(1)=0, \tag{9a}$$

$$\overline{w}_i'(0)=\overline{w}_i'(1)=0, \tag{9b}$$

$$p_{i-1}'(0)=0\quad(i=1,2,3,\ldots). \tag{9c}$$

where

$$f_i(\xi)=\sum_{n=1}^{i}\left[(p_n\overline{w}_{i-n}')'+\frac{1}{\xi}p_n\overline{w}_{i-n}'+\overline{\beta}\overline{w}_n\overline{w}_{i-n}\right]$$
$$(i=1,2,3,\ldots) \tag{10}$$

and noting that $\overline{w}_0=0$ in (8b) and (10). Letting $\xi=0$, (7a) becomes

$$\overline{W}_m=\overline{w}(0)=\overline{w}_1(0)\overline{W}_m+\overline{w}_2(0)\overline{W}_m^2+\overline{w}_3(0)\overline{W}_m^3+\ldots.$$

It follows from the above expression that

$$\overline{w}_1(0)=1, \tag{11a}$$

$$\overline{w}_i(0)=0\quad(i=2,3,\ldots). \tag{11b}$$

Obviously, Eqs. (8)–(11) with $i=1$ represent an eigenvalue problem of a linear homogeneous differential equation corresponding to the buckling problem of the delamination under consideration. It is the same as the buckling problem of a delamination bridged by linear fibers $[q(r)=\alpha w(r)]$ since the equilibrium equations and the boundary conditions do not include the nonlin-

ear factor $\overline{\beta}$. The buckling load of the delamination bridged by linear fibers has been obtained as a series of $\overline{\alpha}$ ([16,17])

$$p_0 = 14.682(1 + 0.009278\overline{\alpha} - 3.5423\times 10^{-6}\overline{\alpha}^2 - 1.01836\times 10^{-8}\overline{\alpha}^3 + \dots) \tag{12a}$$

$$\overline{w}_1(\xi) = \sum_{i=0}^{\infty} \overline{w}_{1i}(\xi)\overline{\alpha}^i \tag{12b}$$

where $\overline{w}_{1i}(\xi)$ $(i=0,1,2,3,\dots)$ are given in Appendix A.

For a bridging fiber with a quadratic hardening law, the buckling load of the delamination can be obtained from the bending theory for circular plates

$$p_0 = 14.682, \tag{13a}$$

$$\overline{w}_1(\xi) = 0.2871 + 0.7129 J_0(3.8317\xi). \tag{13b}$$

where $J_0(\ )$ denotes the Bessel function of the zero order.

The second-order perturbation Eqs. (8)–(11) are

$$\overline{w}_2^{(4)} + \frac{2}{\xi}\overline{w}_2^{(3)} + \left(p_0 - \frac{1}{\xi^2}\right)\overline{w}_2'' + \left(\frac{p_0}{\xi} + \frac{1}{\xi^3}\right)\overline{w}_2' + \overline{\alpha w_2}$$
$$= p_1'\overline{w}_1' + p_1\left(\overline{w}_1'' + \frac{1}{\xi}\overline{w}_1'\right) + \overline{\beta}\overline{w}_1^2 \tag{14a}$$

$$\xi^{-1}(\xi^3 p_1')' = 0 \tag{14b}$$

$$\overline{w}_2(1) = 0, \tag{15a}$$

$$\overline{w}_2'(0) = \overline{w}_2'(1) = 0, \tag{15b}$$

$$p_1'(0) = 0. \tag{15c}$$

The general solution of Eq. (14*b*) is $p_1(\xi) = c_1\xi^{-2} + c_2$. Since $p_1(\xi)$ should be finite at $\xi=0$, one finds $c_1=0$ and $p_1(\xi) = \text{const}$.

In addition, the solution of the homogeneous equation corresponding to (14*a*) is

$$\bar{\bar{w}}_2(\xi) = A\overline{w}_1(\xi) \quad (A \neq 0). \tag{16}$$

The necessary and sufficient condition for the existence of the particular solution is that the inner product of $f_2(\xi)$ in (10) and $\bar{\bar{w}}_1(\xi)$ in (16) should be zero ([23]), that is

$$\langle f_2(\xi), \bar{\bar{w}}_1(\xi)\rangle = \int_0^1 f_2(\xi)\cdot\bar{\bar{w}}_1(\xi)\cdot 2\pi\xi d\xi = 0.$$

From the above one finds

$$p_1 = \frac{\overline{\beta}\int_0^1[\overline{w}_1(\xi)]^3\xi d\xi}{\int_0^1[\overline{w}_1'(\xi)]^2\xi d\xi}. \tag{17}$$

After having found the value of $p_1$, the approximate solution of Eq. (14*a*) can be obtained by the weighted residual method. On the basis of the boundary conditions (15*a*,*b*), $\overline{w}_2(\xi)$ possesses the following form:

$$\overline{w}_2(\xi) = B\xi^2(1-\xi)^2. \tag{18}$$

Applying weighted integration to (14*a*) results in

$$B = \frac{630\times\int_0^1\xi^2(1-\xi)^2\cdot f_2(\xi)d\xi}{399 - 9p_0 + \overline{\alpha}}. \tag{19}$$

The third-order perturbation equations of (8)–(11) are

$$\overline{w}_3^{(4)} + \frac{2}{\xi}\overline{w}_3^{(3)} + \left(p_0 - \frac{1}{\xi^2}\right)\overline{w}_3'' + \left(\frac{p_0}{\xi} + \frac{1}{\xi^3}\right)\overline{w}_3' + \overline{\alpha w_3}$$
$$= p_1\left(\overline{w}_2'' + \frac{1}{\xi}\overline{w}_2'\right) + p_2\left(\overline{w}_1'' + \frac{1}{\xi}\overline{w}_1'\right) + p_2'\overline{w}_1' + 2\overline{\beta}\overline{w}_1\overline{w}_2 \tag{20a}$$

$$\xi^{-1}[\xi^3 p_2'(\xi)]' = [\overline{w}_1'(\xi)]^2 \tag{20b}$$

$$\overline{w}_3(1) = 0, \tag{21a}$$

$$\overline{w}_3'(0) = \overline{w}_3'(1) = 0, \tag{21b}$$

$$p_2'(0) = 0. \tag{21c}$$

The general solution of Eq. (20*b*) is

$$p_2(\xi) = S(\xi) + A_1\xi^{-2} + A_0 \tag{22}$$

where

$$S(\xi) = \int_0^{\xi}\left\{x_1^{-3}\int_0^{x_1} x[\overline{w}_1'(x)]^2 dx\right\}dx_1. \tag{23}$$

Similarly, since the value of $p_2(\xi)$ should be finite at $\xi=0$, the constant $A_1$ in (22) equals zero. From the existence condition for the particular solution of Eq. (20*a*), $A_0$ can be determined as

$$\langle f_3(\xi), \bar{\bar{w}}_1(\xi)\rangle = \int_0^1 f_3(\xi)\cdot\bar{\bar{w}}_1(\xi)\cdot 2\pi\xi d\xi = 0$$

$$A_0 = \frac{\int_0^1\overline{w}_1\cdot\{p_1(\xi\overline{w}_2')' + (\xi S\overline{w}_1')' + 2\overline{\beta}\xi\overline{w}_1\overline{w}_2\}d\xi}{\int_0^1\overline{w}_1'^2\xi d\xi}. \tag{24}$$

Finally, the deflection $\overline{w}_3(\xi)$ of the third-order perturbation can be obtained using (20*a*) and (21*a*,*b*) in the same manner as above.

In view of the fast decaying nature of the expansions (7*a*,*b*), the higher order perturbation solutions have little effect on the initial post-buckling of the delamination. Thus, the second-order expression for the deflection $\overline{w}(\xi)$ and the third-order expression for the membrane force $p(\xi)$

$$\overline{w}(\xi) = \overline{w}_1(\xi)\overline{W}_m + \overline{w}_2(\xi)\overline{W}_m^2, \tag{25a}$$

$$p(\xi) = p_0 + p_1\overline{W}_m + p_2(\xi)\overline{W}_m^2, \tag{25b}$$

are sufficiently accurate for practical use (compared with the higher order solutions, the results from three terms of $p(\xi)$ and two terms of $\overline{w}(\xi)$ show good convergence within the limits of $\overline{W}_m < 3$ or $\sigma < 1.5\sigma_{cr}$). $\overline{w}_1(\xi)$ and $\overline{w}_2(\xi)$ are given, respectively, by (13) and (18); $p_0$, $p_1$, and $p_2(\xi)$ are given, respectively, by (12), (17), and (22).

According to the continuity condition for radial displacement between the delamination and the substrate at the tip of the delamination $(r=a)$, the uniform radial stress acting on the delamination along its boundary can be obtained:

$$\sigma = \frac{D}{a^2 h}\left[\frac{p'(1)}{1-\mu} + p(1)\right]. \tag{26}$$

Letting $\overline{W}_m = 0$, the critical buckling stress for the delamination can be derived from formulas (25*b*) and (26) as

$$\sigma_{cr} = \frac{P}{h} = \frac{D}{a^2 h}p_0. \tag{27}$$

Therefore, the membrane force and bending moment exerted on the boundary of the delamination can be written as

$$P^* = \sigma h - P(a) = \sigma h - \frac{D}{a^2}p(1) = \frac{D}{a^2(1-\mu)}p'(1), \tag{28a}$$

$$M^* = D\left[\frac{d^2 w}{dr^2} + \mu\frac{1}{r}\frac{dw}{dr}\right]\Bigg|_{r=a} = \frac{Dh}{a^2\sqrt{6(1-\mu^2)}}\overline{w}''(1). \tag{28b}$$

## 4 Properties of Delamination Growth

The mode I and II stress intensity factors for a delamination under the boundary loads $P^*$ and $M^*$ are, respectively,

$$K_{\mathrm{I}}=\frac{1}{\sqrt{2}}[P^{*}h^{-1/2}\cos\omega-2\sqrt{3}M^{*}h^{-3/2}\sin\omega], \tag{29}$$

$$K_{\mathrm{II}}=\frac{1}{\sqrt{2}}[P^{*}h^{-1/2}\sin\omega+2\sqrt{3}M^{*}h^{-3/2}\cos\omega]. \tag{30}$$

where $\omega=52.1$ deg ([6,24]). Substituting (28) into (29) and (30) leads to

$$K_{\mathrm{I}}=\frac{D}{a^{2}\sqrt{h}}\left[\frac{p'(1)}{\sqrt{2}(1-\mu)}\cos\omega-\frac{\overline{w}''(1)}{\sqrt{1-\mu^{2}}}\sin\omega\right], \tag{31}$$

$$K_{\mathrm{II}}=\frac{D}{a^{2}\sqrt{h}}\left[\frac{p'(1)}{\sqrt{2}(1-\mu)}\sin\omega+\frac{\overline{w}''(1)}{\sqrt{1-\mu^{2}}}\cos\omega\right]. \tag{32}$$

The mixed-mode stress intensity ratio and the local energy release rate at the crack tip are, respectively,

$$\psi=\tan^{-1}\left[\frac{K_{\mathrm{II}}}{K_{\mathrm{I}}}\right]=\tan^{-1}\left[\frac{\sqrt{\frac{1+\mu}{2(1-\mu)}}p'(1)\sin\omega+\overline{w}''(1)\cos\omega}{\sqrt{\frac{1+\mu}{2(1-\mu)}}p'(1)\cos\omega-\overline{w}''(1)\sin\omega}\right] \tag{33}$$

$$G_{0}=\frac{1-\mu^{2}}{E}(K_{\mathrm{I}}^{2}+K_{\mathrm{II}}^{2})=\frac{D^{2}}{Eha^{4}}\overline{G}_{0} \tag{34}$$

where ($\mu=1/3$)

$$\overline{G}_{0}=\frac{Eha^{4}}{D^{2}}G_{0}=[p'(1)]^{2}+[\overline{w}''(1)]^{2}. \tag{35}$$

The global energy release rate, for delamination growth under external stress $\sigma$ is ([24])

$$G=G_{0}-\Delta G_{b}\quad\text{or}\quad\overline{G}=\overline{G}_{0}-\Delta\overline{G}_{b} \tag{36}$$

where

$$\Delta\overline{G}_{b}=\frac{Eha^{4}}{D^{2}}\Delta G_{b}=2\int_{0}^{1}\left(\overline{\alpha}\overline{w}(\xi)^{2}+\frac{2}{3}\overline{\beta}\overline{w}(\xi)^{3}\right)\xi d\xi \tag{37}$$

is the counteracting part of the energy release rate by the bridging force (see Appendix B).

Obviously, it follows from (18), (19), (10), and (17) that $\overline{w}_{2}(\xi)=p_{1}(\xi)=0$ when $\alpha=\beta=0$ (in the absence of the bridge). Substituting this result into (25) and (33), one can get the mixed-mode stress intensity ratio in the form

$$\psi=\tan^{-1}\left[\frac{K_{\mathrm{II}}}{K_{\mathrm{I}}}\right]=\tan^{-1}\left[\frac{0.14356\overline{W}_{m}\sin\omega+\cos\omega}{0.14356\overline{W}_{m}\cos\omega-\sin\omega}\right]. \tag{38}$$

The result is the same as that obtained by Hutchinson and Suo [24].

## 5 Results and Analysis

Using the above results obtained by perturbation analysis, we proceed to discuss buckling, initial post-buckling, and growth properties of a delamination bridged by nonlinear fibers.

The dependence of the central deflection $\overline{W}_{m}$ of the delamination on the uniform stress $\sigma$ is shown in Fig. 2 for the hardening-type bridging model ($\overline{\beta}=15$) with different values of the linear parameter $\overline{\alpha}$ of the bridging force. This figure shows that the $\sigma-\overline{W}_{m}$ curves form a set of nearly parallel, equidistant lines when $\overline{\alpha}$ increases by equally spaced increments. This demonstrates that the critical stress $\sigma_{cr}$, and the initial post-buckling stress $\sigma$ are proportional to the linear parameter $\overline{\alpha}$ of the bridging force.

Figures 3–5 present a set of curves describing the dependence of the central deflection $\overline{W}_{m}$, the energy release rate $G$, and the mixed-mode stress intensity ratio $\psi$ on the external stress $\sigma$ for $\overline{\alpha}=50$ and various values of $\overline{\beta}$. In Fig. 3, the slope of the curve $\overline{W}_{m}-\sigma/\sigma_{cr}$ at the origin has the same sign as $\overline{\beta}$, and decreases with increasing $\overline{\beta}$. In the case of $\overline{\beta}>0$, $\overline{W}_{m}$ increases with increasing $\sigma$, which indicates that the equilibrium state of the delamination during initial post-buckling is stable; and $\overline{W}_{m}$ decreases with increasing $\overline{\beta}$ when $\sigma=$const. (no change), which indicates that the residual strength of the composite increases with increasing $\overline{\beta}$. However, when $\overline{\beta}<0$, there exists an unstable region during initial post-buckling of the delamination, which means that the external stress $\sigma$ maintaining the equilibrium of the delamination is smaller than the critical stress $\sigma_{cr}$ after buckling. In fact, a sudden change (jump) of the central deflection $\overline{W}_{m}$ occurs at the instant of buckling. Figure 3 shows that $\overline{W}_{m}$ will change suddenly from 0 to 1.2 for $\overline{\beta}=-20$ at the instant of delamination buckling if the critical stress $\sigma_{cr}$ is maintained. For the linearly bridged delamination ($\overline{\beta}=0$), the slope of the $\overline{W}_{m}-\sigma/\sigma_{cr}$ curve at the origin is infinite (see the dashed line in Fig. 3). Hence, the equilibrium of the linearly bridged delamination at the instant of buckling moment forms a boundary between stable and unstable states. Also for $\overline{\beta}=-20$, the global energy release rate changes suddenly from 0 to 5.2 (Fig. 4(*a*)), the local energy release rate changes from 0 to 14 (Fig. 4(*b*)), and the mixed mode stress intensity ratio changes from $-38$ deg to $-51.5$ deg (Fig. 5). (This means that the state of $|K_{1}|>|K_{2}|$ for the mixed crack converts suddenly to that of $|K_{1}|<|K_{2}|$.) This again illustrates the instability of delamination growth for the softening-type bridging model at the instant of buckling. Figure 4 shows that for the hardening-type bridging model, the energy release rate $G$ decreases with increasing nonlinear parameter $\overline{\beta}$ for a given stress $\sigma$, especially at the beginning of post-buckling. Figure 5 indicates that the ratio $\psi$ is independent of the nonlinear parameter $\overline{\beta}$ at the instant of buckling for a given linear parameter $\overline{\alpha}$ ($\overline{\alpha}=50$,

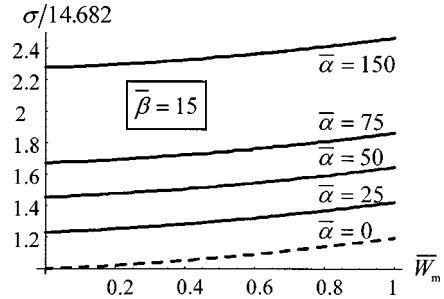



**Fig. 2 Curves of $\sigma-\overline{W}_{m}$ for the hardening-type bridge model with different values of $\overline{\alpha}$**

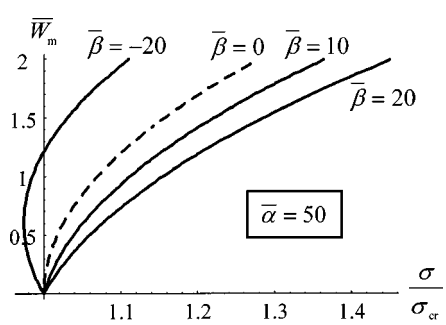



**Fig. 3 Curves of $\overline{W}_{m}-\sigma/\sigma_{cr}$ for $\overline{\alpha}=50$ and different values of $\overline{\beta}$**

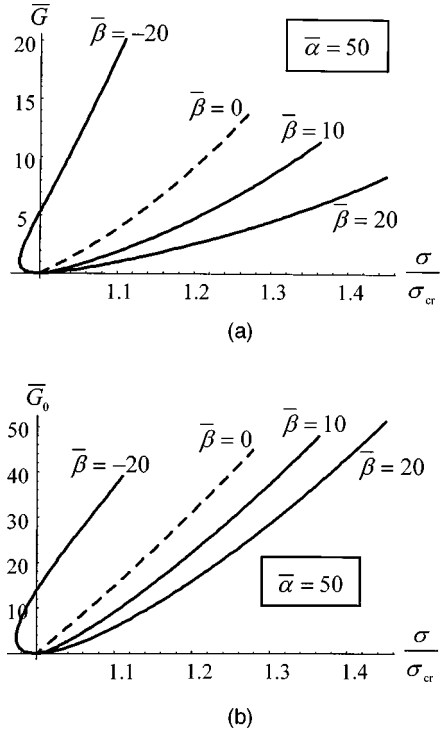



**Fig. 4 Curves of $G-\sigma/\sigma_{cr}$ for different values of $\bar{\beta}$. (*a*) Global energy release rate, (*b*) local energy release rate.**

$\psi\approx-38$ deg). In addition, the dominance of mode II for the mixed-mode crack growth weakens with increasing nonlinear parameter $\bar{\beta}$ for a given stress $\sigma$.

Figure 6 presents the curves of $\bar{W}_m-\sigma/\sigma_{cr}$ for the delamination bridged by fibers of a quadratically hardening type ($\bar{\alpha}=0$, $\bar{\beta}\geqslant 0$) during initial post-buckling. The slope of the $\bar{W}_m-\sigma/\sigma_{cr}$ curve at the origin decreases gradually from infinity as $\bar{\beta}$ increases gradually from zero. For a given $\sigma$, the central deflection of the delamination decreases gradually with increasing $\bar{\beta}$, which indicates that residual stiffness of the composite increases. Figure 7 shows the dependence of the energy release rate $G$ of the delamination bridged by fibers with quadratically hardening on $\sigma$ during the initial post-buckling, illustrating that $G$ decreases gradually with increasing $\bar{\beta}$. Figure 8 presents the curve of $\psi-\sigma/\sigma_{cr}$, which shows that the potential for mode II delamination growth decreases with increasing $\bar{\beta}$.

In the above analysis, the deformation of the delamination during the initial post-buckling stage was determined for a given radius $a$ and different bridging parameters, $\alpha$ and $\beta$. The effect of changing fracture parameters due to applied stress field at the

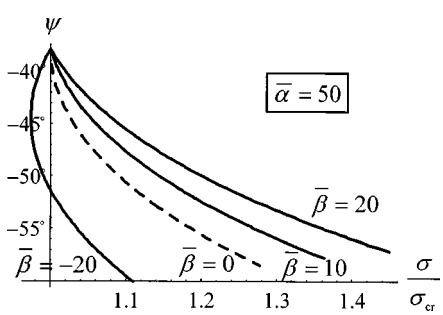



**Fig. 5 Curves of $\psi-\sigma/\sigma_{cr}$ for different values of $\bar{\beta}$**

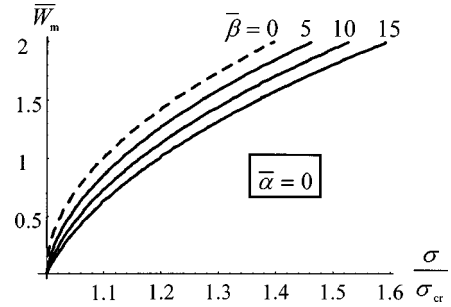



**Fig. 6 Curves of $\bar{W}_m-\sigma/\sigma_{cr}$ for the delamination bridged by fibers with quadratically hardening type**

delamination tip was also elucidated. To further investigate the characteristics of delamination propagation, the following variables are defined:

$$\hat{a}=\frac{a}{h},\quad \hat{\alpha}=\frac{\alpha h^4}{D},\quad \hat{\beta}=\frac{\beta h^5}{D\sqrt{6(1-\mu^2)}},\quad \hat{G}=\frac{Eh^5}{D^2}G. \tag{39}$$

Using the results obtained in Sections 3 and 4, the change in the energy release rate and the mixed-mode stress intensity ratio at the crack tip can be determined for different delamination radii under constant stress field (Figs. 9–14).

Fixing the applied stress $\sigma=1.2\sigma_{cr}$ during the delamination post-buckling stage, the curves of $\hat{G}-\hat{a}$ and $\psi-\hat{a}$ are plotted for the different bridging parameters, $\hat{\alpha}$ and $\hat{\beta}$ (Figs. 9–13). There exist three types of $\hat{G}-\hat{a}$ curves: Type 1: monotonously increasing, Type 2: monotonously decreasing, and Type 3: an initial increase followed by decrease. When the values of $\hat{\alpha}$ and $\hat{\beta}$ are comparatively small, but not zero, the curve $\hat{G}-\hat{a}$ is composed of an ascending region and a descending region, i.e., the Type 3

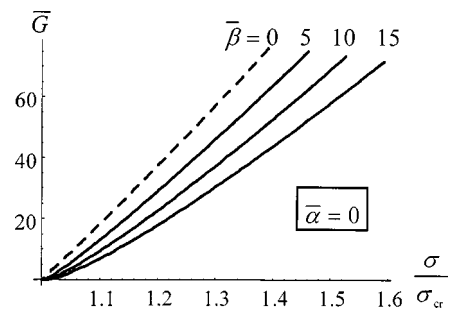



**Fig. 7 Curves of $G-\sigma/\sigma_{cr}$ for the delamination bridged by fibers of a quadratically hardening type**

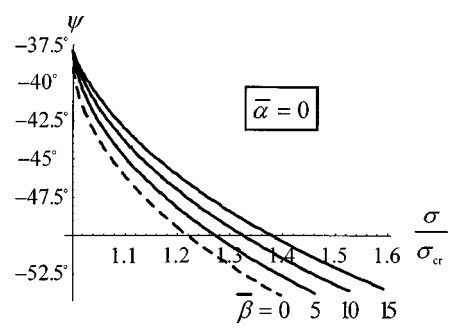



**Fig. 8 Curves of $\psi-\sigma/\sigma_{cr}$ for the delamination bridged by fibers of a quadratically hardening type**

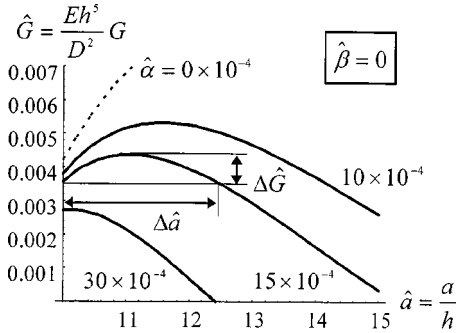



**Fig. 9 Curves of $\hat{G}-\hat{a}$ for $\hat{\beta}=0$ and different values of $\hat{\alpha}$**

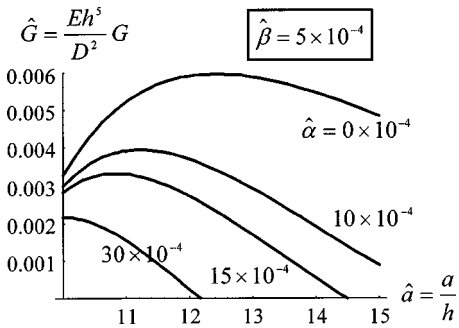



**Fig. 10 Curves of $\hat{G}-\hat{a}$ for $\hat{\beta}=5\times10^{-4}$ and different values of $\hat{\alpha}$**

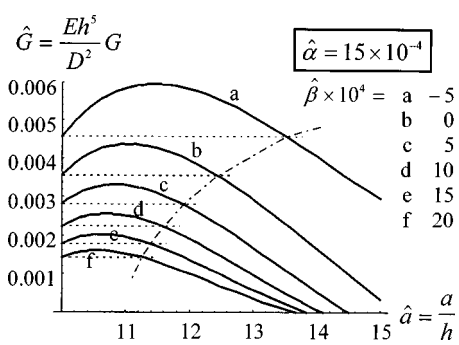



**Fig. 11 Curves of $\hat{G}-\hat{a}$ for $\hat{\alpha}=5\times10^{-4}$ and different values of $\hat{\beta}$**

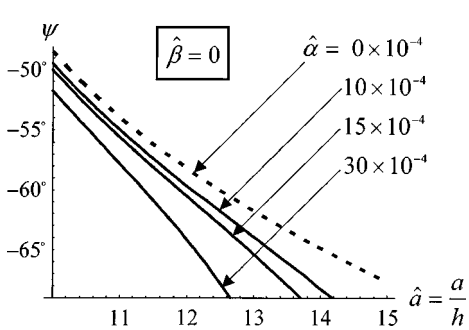



**Fig. 12 Curves of $\psi-\hat{a}$ for $\hat{\beta}=0$ and different values of $\hat{\alpha}$**

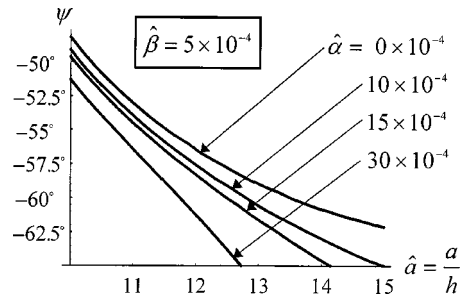



**Fig. 13 Curves of $\psi-\hat{a}$ for $\hat{\beta}=5\times10^{-4}$ and different values of $\hat{\alpha}$**

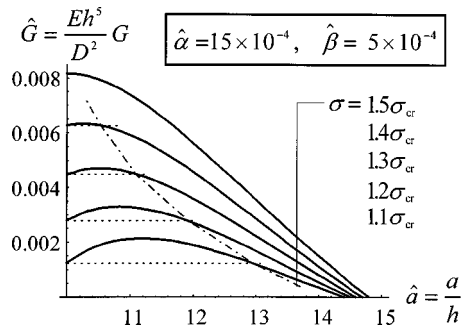



**Fig. 14 Curves of $\hat{G}-\hat{a}$ for $\hat{\alpha}=15\times10^{-4}$, $\hat{\beta}=5\times10^{-4}$ and different values of $\sigma$**

curve. This demonstrates that during the post-buckling stage ($\sigma = 1.2\sigma_{cr}$), if the energy release rate reaches a critical value (i.e., fracture toughness) keeping $\sigma$ fixed, the delamination (with initial radius $\hat{a}=10$) will reach a stable state after experiencing a dynamic growth in a certain range. As the bridging parameter $\hat{\alpha}$ increases, the dynamic effect $\Delta\hat{G}$ and the range $\Delta\hat{a}$ of unstable delamination growth decrease (see the curve for $\hat{\alpha}=15\times10^{-4}$ in Fig. 9). For an unbridged delamination (dashed line in Fig. 9), i.e., $\hat{\alpha}=\hat{\beta}=0$, the energy release rate $\hat{G}$ increases monotonously with increasing delamination radius $\hat{a}$, which illustrates that the delamination growth will be catastrophic once it initiates. When $\hat{\alpha}$ increases to a certain value, the curve $\hat{G}-\hat{a}$ descends monotonously, indicating that the delamination is stable, i.e., further growth will occur only if the stress $\sigma$ is increased. This conclusion is the same as the one drawn in our previous work dealing with the linear bridging model ([17,18,21]), that is, the presence of the fiber bridge impedes catastrophic delamination growth.

Figures 9, 10, and 11 also show that the energy release rate during delamination growth decreases with increasing nonlinear parameter $\hat{\beta}$, which increases the toughness of the initial delamination growth. Moreover, the range $\Delta\hat{a}$ and the dynamic effect $\Delta\hat{G}$ of unstable delamination growth significantly decrease with increasing $\hat{\beta}$. That is, increasing the nonlinear parameter improves the stability of the delamination growth (dashed line in Fig. 11 is the boundary between the stable and unstable regions). Figs. 12 and 13 show the variation of the mixed mode stress intensity ratio $\psi$ corresponding to Figs. 9 and 10, respectively.

The relation between the energy release rate and the delamination radius under different stress fields for $\hat{\alpha}=15\times10^{-4}$, $\hat{\beta}=5\times10^{-4}$ is shown in Fig. 14. As the energy release rate increases with increasing stress field, the ascending part of the curve $\hat{G}-\hat{a}$ shortens and the curve changes gradually to the monoto-

nously descending type. In other words, for the bridged delamination, the higher the material fracture toughness, the higher the stability of the delamination growth, and the smaller its range and dynamic effect of the unstable growth.

## 6 Conclusion

Based on to the above analysis, the following conclusions can be drawn:

1 There exists an unstable region for the softening-type fiber bridging model during initial post-buckling. At the beginning of post-buckling, the central deflection of the delamination, the energy release rate at the delamination tip and the mixed-mode stress intensity ratio experience sudden changes. Thereafter, post-buckling becomes stable. This phenomenon results likely in unstable delamination growth for materials with lower fracture toughness when the delamination buckles rather than post-buckles. The initial post-buckling for the hardening-type model is stable. The residual stiffness of the composite increases with an increase of the nonlinear parameter $\bar{\beta}$. The linear bridging model forms the boundary between stable and unstable states during initial postbuckling.

2 The critical buckling stress of a delamination bridged by quadratically hardening fibers is the same as that of an unbridged delamination. The potential for mode II delamination growth decreases with increasing nonlinear bridging parameter.

3 Bridging force changes the catastrophic nature of delamination growth and increases the stability of delamination. With the bridging parameters increasing and the energy release rate decreasing, the initial delamination growth toughness is improved. In addition, the range and the dynamic effect of unstable delamination growth diminish or disappear as the bridging parameters increase.

4 For the bridged delamination, the higher the material fracture toughness, the higher the stability of the delamination growth, and the smaller the range and dynamic effect of its unstable growth.


## Acknowledgment

This research has been supported by the China Postdoctoral Science Foundation.


## Appendix A

Expressions $\bar{w}_{1i}(\xi)$ $(i=0,1,2,3,\ldots)$ for the delamination buckling deflection $\bar{w}_1(\xi)$ appearing in Eq. (12*b*) are as follows:

$$\bar{w}_{10}(\xi)=0.2871+0.7129\text{Bessel}J(0,3.8317\xi)$$

$$\begin{aligned}\bar{w}_{11}(\xi)=\xi^2(1-\xi^2)^2(&-0.00423354+0.00206874\xi^2\\&-0.000465723\xi^4+0.0000496396\xi^6+0.0000269719\xi^8\\&-0.0000171247\xi^{10}+\ldots)\end{aligned}$$

$$\begin{aligned}\bar{w}_{12}(\xi)=\xi^2(1-\xi^2)^2(&-0.0000125403+0.0000139607\xi^2\\&-4.5087\times10^{-6}\xi^4+8.133\times10^{-7}\xi^6-1.36957\times10^{-7}\xi^8\\&+2.97274\times10^{-8}\xi^{10}+\ldots)\end{aligned}$$

$$\begin{aligned}\bar{w}_{13}(\xi)=\xi^2(1-\xi^2)^2(&-2.73152\times10^{-8}+4.578\times10^{-8}\xi^2\\&-2.56008\times10^{-8}\xi^4+6.09799\times10^{-9}\xi^6\\&-8.94374\times10^{-10}\xi^8+9.72303\times10^{-11}\xi^{10}+\ldots)\end{aligned}$$

$$\begin{aligned}\bar{w}_{14}(\xi)=\xi^2(1-\xi^2)^2(&-3.4593\times10^{-11}+9.19294\times10^{-11}\xi^2\\&-7.79477\times10^{-11}\xi^4+2.94059\times10^{-11}\xi^6\\&-5.43832\times10^{-12}\xi^8+5.37387\times10^{-13}\xi^{10}+\ldots)\end{aligned}$$

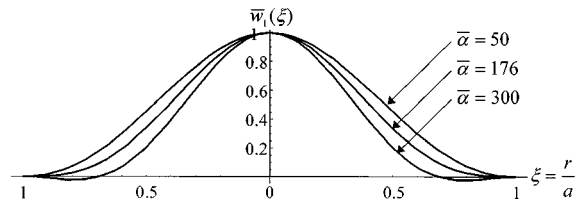



**Fig. 15 The deflections of the delamination buckling for different $\bar{\alpha}$**

$$\begin{aligned}\bar{w}_{15}(\xi)=\xi^2(1-\xi^2)^2(&3.39783\times10^{-14}+5.70367\times10^{-14}\xi^2\\&-1.23455\times10^{-13}\xi^4+7.69192\times10^{-14}\xi^6\\&-2.16369\times10^{-14}\xi^8+2.50684\times10^{-15}\xi^{10}+\ldots)\end{aligned}$$

$$\begin{aligned}\bar{w}_{16}(\xi)=\xi^2(1-\xi^2)^2(&3.69661\times10^{-16}-4.15562\times10^{-16}\xi^2\\&+9.06572\times10^{-17}\xi^4+6.7624\times10^{-17}\xi^6\\&-4.19366\times10^{-17}\xi^8+7.21486\times10^{-18}\xi^{10}+\ldots)\end{aligned}$$

$$\begin{aligned}\bar{w}_{17}(\xi)=\xi^2(1-\xi^2)^2(&1.28208\times10^{-18}-2.12341\times10^{-18}\xi^2\\&+1.31773\times10^{-18}\xi^4-3.6009\times10^{-19}\xi^6\\&+2.73422\times10^{-20}\xi^8+3.69499\times10^{-21}\xi^{10}+\ldots)\end{aligned}$$

$$\begin{aligned}\bar{w}_{18}(\xi)=\xi^2(1-\xi^2)^2(&2.61627\times10^{-21}-5.72007\times10^{-21}\xi^2\\&+4.77626\times10^{-21}\xi^4-2.04609\times10^{-21}\xi^6\\&+4.77547\times10^{-22}\xi^8-5.19869\times10^{-23}\xi^{10}+\ldots)\end{aligned}$$

$$\begin{aligned}\bar{w}_{19}(\xi)=\xi^2(1-\xi^2)^2(&1.12803\times10^{-24}-7.58201\times10^{-24}\xi^2\\&+9.65879\times10^{-24}\xi^4-5.64269\times10^{-24}\xi^6\\&+1.75687\times10^{-24}\xi^8-2.47349\times10^{-25}\xi^{10}+\ldots)\end{aligned}$$

$\ldots$

$\ldots$

Letting $\bar{w}_1'(1)=\Sigma_{i=0}^{\infty}\bar{w}_{1i}'(1)\bar{\alpha}^i=0$, the characteristic delamination radius for initial buckling can be obtained:

$$\bar{\alpha}^*=\frac{a_0^4}{D}\alpha=176.502 \quad\text{or}\quad a_0=3.6499\sqrt{D/\alpha}. \qquad (A1)$$

Its physical meaning is shown in Fig. 15.

## Appendix B

In the post-buckled delamination state, the bridging force (traction due to the bridge) acting on the delamination surface will bring about closure of delamination cracking. It contributes to the fracture toughness $\Delta G_b$, i.e., counteracts a part of the energy release rate at the crack tip, which can be determined by calculating the virtual work of the bridging force when the delamination propagates virtually. Assuming that the delamination propagates to $a+\delta a$, the virtual displacement of the delamination is $\delta w$, and the virtual work $\delta G_b$ (per unit length) of the bridging force is

$$\begin{aligned}\delta G_b&=\frac{1}{2\pi a}\left(\int_0^{a+\delta a}q(w+\delta w)2\pi r\,dr-\int_0^a qw2\pi r\,dr\right)\\&=\frac{1}{a}\left(\int_0^a q\frac{\delta w}{\delta a}r\,dr+[q(w+\delta w)r]\Big|_{r=a}\right)\delta a. \qquad (B1)\end{aligned}$$

Substituting (1) into the above equation and considering the displacement boundary condition $w(a)=0$ and the relation

$$\lim_{\delta a \to 0} \frac{\delta w}{\delta a} = \frac{\partial w}{\partial a} = -\frac{r}{a}\frac{\partial w}{\partial r} \tag{B2}$$

yields, after some simplification,

$$\Delta G_b = \lim_{\delta a \to 0} \frac{\delta G_b}{\delta a} = \frac{1}{a}\int_0^a q \frac{\partial w}{\partial a} r dr = \frac{1}{a^2}\int_0^a \left( \alpha w^2 + \frac{2}{3}\beta w^3 \right) r dr. \tag{B3}$$

Substituting (4) into ($B3$) leads to

$$\Delta \overline{G}_b = \frac{Eha^4}{D^2}\Delta G_b = 2\int_0^1 \left( \overline{\alpha w}(\xi)^2 + \frac{2}{3}\overline{\beta w}(\xi)^3 \right) \xi d\xi. \tag{B4}$$

## References


[1] Kardomateas, G. A., and Schmueser, D. W., 1988, ''Buckling and Postbuckling of Delaminated Composites under Compressive Loads Including Transverse Shear Effects,'' AIAA J., **26**, pp. 337–343.

[2] Chai, H., Babcock, C. D., and Knauss, W. G., 1981, ''One Dimensional Modeling of Failure in Laminated Plates by Delamination Buckling,'' Int. J. Solids Struct., **17**, pp. 1069–1083.

[3] Bottega, W. J., and Maewal, A., 1983, ''Delamination and Buckling in Laminates,'' ASME J. Appl. Mech., **50**, pp. 185–189.

[4] Evans, A. G., and Hutchinson, J. W., 1984, ''On the Mechanics of Delamination and Spalling in Compressed Films,'' Int. J. Solids Struct., **20**, pp. 455–466.

[5] Yin, W.-L., 1985, ''Axisymmetric Buckling and Growth of a Circular Delamination in a Compressed Laminate,'' Int. J. Solids Struct., **21**, pp. 503–514.

[6] Chai, H., 1990, ''Three Dimensional Fracture Analysis of Thin Film Debonding,'' Int. J. Fract., **46**, pp. 237–256.

[7] Chai, H., 1990, ''Buckling and Post-Buckling Behavior of Elliptical Plates: Part I—Analysis, Part II—Results,'' ASME J. Appl. Mech., **57**, pp. 981–994.

[8] Suo, Z., 1990, ''Delamination Specimens for Orthotropic Materials,'' ASME J. Appl. Mech., **57**, pp. 627–634.

[9] Peck, S. O., and Springer, G. S., 1991, ''The Behavior of Delaminations in Composite Plates: Analytical and Experimental Results,'' J. Compos. Mater., **25**, pp. 907–929.

[10] Kardomateas, G. A., 1993, ''The Initial Postbuckling and Growth Behavior of Internal Delaminations in Composite Plates,'' ASME J. Appl. Mech., **60**, pp. 903–910.

[11] Giannakopoulos, A. E., Nilsson, K. F., and Tsamasphyros, G., 1995, ''The Contact Problem at Delamination,'' ASME J. Appl. Mech., **62**, pp. 989–996.

[12] Zhang, X., and Yu, S., 1996, ''An Analysis of Axisymmetric Buckling and Growth of Circular-Shaped Delamination,'' Acta Mech. Solida Sinica, **9**, pp. 201–209.

[13] Shu, D., and Mai, Y.-W., 1993, ''Effect of Stitching on Interlaminar Delamination Extension in Composite Laminates,'' Compos. Sci. Technol., **49**, pp. 165–171.

[14] Cox, B. N., 1994, ''Delamination and Buckling in 3D Composites,'' J. Compos. Mater., **28**, pp. 1114–1126.

[15] Cox, B. N., and Marshall, D. B., 1994, ''Concepts for Bridged Cracks in Fracture and Fatigue,'' Overview No. 111, Acta Metall. Mater., **42**, pp. 341–363.

[16] Li, S., Huang, Y., Hu, Y., and Zhong, W., 1997, ''Analysis on Buckling and Postbuckling of Delamination in 3D Composites,'' Acta Mech. Solida Sinica, **10**, pp. 212–219.

[17] Li, S., Hu, Y., Huang, Y., and Zhong, W., 1997, ''The Fracture Analysis of a Bridged Delamination in 3D Composites,'' Acta Mech. Solida Sinica, **18**, No. 3, pp. 94–99 (in Chinese).

[18] Hu, Y., Huang, Y., Li, S., and Zhong, W., 1999, ''The Effects of Bridging in a 3D Composite on Buckling, Postbuckling and Growth of Delamination,'' Arch. Appl. Mech., **69**, pp. 419–428.

[19] Massabo, R., and Cox, B. N., 1999, ''Concepts for Bridged Mode II Delamination Cracks,'' J. Mech. Phys. Solids, **47**, No. 6, pp. 1265–1300.

[20] Li, S., Huang, Y., Hu, Y., and Li, D., 1998, ''Effects of Fiber Bridge on the Delamination Buckling in 3D Composites,'' Eng. Mech., **15**, No. 4, pp. 127–132 (in Chinese).

[21] Li, S., Li, D., Huang, Y., and Hu, Y., 1998, ''Analysis on Stability Characteristics of Quasi-Static Growth of a Bridged Delamination in 3D Composites,'' Acta Mater. Compos. Sinica, **15**, No. 4, pp. 143–150 (in Chinese).

[22] Xu, Z., 1982, *Theory of Elasticity (II)*, Higher Education Press, Beijing, (in Chinese).

[23] Gu, D., 1993, *Perturbation Method*, Higher Education Press, Beijing (in Chinese).

[24] Hutchinson, J. W., and Suo, Z., 1992, ''Mixed Mode Cracking in Layered Materials,'' Adv. Appl. Mech., **29**, pp. 63–191.

**E. J. Berger**[1]
Assistant Professor,
Department of Mechanical, Industrial, and
Nuclear Engineering,
University of Cincinnati,
Cincinnati, OH 45221-0072
Assoc. Mem. ASME

**M. R. Begley**
Assistant Professor,
Department of Mechanical Engineering,
University of Connecticut,
Assoc. Mem. ASME

**M. Mahajani**
Research Assistant,
Department of Mechanical, Industrial, and
Nuclear Engineering,
University of Cincinnati,
Cincinnati, OH 45221-0072

# Structural Dynamic Effects on Interface Response: Formulation and Simulation Under Partial Slipping Conditions

*A new formulation for dynamic sliding contact problems with partial slipping is presented and used to investigate the influence of structural dynamic response on interface behavior. The mixed differential-algebraic equation (MDAE) approach uses differential equations to describe the slipping dynamics and algebraic (constraint) equations to model interfacial sticking. An efficient method for solving the case of partial interface slipping has been developed, and special consideration has been given to the changing equations of motion (at the transition from stick-to-slip and slip-to-stick). An example is presented for the case of an elastic block pressed into a rigid foundation and loaded with cyclic tangential tractions at the top of the block. The full elastodynamic transient simulations illustrate that interface slip response is a strong function of loading frequency, reaching a maximum when the external loading frequency is near the theoretical (shear-mode) natural frequency of the structure.* [S0021-8936(00)02204-2]

## 1 Introduction

Many mechanical systems rely upon friction elements for proper and efficient function. The details of surface interactions in localized regions of the nominal contact area are crucial to the efficiency and durability of the components. For example, brakes and clutches transmit power across friction interfaces, with unfavorable local interface conditions often resulting in ''hot-spotting'' ([1]). Machine tools slideways, which guide cutting tools with accuracy and precision, and joints in robotic manipulators experience performance limitations due to small-scale stick-slip (see, for example, [2]). Riveted and bolted connections and turbine blade mounts for power generation or aerospace applications are susceptible to fretting fatigue failure, with local interface mechanics playing an important role (see, for example, [3]) on crack initiation and propagation. Frequency dependence of fretting experiment results has also been demonstrated in the literature by Söderberg, Bryggman, and McCullough [4], Bryggman and Söderberg [5], and Schouterden, Blanpain, Celis and Vingsbo [6].

As such, understanding the implications of interfacial friction in mechanical systems is important—particularly the impact of system dynamics on interface response. Coupling between in-plane (tangential) and out-of-plane (normal) response in sliding systems plays an important role in interface behavior. The importance of normal motion in contact systems with friction was first elaborated by Tolstoi [7], in which he described synchronous tangential and normal motion during stick-slip vibrations. Further efforts to determine the role of normal motion on friction system performance include those of Budanov, Kudinov, and Tolstoi [8], Nayak [9], and Gray and Johnson [10]. Contemporary models of friction-excited vibrations and friction interface phenomena still generally rely upon lumped parameter *discrete* models, although many include both translational and angular degrees-of-freedom (Hess and Soom [11], Oden and Martins [12], Martins, Oden, and Simoes [13], and Tworzydlo, Becker, and Oden [14]). Other recent works have included multiple inertia elements as well ([15]). Although various mechanisms of normal-tangential coupling have been established and analyzed, these lumped parameter dynamic models do not allow detailed examination of the spatial variations in contact pressure, friction tractions, and stick or slip at the contact interface. Furthermore, some stick-slip models actually use a friction regularization procedure, in which the friction discontinuity at zero relative velocity is approximated by a smooth, continuously differentiable function (e.g., arctan-type models). This approximation precludes stick-slip behavior (which requires a multivalued friction force at zero relative velocity), but it can provide some interesting insight into friction-excited system dynamics.

Models capturing the highly localized details of interface behavior demand continuous treatments of the components. For example, an approach to continuous dynamic contact modeling is presented by Oden and Martins [12]. A third-body approach, in which the contact interface possesses constitutive properties, is used in numerical simulations of contact systems. The study employs friction modeling which approximates the discontinuous behavior at zero relative velocity using a regularization procedure based upon a small parameter $\epsilon$. The result of the regularization is the nonexistence of the zero velocity sticking state, in favor of a near-zero velocity condition which allows interfacial creep with very small relative velocity (smaller than the regularization parameter). The authors demonstrate some spatial dependence of friction and normal force results, but focus more on other issues, including rotational degrees-of-freedom, and the dynamic aspects of frictional contact such as eigenvalues and successive linearization around the steady sliding equilibrium point. Other related works include Martins et al. [13] and the low-order continuum model of Mahajani [16].

A variety of rate and state-dependent friction models have been applied to investigate the stability of steady-sliding conditions, as applied to geomechanics and unstable fault slip (Dieterich [17], Dieterich [18], Dieterich [19], Rice and Ruina [20], Gu, Rice, Ruina, and Tse [21]). In these models, frictional shear stress is viewed as a function of slip rate (i.e., relative velocity), slip his-

[1]To whom all correspondence should be addresed.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, May 21, 1999; final revision, June 16, 2000. Associate Technical Editor: A. A. Ferri. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME Journal of Applied Mechanics.

 **Copyright © 2000 by ASME** 

tory, and possibly other interfacial ''states.'' As a result, changes in sliding velocity produce two components of interfacial shear stress response: one instantaneous, and one evolutionary. After sliding over some critical distance $d_c$, the interfacial shear stress achieves a steady-state value and remains there until the next disturbance in sliding velocity. Step changes in normal stress show a similar instantaneous-plus-evolutionary response of shear stress. The implications of these models on systems with smoothly varying velocity or normal force are unclear. The models have been used to examine stability of steady sliding conditions, as applied to geomechanics and unstable fault slip in a single-degree-of-freedom system ([21]) as well as continuous contact elastic systems ([22]).

Recently, Adams, in a series of articles, has examined continuous elastic components in sliding contact ([13,24,25]). Using elastic half-spaces in dry contact as a candidate system, he concludes that the material response along the interface, and in particular slip wave propagation, can be a potentially destabilizing mechanism for steady sliding. Adams [23] used analytical tools to determine conditions under which the onset of loss-of-contact or stick-slip interface response can be expected; actual stick-slip solutions are not presented. Adams [24] demonstrates that interface slip waves can account for the apparent velocity-dependence of friction coefficient measurements. Finally, in Adams [25], relative motion between the half-spaces without gross slipping at the interface is demonstrated for cases in which separation pulses propagate along the interface. In each case, he uses a velocity-independent friction coefficient, and he makes no distinction between static and kinetic friction.

A final important contribution has been the work Menq and co-workers, who have analyzed microslip (regions of interface slip which are small relative to the contact size) in friction dampers used in aerospace turbine blades. Menq has investigated point contact stick/slip models ([26,27]) using a linearization technique (one-term harmonic balance) to develop approximate analytical solutions for the blade response. In a more detailed analysis, Menq again utilized the harmonic balance approach in a continuous blade-damper model composed of an elastic bar, retaining spring, and contact shear layer with spatially constant normal interface pressure ([28,29]). The contact normal pressure across the damper interface is assumed constant, and microslip is implemented using an elastoplastic shear layer at the contact interface. Menq's results clearly illustrate the importance of considering microslip in blade life calculations, but rely substantially on judicious choice of microslip model parameters (shear layer stiffness and strength, equivalent bar modulus and area) to achieve reasonable quantitative agreement with experimental results.

In light of the importance of accurate determination of interface slip amplitude under a variety of contact conditions, this paper presents numerical solutions of the piecewise nonlinear dynamic equations of motion for partial slip dynamic contact problems in a finite domain. This work allows zero relative velocity response of the contacting nodes, and computes the consequent effect on tangential interface traction. This approach does not rely upon a friction regularization or smoothing procedure, and it includes detailed spatial discretization of the contact interface. The numerical tool described in this paper is capable of quantifying interface slip amplitude under truly dynamic operating conditions which include the full range of contact scenarios, including loss of contact. Sample results are presented for an idealized loading scenario, whereby the normal pressure across the contact interface due to a constant far-field pressure (i.e., away from the contact interface) is high enough to prevent loss of contact. It should be emphasized that the temporal and spatial variation in normal pressure due to system dynamics is explicitly accounted for. Despite the apparent simplicity of the geometry and loading, the results indicate that component dynamic response can fundamentally affect the interface behavior, including both slip amplitude and size of slipping region.

## 2 Mathematical Model

The motivating system for this analysis is shown in Fig. 1(*a*). This system originates in the traditional one-degree-of-freedom rigid-body system used in the literature to study a variety of friction-excited systems. However, in this research, the sliding component is considered to be a *continuous* elastic system, attached to ground via a discrete spring and damper. The numerical approach outlined here is for a two-dimensional body in contact with a rigid surface that translates at velocity $V_{\mathrm{ref}}$. The body is loaded by a distributed normal pressure, $P_{\mathrm{app}}(x,t)$ and a distributed tangential traction $Q_{\mathrm{app}}(x,t)$ along its top edge; both of the applied loads can be functions of time $t$. Interface friction is described by a general frictional shear stress constitutive relation. The properties of the system include Young's modulus $E$ and Poisson's ratio $\nu$, as well as the total inertia of the moving surface $I$. Note the two most crucial features of the $n$-degree-of-freedom discretized system, shown in Fig. 1(*b*): (i) dynamic effects and localized interface behavior are captured, and (ii) normal motion is included in a physically realistic way. In the formulation presented here, loss of contact is not considered; this phenomenon is not relevant to the scenarios of primary interest, which involve large normal pressures generally prohibiting loss of contact. Still, variations in normal pressure at the interface due to dynamic re-

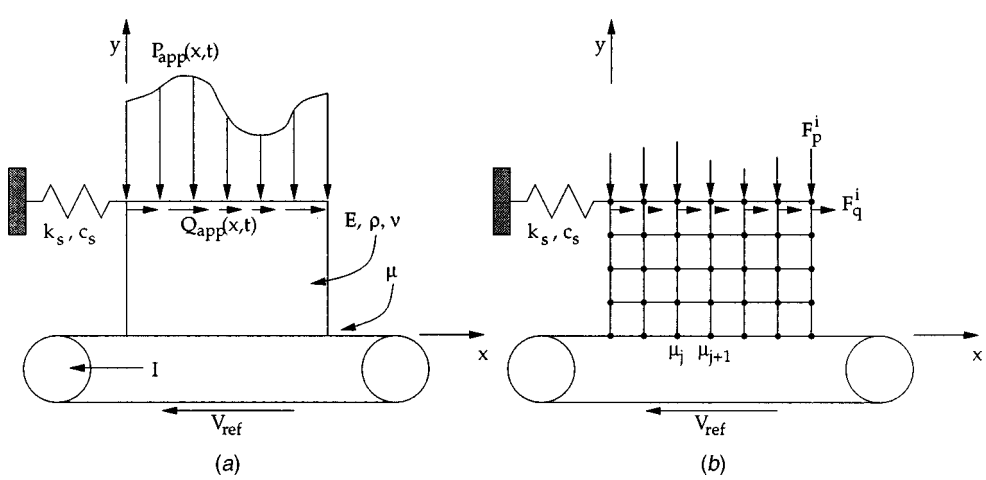



**Fig. 1 System for partial slip contact dynamics study: (*a*) continuous contact physical system, (*b*) discrete system for partial slip contact study**

sponse are rigorously accounted for. The contact pressure is monitored to insure the assumption of continuous contact is valid, and the approach here can be generalized to include loss of contact. Shown on the figure are discrete applied forces at node $i$ ($F_p^i$ and $F_q^i$, in the normal and tangential directions) as well as the nodal interface friction coefficients $\mu_j$ and $\mu_{j+1}$ for the neighboring $j$ and $j+1$ interface nodes.

**2.1 Continuous Description and Discrete Equations of Motion.** The governing equation for the dynamic response of the continuous system can be derived in the usual manner from the principal of virtual work, with the inertial and damping terms included as body forces. The resulting statement is

$$\int_V (\sigma_{ij,j}+\rho\ddot{u}_i+c_m\dot{u}_i)\,\delta u_i dV=\int_S T_i\,\delta u_i dS \tag{1}$$

in which $u_i$ is displacement, $c_m$ is a material damping parameter, and $T_i$ is a surface traction. The friction shear stress in a slipping region along the interface can be represented in a general way by a rate and state-dependent friction law (for example, Rice and Ruina [22], Rice [20], or Ben-Zion and Rice [30]):

$$\tau_f=\tau_f(V_{\text{rel}},\theta_1,\theta_2,\ldots) \tag{2}$$

where $\theta_i$ are the interface ''states,'' each with its own equation governing its rate of change. Note that this approach may incorporate state variables that depend on the history of deformation, by defining appropriate evolution equations for $\theta_i$; thus Eq. (2) is an implicit function of time. The integral of the frictional shear is related to the reference velocity $V_{\text{ref}}$ of the moving surface as follows:

$$I\dot{V}_{\text{ref}}+c_bV_{\text{ref}}=\int_0^L \tau_f\,dx \tag{3}$$

where $L$ is the interface length, and $c_b$ is the damping associated with the surface motion (possibly attributed to resistance in the bearings of the pulleys, for example).

In the ensuing discussions, it will be convenient to reference an effective friction coefficient, instead of the shear stress constitutive law described above. The effective friction coefficient is defined as

$$\mu_{\text{eff}}\begin{cases} =\dfrac{\tau_f}{\sigma_{yy}}; & V_{\text{rel}}\neq 0\\ \leqslant(\mu_{\max}); & V_{\text{rel}}=0\end{cases} \tag{4}$$

during the slipping and sticking portions of the response. $\sigma_{yy}$ is the interface normal pressure, *a function of space and time*, and $\mu_{\max}$ is a limiting value for the sticking friction coefficient. Note that during slipping, both $\tau_f$ and $\sigma_{yy}$ are implicit functions of time and space. Calculation of the effective friction coefficient at $V_{\text{rel}}=0$ is nontrivial and is explained in more detail in a subsequent section.

Element matrix formulation is achieved using the standard finite element coordinate transformations with linear shape functions. A supplemental set of elements and interpolation functions has been defined for calculations along the contact interface. The so-called ''surface'' (interface) elements, simply one-dimensional elements using linear shape functions $N_i^s$, are used to derive the boundary loading vector due to contact friction interactions. Therefore, along the contact interface, frictional shear stress is an interpolated quantity, and the local interface effective friction coefficient can be written:

$$\mu_{\text{eff}}^e(\xi)=\sum_{i=1}^{i=2}\mu_{\text{eff},i}^e N_i^s(\xi) \tag{5}$$

where $\xi$ is a natural coordinate along the element.

After performing the necessary finite element matrix manipulations, the resulting discrete equations of motion for a system of $n/2$ nodes and $m$ interface nodes are

$$[M]\ddot{\mathbf{x}}+[C]\dot{\mathbf{x}}+[K]\mathbf{x}=\mathbf{F}=\mathbf{F_f}+\mathbf{F_p}+\mathbf{F_q} \tag{6}$$

where the vector $\mathbf{x}$ contains both $x$ and $y$-displacements of the nodes, and therefore has size $(n\times 1)$. The right-hand side of the equation contains the friction force vector $\mathbf{F_f}(n\times 1)$ and also the normal and tangential traction vectors $\mathbf{F_p}(n\times 1)$ and $\mathbf{F_q}(n\times 1)$, associated with the known pressure distribution $P_{\text{app}}(x,t)$ and known shear traction $Q_{\text{app}}(x,t)$, respectively. In the nonconstant reference velocity case, a discretized version of Eq. (3) can be added to the system above, yielding an augmented system of $n+1$ second-order differential equations. It should be noted again that the velocity-dependent terms in the discretized equation do not arise from a rate-dependent stress-strain relationship, but rather from the assumption that there is a dissipative body force proportional to velocity (see Eq. (1)).

**2.2 Boundary Tractions: Interface Shear Stress Calculations.** The right-hand side of Eq. (6) involves contributions from applied loads ($\mathbf{F_p}$ and $\mathbf{F_q}$, derived following usual finite element procedures) and contact interface behavior ($\mathbf{F_f}$). The continuous equation for frictional shear traction at the contact surface was given by Eq. (2). In order to determine nodal friction force values for each surface element, the surface element interpolation functions must be used in the following integral expression:

$$\mathbf{F_f^e}=\int_{L_e}[N^s]^T\tau_f\,dx \tag{7}$$

where $[N^s]$ is the interpolation function matrix, $L_e$ is the interface element length, and $\tau_f=\mu_{\text{eff}}\sigma_{yy}$ is the frictional shear traction on the contact surface.

The stress component $\sigma_{yy}$ is in general a function of space and time, and it can be determined along the contact interface using the stress-strain strain-displacement, and interpolation function expressions above:

$$\{\sigma(x(t))\}=[E][L][N]x(t) \tag{8}$$

where $[N]$ is the matrix of shape functions for the domain elements. Using the friction coefficient expression, Eq. (5), the integral (7) becomes

$$\mathbf{F_f^e}=\int_{L_e}\sigma_{yy}(x(t))[N^s]^T[N^s]\{\mu_{\text{eff}}^e\}dx \tag{9}$$

$$=\left\{\int_{L_e}\sigma_{yy}(x(t))[N^s]^T[N^s]dx\right\}\{\mu_{\text{eff}}^e\} \tag{10}$$

where $\{\mu_{\text{eff}}^e\}$ is the vector of nodal effective friction coefficients for surface element $e$. The integral can be evaluated using Gaussian quadrature for each element, resulting in the following expression for the global friction force vector, assembled from the elemental expressions:

$$\mathbf{F_f}=[\hat{F}_f]\{\mu_{\text{eff}}\}. \tag{11}$$

$[\hat{F}_f](n\times m)$ is in general not a diagonal matrix (a banded matrix of bandwidth 3 for line elements on the contact surface), and $\{\mu_{\text{eff}}\}(m\times 1)$ is a vector containing the effective contact friction coefficients for the $m$ interface nodes. Because of the banded structure of the $[\hat{F}_f]$ matrix, this equation implies that the friction force acting at a particular contact node depends not only upon the effective friction coefficient at that node, but *also upon the effective friction coefficients of neighboring nodes.* The important implication for systems in which partial slipping occurs is that the friction force at a slipping node depends upon not only the known effective friction coefficient at the slipping node, but also the unknown sticking effective friction coefficient at neighboring sticking nodes.

As a result, if the number of sticking nodes is $p$, the contact friction force equation can be rewritten as follows:

$$\mathbf{F_f}=[\hat{F}_f^{\text{stick}}]\{\mu_{\text{eff}}^{\text{stick}}\}+[\hat{F}_f^{\text{slip}}]\{\mu_{\text{eff}}^{\text{slip}}\} \tag{12}$$

where $\{\mu_{\text{eff}}^{\text{stick}}\}(p\times 1)$ contains the effective sticking friction coefficients, and $\{\mu_{\text{eff}}^{\text{slip}}\}((m-p)\times 1)$ contains the effective slipping friction coefficients. Therefore, the other matrices have the following sizes:

$$[\hat{F}_f^{\text{stick}}]; \quad n\times p \tag{13}$$

$$[\hat{F}_f^{\text{slip}}]; \quad n\times(m-p). \tag{14}$$

In this way, the sticking and slipping effective friction coefficients have been isolated for later use; note that the reduced-order matrices $[\hat{F}_f^{\text{stick}}]$ and $[\hat{F}_f^{\text{slip}}]$ are constructed directly using the appropriate columns of the full matrix $[\hat{F}_f]$. Each of these matrices is updated at every time-step to reflect the changing interface normal forces (see Eqs. (10) through (11)).

**2.3 Sticking Motion and Constraint Equations.** During sticking of a contact node, the motion of that node is constrained by the following relation:

$$V_{\text{ref}}-\dot{x}_i=0 \tag{15}$$

where $\dot{x}_i$ is the $x$-direction velocity of the sticking contact node. In general, the constraint equations for all the sticking nodes can be written in matrix form

$$[\Phi]\left\{\frac{\dot{\mathbf{x}}}{V_{\text{ref}}}\right\}=\mathbf{0} \tag{16}$$

where the coefficient matrix $[\Phi]$ $(p\times(n+1))$ has as many rows are there are sticking nodes. The rows of $[\Phi]$ have $(-1)$'s in the column corresponding to $\dot{x}_i$, and (1)'s in the last column (corresponding to the reference velocity). For use in the numerical scheme detailed next, the constraint equations on the acceleration level must be developed by differentiating Eq. (16):

$$[\Phi]\left\{\frac{\ddot{\mathbf{x}}}{\dot{V}_{\text{ref}}}\right\}=\mathbf{0}. \tag{17}$$

Finally, note that in the constant reference velocity case, the constraint equations given by Eq. (15) become

$$\ddot{x}_i=0 \tag{18}$$

and the reference velocity is removed from the system of constraint equations (Eq. (17)).

**2.4 The Mixed Differential-Algebraic Equation (MDAE): Global System of Equations.** Stick-slip calculations require special care because of the following observations: (i) for a slipping contact node, the effective friction coefficient is known, while its motion is unknown, and (ii) for a sticking contact node, the motion is known, while the effective friction coefficient is unknown. In the first case, differential equations arise for the dynamic motion of the system, while in the second case, algebraic equations describing the constraint forces on the sticking contact nodes can be derived. A convenient means of handling these equations is the mixed differential-algebraic equation, described by Haug [31] and stated as follows:

$$\left[\begin{array}{c|c}[M] & -[\hat{F}_f^{\text{stick}}]\\ \hline [\Phi] & [0]\end{array}\right]\left\{\begin{array}{c}\ddot{\mathbf{x}}\\ \hline \{\mu_{\text{eff}}^{\text{stick}}\}\end{array}\right\}=\left\{\begin{array}{c}-[C]\dot{\mathbf{x}}-[K]\mathbf{x}+\mathbf{F_p}+\mathbf{F_q}+[\hat{F}_f^{\text{slip}}]\{\mu_{\text{eff}}^{\text{slip}}\}\\ \hline \mathbf{0}\end{array}\right\}. \tag{19}$$

This formulation allows the effective contact sticking friction coefficients to be solved for directly. The partitioned system of Eq. (19) includes both the differential and algebraic equations, and couples them together so that simultaneous solution for the vector of unknowns insures a set of ''compatible'' accelerations and sticking effective friction coefficients. Using the two reduced-order matrices $[\hat{F}_f^{\text{stick}}]$ and $[\hat{F}_f^{\text{slip}}]$ allows decoupling of the known and unknown effective friction coefficients. The result of this procedure is that both the differential and algebraic equations are guaranteed to be satisfied.

## 3 Solution Algorithm and Stick-Slip Detection

Stick-slip transition times for the contact nodes during system motion must be determined accurately in order to avoid numerical difficulties; recall that the equations of motion are explicitly changed from differential to algebraic at a slip-to-stick transition (and the opposite way for a stick-to-slip transition). Relatively large time-steps can be used in regions away from the transition points where the system dynamics are well behaved, while very small time-steps must be used in the vicinity of the transition points. Guidelines for choosing an appropriate time-step are illustrated in the context of a specific example presented in the next section. An efficient algorithm has been developed to accurately determine the transition points for both slip-to-stick and stick-to-slip.

**3.1 Stick-Slip Transitions.** The transition point from slip to stick can be determined using a simple zeroth-order optimization procedure as follows. The relative velocity of a node $V_{\text{rel},i}$ can be used to define a cost function for optimization when considered as shown in Fig. 2. If at a particular time-step, say time-step $i$, the sign of the relative velocity changes, the sticking boundary has been crossed. At this point, it is useful to iterate and find (to some pre-determined numerical accuracy) the exact time of crossing, $t_{\text{cross}}$. The unimodal cost function which can be used for this purpose is

$$\mathcal{F}_{sl}=|V_{\text{ref}}-\dot{x}_i|. \tag{20}$$

The minimum of this function is $\mathcal{F}_{sl}=0$, which occurs at the time $t=t_{\text{cross}}$. This procedure in practice works well, because large time-steps can be used until a crossing point has been detected; then, the time-step is iteratively decreased until the exact crossing time is determined, and the integration again proceeds with the larger time-step.

A similar procedure is used to detect changes from sticking to slipping. In this case, the cost function for optimization is

$$\mathcal{F}_{st}=[\mu_{\max}-\mu_{\text{eff},i}]^2 \tag{21}$$

which once again has a minimum value of $\mathcal{F}_{st}=0$ at the transition point. Both of these optimization procedures can be implemented using any of the zeroth-order methods described in, for example, Haftka and Gürdal [32], and a golden section line search algorithm has been used here.

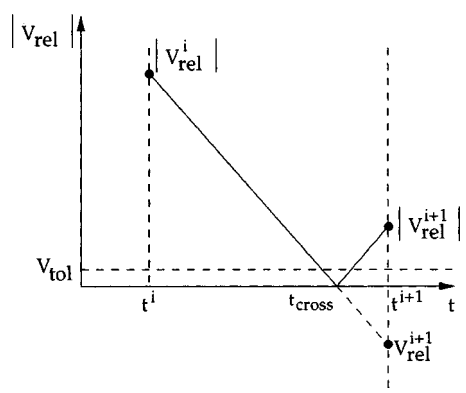



**Fig. 2 Stick-slip detection algorithm using zeroth-order optimization**

**3.2 Equation Normalization.** For the $n$ discrete equations of motion, there exist $n$ natural frequencies; for realistic material stiffness and density parameters and reasonable geometries, the natural frequencies of the structure will be widely spaced (perhaps over several orders of magnitude). The requirements for temporal resolution in the numerical simulation then make the computational cost of the procedure quite high. It is important to develop a procedure which helps to reduce the cost, and the following equation normalization has been employed.

The key scaling parameters (with their units) for the equations are the following:

$$\hat{c}=\sqrt{\frac{E}{\rho L^2}};\quad \text{material wave frequency}\ \left[\frac{1}{s}\right]$$

$$\tau=\hat{c}t;\quad \text{dimensionless time}\ [\ ]$$

$$\hat{\mathbf{x}}=\frac{\mathbf{x}}{L};\quad \text{dimensionless displacement}\ [\ ]$$

where $A$ is the domain area and $L$ is the domain length (measured along the interface). Using these quantities as scaling factors, time differentiation becomes

$$\frac{d}{d\tau}(\ldots)=\frac{1}{\hat{c}}\frac{d}{dt}(\ldots)=(\ldots)' \tag{22}$$

$$\frac{d^2}{d\tau^2}(\ldots)=\frac{1}{\hat{c}^2}\frac{d^2}{dt^2}(\ldots)=(\ldots)''. \tag{23}$$

In addition, common factors are extracted from each of the dynamic matrices as follows:

$$[\hat{M}]=\frac{[M]}{\rho AL};\quad \text{scaled mass matrix}$$

$$[\hat{C}]=\frac{[C]}{c_m AL};\quad \text{scaled damping matrix}$$

$$[\hat{K}]=\frac{[K]}{EA};\quad \text{scaled stiffness matrix.}$$

Each of these scaled matrices is dimensionless. Using the time and length scaling as described above, and factorization of the mass, stiffness, and damping matrices, the dimensionless slipping equations of motion can be rewritten as

$$[\hat{M}]\hat{\mathbf{x}}''+\frac{c_m}{\rho\hat{c}}[\hat{C}]\hat{\mathbf{x}}'+[\hat{K}]\hat{\mathbf{x}}=\frac{1}{\rho AL\hat{c}^2}[\mathbf{F_f}+\mathbf{F_p}+\mathbf{F_q}] \tag{24}$$

where $c_m$ is a characteristic volumetric material damping parameter with units of $[(N-s)/m](1/m^3)$. Similar scaling procedures, using the defined parameters, can be applied to the other equations of motion for the system, including sticking equation, the mixed differential-algebraic equation, constraint equations, etc.

The crucial implication of this scaling procedure on solution of the equations is that the natural frequencies of the discrete system, which strongly influence the temporal resolution requirement for simulations, are *independent* of material parameters $E$ and $\rho$. Instead, they are a function of only the domain geometry and node positions. As a result, the bandwidth over which the natural frequencies exist is much smaller than for the case of fully dimensional equations of motion (as in Eq. (6) above). Therefore, time-stepping through the dimensionless equations of motion becomes much less expensive than solving the fully dimensional equations of motion.

## 4 Application to a Dynamic Partial Slip Problem

In this section, the approach for dynamic contact calculations, including the piecewise equations of motion related to stick-slip oscillations, are applied to an example problem that is relevant to partial slip contact in blade dampers in turbines ([33]). A schematic of the geometry and loading parameters for the example problem is shown in Fig. 3. In the following example, the temporally constant applied normal pressure $P_{\text{app}}(x)$ is related to the time-varying applied tangential loading $Q_{\text{app}}(x,t)$ as follows:

$$Q_{\text{app}}(x,t)=\mu_{\text{far}}P_{\text{app}}(x)\sin(\omega t) \tag{25}$$

where $\mu_{\text{far}}$ is a far-field loading parameter, and $\omega$ is the tangential loading frequency. Their discrete (finite element method) counterparts are shown as the far-field loading vectors $\mathbf{F_p}$ and $\mathbf{F_q}$ on Fig. 3. In all cases, the interface normal pressure $\sigma_{yy}$ is closely monitored to insure that the no-loss-of-contact assumption is accurate. In the cases presented here, $\sigma_{yy}<0$ for all $x$, and contact occurs along the entire interface throughout the full simulation.

For the loading parameters used here, which result in partial slipping of the interface, small slip velocities and displacements are observed. As a result, the history-dependence of the generalized friction law of Eq. (2) is negligible; in this scenario, a simple rate-dependent friction law is appropriate:

$$\mu_{\text{eff}}(V_{\text{rel}})=\mu_0+\mu_1\exp\left(-\frac{1}{V_o}|V_{\text{rel}}|\right) \tag{26}$$

which is valid only for nonzero relative velocity. This approach is equivalent to assuming the slip displacements observed in partial slip contact are much smaller than any characteristic slip length as defined by the rate and state-dependent friction equations. In addition, because evolutionary models are typically described for *step* changes in velocity, and the current work considers *forced response* to harmonic excitation, the definition of friction evolution in reciprocating contact is unclear. The results presented here are restricted to the case for which gross slip is not considered. An estimate for the maximum allowable tangential traction magnitude (such that gross slip does not occur) is found by treating the deformation as quasi-static:

$$\mu_{\text{far}}<(\mu_0+\mu_1). \tag{27}$$

Simulation parameters for the results shown in this section are listed in Table 1. In this case, a stationary reference surface $V_{\text{ref}}=0$ is assumed. The dimensionless first shear mode natural frequency for the partially constrained (i.e., partial sticking boundary condition) elastic body has been calculated; $\omega_1\approx 0.65$. The actual value of first natural frequency depends upon how large the sticking zone is. The eigenvalue problem has been solved using the discrete mass, damping, and stiffness matrices determined from the finite element procedure above, with only one node (on the center line of symmetry) sticking, and a frictionless boundary condition elsewhere. This is a limiting behavior which predicts a

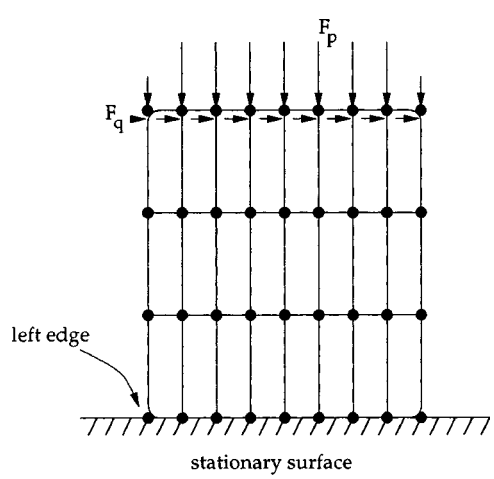



**Fig. 3 Schematic of domain geometry and loading conditions for an example problem**

**Table 1 Simulation parameters used for dynamic stick-slip calculations**

| Parameter | Units | Value |
|---|---|---|
| Applied pressure | Pa | $P_{\text{app}}(x)=P_0$ |
| | | $=1\times10^7$ (constant) |
| Friction parameters | - | $\mu_0=0.1$ |
| | - | $\mu_1=0.02$ |
| | m/s | $V_o=2.0$ |
| Material modulus | Pa | $E=200\times10^9$ |
| Material density | kg/m$^3$ | $\rho=7850$ |
| Damping parameter | N−s/m$^4$ | $c_m=250$ |
| Reference velocity | m/s | $V_{\text{ref}}=0$ |
| Length | m | $L=1$ |
| Height | m | $h=1$ |
| First shear mode natural frequency | - | $\omega_1\approx0.65$ |
| Time step | - | $\Delta\tau=0.01$ |
| Number $x$-nodes | - | 9 |
| Number $y$-nodes | - | 4 |

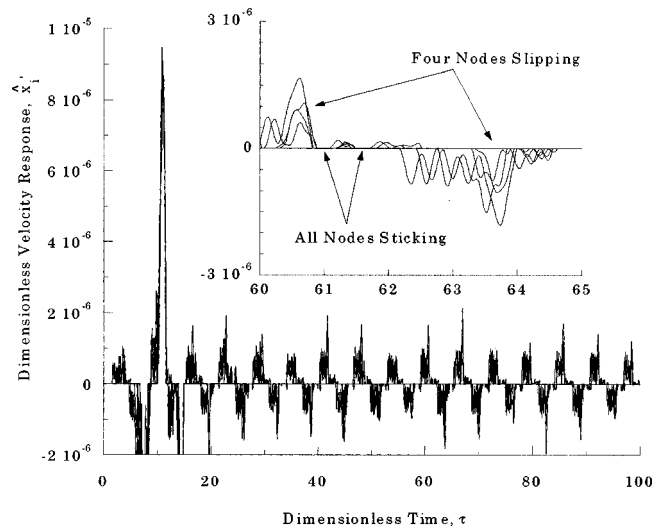



**Fig. 4 Dimensionless interface velocity response for $\omega$=1 and $\mu_{\text{far}}$=0.05; inset: close-up showing detail of partial slipping response**

*lower bound* for the first shear mode natural frequency; larger sticking zones imply a stiffer system, and therefore a higher natural frequency.

The time-step $\Delta\tau$ for the numerical simulations requires some trial and error to determine. In particular, the transition from stick to slip, and slip to stick, of a node is a very high frequency phenomenon, requiring a very small time-step to capture accurately. However, as discussed in Section 3.1, an efficient procedure for accurately determining the transition time has been implemented. Away from the stick-slip transition points, the required time-step is governed by the natural frequencies of the system. The fundamental time-step (for calculations away from the stick-slip transition points) should be small enough to capture the highest frequency of interest. For the simulations presented here, the time-step is $\Delta\tau$=0.01 (dimensionless seconds). By contrast, the forced response for the range of frequency examined here has a minimum fundamental period of $T\approx2\pi$ (dimensionless seconds), so at least 600 time points are captured per forcing cycle.

Figure 4 shows the normalized interface node velocity response as a function of time. Note that there are periods during which as many as four interface nodes experience slipping (see the inset to Fig. 4). More importantly, there also exist operating periods of total stick (all nodes at zero velocity), as well as operation with only the contact edges slipping. The large response peak near the beginning of the time history is a start-up transient which is quickly damped. A different view of this same data is shown in the waterfall plot of Fig. 5, which clearly shows the contact center experiencing sustained stick, as expected. In the waterfall plot, every 50th data point has been plotted to improve the clarity of the figure. In addition, the transient period toward the beginning of the time history has been neglected in this plot. Note that the edge velocities vary with the same frequency ($\omega$=1) as the tangential

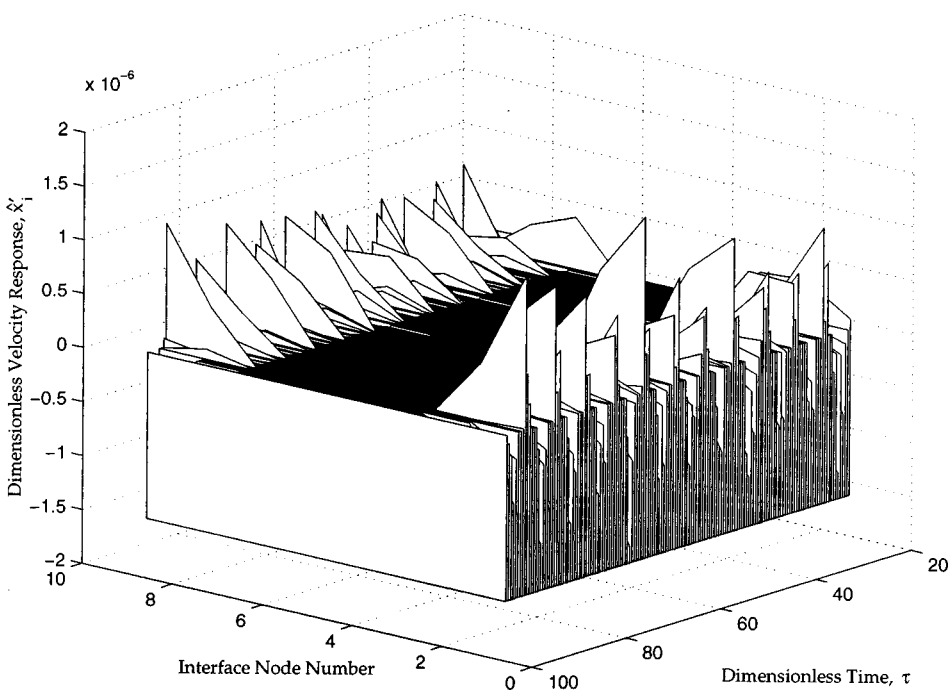



**Fig. 5 Waterfall plot showing regions of consistent stick (center of contact) and regions of nonzero dynamic response in phase with tangential forcing ($\omega$=1 and $\mu_{\text{far}}$=0.05)**

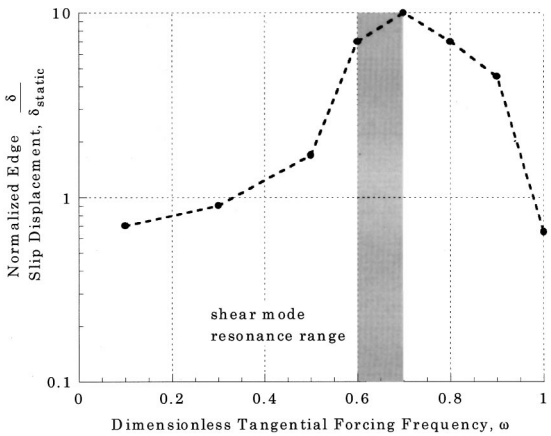



**Fig. 6 Edge node maximum normalized slip displacement versus tangential forcing frequency**

forcing. The numerical algorithm produces very stable predictions for slip amplitude, interface (partial) stick-slip, and effective friction coefficient (not shown).

Figure 6 shows the variation of maximum normalized displacement response (for the node at the left contact edge) as a function of tangential forcing frequency $\omega$. The displacements are normalized by the slip displacement resulting from a quasi-static analysis of the type typically used in the contact mechanics community (see Johnson [34] for details). (Obviously, the number of interface nodes is limited here to reduce computational expense; more finely meshed quasi-static analyses using the commercial code ABAQUS indicate that the quasi-static slip displacements for the relatively coarse mesh presented here are within ten percent of those resulting from more detailed analyses.) The choice of normalization emphasizes the dramatic increase in slip displacement as the excitation frequency approaches the first shear mode natural frequency ($\omega_1 \approx 0.65$, marked on the figure). It is interesting to note that the peak in response amplitude is observed slightly above the lower bound estimate of the natural frequency. The interplay between dynamic excitation and slip behavior is further emphasized in Fig. 7, in which the time-averaged percent of the interface experiencing sticking is plotted as a function of tangential forcing frequency. Slip behavior which is confined to the contact edges at low frequencies extends towards the center of contact near the first shear natural frequency.

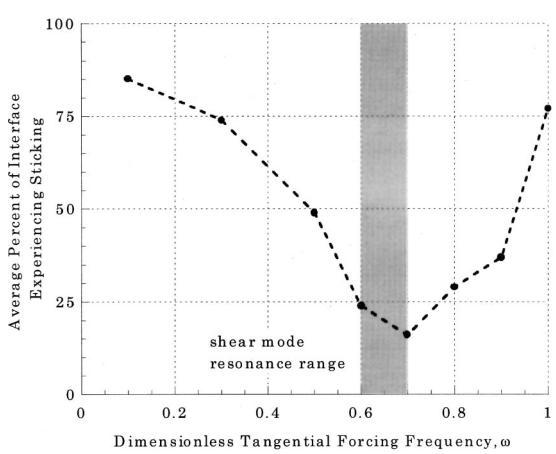



**Fig. 7 Average percent of the interface experiencing sticking versus tangential forcing frequency**

## 5 Discussion and Implications

The analytical modeling and numerical solution procedures developed here, including the MDAE formulation for the partial slipping dynamics problem, show excellent performance in capturing complicated interface behavior. For finite-domain systems experiencing dynamic loading, or for which system dynamic effects are likely to play a significant role in interface response, the techniques developed here provide powerful tools for fully transient interface calculations. The partial slip regime can be captured, and the slip amplitudes can be calculated directly. In addition, interface friction force variations under partial slipping conditions can be determined.

The numerical algorithm for dynamic stick-slip interface response calculations presented here has also produced interesting results with regard to frequency-dependence of interface response. Indeed, it appears that under certain frequency tuning relationships, i.e., $(\omega/\omega_1) \approx 1$, the interface response can be large. This resonance phenomenon clearly has strong implications on interface response, and nonresonant excitation results in a smaller interface response with a larger portion of the interface experiencing sticking (Figs. 6 and 7). In addition, the numerical algorithm captures out-of-plane motion as well as spatial and temporal variations in interface normal forces. However, for the range of loading presented here, normal motion of the body and interface normal force variations are small.

It is important to point out that the spatially and temporally complicated (and, in fact, frequency-dependent) interface behavior is not attributable to a self-excited sliding instability. Indeed, it is the dynamic response of the structure itself which influences the interface response. Self-excited interface vibrations of the kind examined previously in the literature are precluded here by sufficiently large damping terms on the left-hand side of Eq. (24). In addition, for the system response under external harmonic excitation considered here, complicated localized behavior near the stick-slip boundary can be expected. The completely reversed tangential load produces reciprocating velocities, which pass through zero relative velocity twice per loading cycle; at the interface, this implies that transitions between sticking (zero relative velocity) and slipping occur frequently. In addition, the stick-slip boundary itself moves during the loading cycle in response to changes in both tangential and normal stress. The result is extremely rich dynamic behavior along the contact interface.

There is an unresolved issue as to whether or not Rice's argument of a critical cell size ([20]) is applicable here; the problem considered here involves forced response of a damped system that does not exhibit slip instabilities. The question of whether or not the system considered is ''inherently discrete'' is largely irrelevant since slip velocities and displacements are limited by the sticking region. Convergence tests have indicated that the frequency dependence of the results is independent of the mesh used, and slip zone sizes are insensitive to mesh refinement; even the slip displacement magnitudes are relatively insensitive to mesh refinement. This is supported by convergence studies for the dynamic model and by comparisons with more detailed quasi-static analyses.

Furthermore, Fig. 7 indicates that the tuning relationship between the tangential forcing frequency and the shear mode natural frequency can have an important impact on the interface dynamics. For example, because there is more energy available for interface in-plane deformation under resonant conditions, the system operates closer to a gross slip condition, with larger slip amplitudes. In this case, proper frequency tuning appears to be favorable (from a fatigue life viewpoint), as the system is shifted from the deleterious partial slipping condition, toward the lower-wear gross slip condition.

The sensitivity of the results to changes in finite element or friction parameters can be estimated by considering a frequency response function (FRF) relating excitation shear traction and interface response (for a suitably linearized system). General con-

clusions for the linearized FRF approach agree with the trends in the full finite element method solution. Damping tends to decrease the FRF amplitude at resonance, and therefore increasing material damping $c_m$ will decrease the interface response. The frequency at which slip response is maximum is related to component geometry, as well as density and modulus. For the small slip amplitudes associated with partial slip response, the results are very *insensitive* to the velocity-dependence of friction (i.e., to $\mu_1$); this is evidenced by the fact that the maximum relative velocity is never larger than several percent of the scaling velocity factor in the friction law. Therefore, it is $\mu_0$ which governs the extent of partial slipping. Finally, the extent of the coupling between structural dynamic and interface response is governed by a combination of two parameters: (i) the magnitude of the structural FRF, and (ii) the maximum allowable sticking friction force. Although the range of parameters considered here is well within the range of typical values seen in practice, a more systematic convergence study is undoubtedly called for. Such numerical simulations are beyond the scope of this paper, and will be presented in future work ([33]).

## 6 Conclusions

This paper presents a new algorithm for contact interface problems with partial slip under dynamic loading. The approach considers the full elastodynamic response of continuous elastic systems with state-dependent friction laws, and as such considers both the differential equations associated with slipping response, and the algebraic constraint equations related to sticking response. This model provides several distinct advantages over other models from the literature, most notably the accurate calculation of the following spatially and temporally varying parameters: interface contact normal and shear forces, interface slip displacement, stick and slip zone sizes, and structural dynamic influence on interface response.

The MDAE formulation provides a numerically stable means of determining the partial slipping response and sticking friction coefficients. The sample results presented here illustrate that the structural dynamics of the system being examined can play an important role in determining interface response. At certain tangential forcing frequencies, a condition of almost pure slip results at the interface, even for $\mu_{\text{far}}=0.05$ in this case; the larger slip amplitudes and smaller stick zone may result in operation in a different (and less severe) wear regime. The implications of the response shown in Fig. 6 are significant in many applications. The results shown in Fig. 6 illustrate that the slip distances and slip zone sizes estimated from quasi-static calculations which ignore the structural dynamic response may significantly underestimate the behavior under dynamic loading.

## References


[1] Zagrodzki, P., 1990, ''Analysis of Thermomechanical Phenomena in Multidisc Clutches and Brakes,'' Wear, **140**, pp. 291–308.

[2] Armstrong-Helouvry, B., 1991, *Control of Machines With Friction*, Kluwer, Boston.

[3] Waterhouse, R. B., 1992, ''The Problems of Fretting Fatigue Testing,'' *Standardization of Fretting Fatigue Test Methods and Equipment-ASTM STP 1159*, R. B. Waterhouse and M. H. Attia, eds., ASTM, Philadelphia, pp. 13–19.

[4] Söderberg, S., Bryggman, U., and McCullough, T., 1986, ''Frequency Effects in Fretting Wear,'' Wear, **110**, pp. 19–34.

[5] Bryggman, U., and Söderberg, S., 1986, ''Contact Conditions in Fretting,'' Wear, **110**, pp. 1–17.

[6] Schouterden, K., Blanpain, B., Celis, J. P., and Vingsbo, O., 1995, ''Fretting of Titanium Nitride and Diamond-Like Carbon Coatings at High Frequencies and Low Amplitude,'' Wear, **181-183**, pp. 86–93.

[7] Tolstoi, D. M., 1967, ''Significance of the Normal Degree of Freedom and Natural Normal Vibrations in Contact Friction,'' Wear, **10**, pp. 199–213.

[8] Budanov, B. V., Kudinov, V. A., and Tolstoi, D. M., 1980, ''Interaction of Friction and Vibration,'' Soviet Journal of Friction and Wear, **1**, pp. 79–89.

[9] Nayak, P. R., 1972, ''Contact Vibrations,'' J. Sound Vib., **22**, No. 3, pp. 297–322.

[10] Gray, G. G., and Johnson, K. L., 1972, ''The Dynamic Response of Elastic Bodies in Rolling Contact to Random Roughness of Their Surfaces,'' J. Sound Vib., **22**, No. 3, pp. 323–342.

[11] Hess, D. P., and Soom, A., 1992, ''Normal and Angular Motions at Rough Planar Contacts During Sliding With Friction,'' ASME J. Tribol., **114**, pp. 567–578.

[12] Oden, J. T., and Martins, J. A. C., 1985, ''Models and Computational Methods for Dynamic Friction Phenomena,'' Comput. Methods Appl. Mech. Eng., **52**, pp. 527–634.

[13] Martins, J. A. C., Oden, J. T., and Simoes, F. M. F., 1990, ''A Study of Static and Kinetic Friction,'' Int. J. Eng. Sci., **28**, pp. 29–92.

[14] Tworzydlo, W. W., Becker, E. B., and Oden, J. T., 1994, ''Numerical Modeling of Friction-Induced Vibrations and Dynamic Instabilities,'' Appl. Mech. Rev., **47**, No. 7, pp. 255–274.

[15] Berger, E. J., Mahajani, M., and Karnik, J., 1998, ''Local Stick-Slip Effects in Dynamic Contact Systems,'' ASME Journal Of Tribology, in press.

[16] Mahajani, M., 1998, ''A Study in Stick-Slip Vibrations,'' Master's thesis, University of Cincinnati, Cincinnati, OH.

[17] Dieterich, J. H., 1978, ''Time-Dependent Friction and the Mechanisms of Stick-Slip,'' Pure Appl. Geophys., **116**, pp. 790–806.

[18] Dieterich, J. H., 1979, ''Modeling of Rock Friction 1: Experimental Results and Constitutive Equations,'' J. Geophys. Res., **84**, pp. 2161–2168.

[19] Dieterich, J. H., 1979, ''Modeling of Rock Friction 2: Simulation of Preseismic Slip,'' J. Geophys. Res., **84**, pp. 2169–2175.

[20] Rice, J. R., 1993, ''Spatio-Temporal Complexity of Slip on a Fault,'' J. Geophys. Res., **98**, No. B6, pp. 9885–9907.

[21] Gu, J.-C., Rice, J. R., Ruina, A. L., and Tse, S. T., 1984, ''Slip Motion and Stability of a Single Degree of Freedom Elastic System With Rate and State Dependent Friction,'' J. Mech. Phys. Solids, **32**, No. 3, 167–196.

[22] Rice, J. R., and Ruina, A. L., 1983, ''Stability of Steady Frictional Slipping,'' ASME J. Appl. Mech., **50**, pp. 343–349.

[23] 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.

[24] Adams, G. G., 1998, ''Steady Sliding of Two Elastic Half-Spaces With Friction Reduction due to Interface Stick-Slip,'' ASME J. Appl. Mech., **65**, pp. 470–475.

[25] Adams, G. G., 1999, ''Dynamic Motion of Two Elastic Half-Spaces in Relative Sliding Without Slipping,'' ASME J. Tribol., **121**, No. 3, pp. 455–461.

[26] Menq, C.-H., Griffin, J. H., and Bielak, J., 1986, ''The Influence of a Variable Normal Load on the Forced Vibration of a Frictionally Damped Structure,'' ASME J. Eng. Gas Turbines Power, **108**, pp. 300–305.

[27] Menq, C.-H., and Griffin, J. H., 1985, ''A Comparison of Transient and Steady State Finite Element Analyses of the Forced Response of a Frictionally Damped Beam,'' ASME J. Vibr. Acoust., **107**, pp. 19–25.

[28] Menq, C.-H., Bielak, J., and Griffin, J. H., 1986, ''The Influence of Microslip on Vibratory Response, Part I: A New Microslip Model,'' J. Sound Vib., **107**, No. 2, pp. 279–293.

[29] Menq, C.-H., Griffin, J. H., and Bielak, J., 1986, ''The Influence of Microslip on Vibratory Response, Part II: A Comparison With Experimental Results,'' J. Sound Vib., **107**, No. 2, pp. 295–307.

[30] Ben-Zion, Y., and Rice, J. R., 1997, ''Dynamic Simulations of Slip on a Smooth Fault in an Elastic Solid,'' J. Geophys. Res., **102**, Vol. B8, pp. 17771–17784.

[31] Haug, E. J., 1989, *Computer-Aided Kinematics and Dynamics of Mechanical Systems*, Allyn and Bacon, Boston.

[32] Haftka, R. T., and Gürdal, Z., 1992, *Elements of Structural Optimization*, 3rd Ed., Kluwer, Boston.

[33] Berger, E. J., Begley, M. R., and Breitenstein, J., 2000, ''Interface Energy Dissipation in Microslip Contacts,'' J. Sound Vib., in review.

[34] Johnson, K. L., 1985, *Contact Mechanics*, Cambridge University Press, Cambridge, UK.

**F. D. Fischer**[1]
Professor of Mechanics,
Mem. ASME

**E. R. Oberaigner**
Associate Professor,

Institute for Mechanics
Montanuniversitat Leoben,
A-8700 Leoben, Austria

# Deformation, Stress State, and Thermodynamic Force for a Transforming Spherical Inclusion in an Elastic-Plastic Material

*The progress of a transformed phase into an elastic-plastic parent phase is simulated by a growing sphere. The transformation is accompanied by a dilatational volume change. The strain and stress state in the full space is presented. In addition, the local and global energy terms are calculated. Finally the thermodynamic forces on the interface are derived. Also strain hardening is considered.* [S0021-8936(00)00304-4]

## 1 Introduction

In material physics a sphere with radius $R(t)$, $t$ being time, is very often used to represent a phase transformation nucleus or a transforming microregion embedded in a parent phase. Roitburd and Temkin [1,2] and Kaganova and Roitburd [3] investigated the martensitic transformation in an elastic-plastic material by such a model. The transformation volume strain $\boldsymbol{\varepsilon}^*=\varepsilon_o\mathbf{I}$, $\mathbf{I}$ unity tensor, gives rise to an internal strain and stress state due to its accommodation by the surrounding material. A slightly more general study for an elastic material with different material properties was recently presented by Nazyrov and Freidin [4].

Spheres as the product phase are also dealt with by several researchers on diffusional transformation see, e.g., Vandermeer [5].

Analytical solutions for the strain and stress state of the sphere, imbedded in an infinite elastic-plastic material, with fixed radius $R$ subjected to a sudden dilatational eigenstrain are usually applied. This problem was surprisingly often treated in an independent way in the literature. However, a comprehensive collection of all results is still lacking. The paper by Lee et al. [6] as well as the review article by Fischer et al. [7] may be cited. Several mistakes exist in the diverse papers; these will be avoided in the following context. Since the pressure at the interface between the internal sphere with radius $R$ and the outer elastic-plastic space (hollow sphere) can be considered as the unknown quantity, the solution is directly related to the stress and strain field of a hollow sphere under internal pressure, for details see, e.g., Lubliner [8] chapter 4.3.2.

The following aspects must be taken into account when dealing with a growing spherical inclusion instead of a fixed spherical inclusion:

- An incremental concept must be followed with respect to the development of the plastic strain and the plastic work.
- Under the assumption of pure radial loading, the incremental and deformation theories of plasticity are identical. If no unloading occurs, the solution offered in the literature for a fixed radius $R$ can be used, as can be seen by an argument by Budiansky [9]. However, with respect to the total plastic work it must be considered that also the material inside the sphere, $0\leqslant r\leqslant R(t)$, has plastified.
- The growth process can also be modeled by a sequence of elastic hollow spheres, see Franciosi et al. [10], introducing a certain scaling factor, however, with much more computational effort than below.

The main goal of this paper is to provide the researcher in the field of phase transformations in an elastic-plastic material with a full and comprehensive set of relations for the strain and stress state as well as energy terms relevant for the understanding of the growth process.

## 2 Analysis

**2.1 Strain, Stress, and Energy Terms.** We solve now the problem of a sphere subjected to a dilational eigenstrain and imbedded in an elastic-plastic space. Since we think of phase transformations, $\varepsilon_o$ is mostly on the order of some few percents, and thus we apply the infinitesmal deformation theory.

We introduce a stress vector $\boldsymbol{\sigma}^T=(\sigma_r,\sigma_\varphi,\sigma_\theta)$, where the superscript $^T$ stands for ''transposed,'' a strain vector $\boldsymbol{\varepsilon}^T=(\varepsilon_r,\varepsilon_\varphi,\varepsilon_\theta)$, the three vectors $\mathbf{e}^T=(1,1,1)$, $\tilde{\mathbf{e}}^T=(0,1,1)$, $\tilde{\tilde{\mathbf{e}}}^T=(2,-1,-1)$, the unity matrix $\mathbf{I}$ ($I_{ij}=\delta_{ij}$, Kronecker delta) and the matrix $\mathbf{J}$, $J_{ij}=1$. The yield stress is $\sigma_f$. Linear hardening is described by $E_p$, being the plastic tangent modulus, and a hardening factor $f=1+2(1-\nu)E_p/E$. If not explicitly mentioned an elastic-ideally plastic material with $f=1$ is assumed.

The strain vector is decomposed into an elastic part $\boldsymbol{\varepsilon}_e$, a plastic part $\boldsymbol{\varepsilon}_p$, and within $0\leqslant r\leqslant R$ a transformation part $\boldsymbol{\varepsilon}^*=\varepsilon_o\mathbf{e}$. Due to Hooke's law $\boldsymbol{\varepsilon}_e$ can be written as

$$E\varepsilon_e=[(1+\nu)\mathbf{I}-\nu\mathbf{J}]\cdot\boldsymbol{\sigma}. \tag{1}$$

Since a hydrostatic load stress $\sigma_o$ does not introduce any stress deviator component, the following results can be superposed on the corresponding elastic solution for $\sigma_o$.

An explicit relation for $R_p$, the radius of the outer boundary of the plastic zone, can be found as

$$\frac{R_p}{R}=\left(\frac{E\varepsilon_o}{(1-\nu)\sigma_f}\right)^{1/3}=\kappa^{1/3}, \tag{2}$$

see, e.g., Fischer et al. [7] or Böhm et al. [11].

Of course, since $R_p\geqslant R$, the limit condition for plastification ensures that the plasticity constant is greater than or equal to one.

Three different zones can be distinguished with respect to the stress state $\boldsymbol{\sigma}$ and the corresponding strain components ($\boldsymbol{\varepsilon}_e$ is directly calculated by applying Hooke's law):

[1]To whom correspondence should be addressed.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, June 23, 1999; final revision, Feb. 14, 2000. Associate Technical Editor: M.-J. Pindera. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

 **Copyright © 2000 by ASME** 

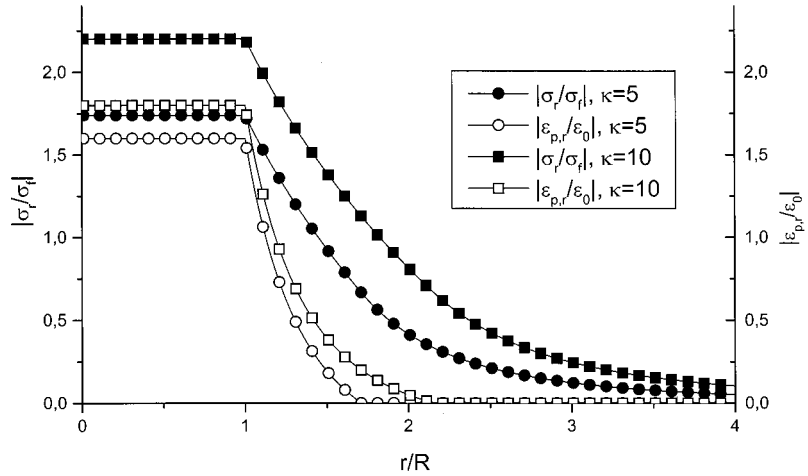



**Fig. 1 Profile of the relative radial stress $|\sigma_r/\sigma_f|$ and the relative plastic strain $|\varepsilon_{p,r}/\varepsilon_o|$ in relation to $r/R$ for two values of $\kappa$**

- The inner elastic zone $0 \leqslant r \leqslant R(t)$:

$$\boldsymbol{\sigma} = -\frac{2\sigma_f}{3}(1+\ln\kappa)\mathbf{e}, \tag{3a}$$

$$E\varepsilon_e = -(1-2\nu)\frac{2\sigma_f}{3}(1+\ln\kappa)\mathbf{e}, \tag{3b}$$

$$\boldsymbol{\varepsilon}_p = \mathbf{0}, \quad \boldsymbol{\varepsilon}^* = \varepsilon_o\mathbf{e}.$$

- The elastic-plastic zone $R(t) \leqslant r \leqslant R_p(t)$:

$$\boldsymbol{\sigma} = -\frac{2\sigma_f}{3}\mathbf{e} - \frac{2\sigma_f}{3}\left[\ln\kappa - \ln\left(\frac{r}{R}\right)^3\right]\mathbf{e} + \sigma_f\tilde{\tilde{\mathbf{e}}}, \tag{3c}$$

$$E\boldsymbol{\varepsilon}_e = -(1-2\nu)\frac{2\sigma_f}{3}\left[\ln\kappa - \ln\left(\frac{r}{R}\right)^3\right]\mathbf{e} - \frac{(1+\nu)}{3}\sigma_f\tilde{\tilde{\mathbf{e}}}, \tag{3d}$$

$$\boldsymbol{\varepsilon}_p = -\left[\left(\frac{R}{r}\right)^3 - \frac{1}{\kappa}\right]\varepsilon_o\tilde{\tilde{\mathbf{e}}}. \tag{3e}$$

- The outer elastic zone $R_p \leqslant r \leqslant \infty$:

$$\boldsymbol{\sigma} = -\frac{\sigma_f}{3}\kappa\left(\frac{R}{r}\right)^3\tilde{\tilde{\mathbf{e}}}, \tag{3f}$$

$$E\boldsymbol{\varepsilon}_e = -(1+\nu)\frac{\sigma_f}{3}\kappa\left(\frac{R}{r}\right)^3\tilde{\tilde{e}}, \tag{3g}$$

$$\boldsymbol{\varepsilon}_p = \mathbf{0}.$$

Figure 1 demonstrates profiles for the relative radial stress $|\sigma_r/\sigma_f|$ and the relative radial plastic strain $|\varepsilon_{p,r}/\varepsilon_o|$ in relation to the dimensionless radius $r/R$ for two values of $\kappa$.

The plastic strain attains a maximum value of $\boldsymbol{\varepsilon}_{p,\max} = -[1-1/\kappa]\varepsilon_o\tilde{\tilde{\mathbf{e}}}$ at $r=R$.

It is important to note that the inner zone $0 \leqslant r \leqslant R(t)$ which behaves elastically, has achieved a plastic strain in the amount of $\boldsymbol{\varepsilon}_p = -[1-1/\kappa]\varepsilon_o\tilde{\tilde{\mathbf{e}}}$ during the growth process from $R=0$ to its actual value. This means that no jump of the plastic strain at the interface $r=R(t)$ can be observed, see (3*e*).

Due to spherical symmetry the total strain components are $\varepsilon_r = \partial u/\partial r$ and $\varepsilon_\varphi = \varepsilon_\theta = u/r$, which means that due to the compatibility $\varepsilon_\varphi$ and $\varepsilon_\theta$ must be continuous. The condition can easily be verified with the above relations. Therefore, the radial displacement $u|_R$ can be calculated from $\varepsilon_\varphi$ at $r=R$, (3*b*), as

$$u|_R = -(1-2\nu)\frac{2\sigma_f}{3E}(\ln\kappa+1)R + \varepsilon_o R$$

$$= \frac{\sigma_f}{E}R\left[(1-\nu)\kappa - \frac{2(1-2\nu)}{3}(\ln\kappa+1)\right]. \tag{3h}$$

The strain energy density $U' = 1/2\boldsymbol{\sigma}^T\cdot\boldsymbol{\varepsilon}_e$ can be calculated for the three regions as follows:

- The inner elastic zone $0 \leqslant r \leqslant R(t)$:

$$U' = \frac{2(1-2\nu)}{3}\cdot\frac{\sigma_f^2}{E}\cdot(1+\ln\kappa)^2. \tag{4a}$$

- The elastic-plastic zone $R(t) \leqslant r \leqslant R_p(t)$:

$$U' = \frac{2(1-2\nu)}{3}\cdot\frac{\sigma_f^2}{E}\cdot\left[\ln\kappa - \ln\left(\frac{r}{R}\right)^3\right]^2 + \frac{(1+\nu)}{3}\frac{\sigma_f^2}{E}. \tag{4b}$$

- The outer elastic zone $R_p \leqslant r \leqslant \infty$:

$$U' = \frac{(1+\nu)}{3}\cdot\frac{\sigma_f^2}{E}\cdot\kappa^2\left(\frac{R}{r}\right)^6. \tag{4c}$$

Note the jump in the specific strain energy at $r=R(t)$!

We can now calculate the total strain energy produced during the growth of the inclusion from the origin to a certain radius $R$. Employing (4*a*)–(4*c*) yields after some integration

$$U = \int_0^\infty U' 4r^2\pi dr = \frac{\sigma_f^2}{E}\cdot\frac{4R^3\pi}{3}\cdot(1-\nu)(2\kappa-1). \tag{5}$$

The increment $dW_p'$ of the specific plastic work in the plastic zone follows as

$$dW_p' = \boldsymbol{\sigma}^T\cdot d\boldsymbol{\varepsilon}_p, \tag{6a}$$

and with (3*c*) and (3*e*)

$$dW_p' = \left\{-\frac{2\sigma_f}{3}\left[\ln\kappa - \ln\left(\frac{r}{R}\right)^3\right]\mathbf{e}^T - \frac{\sigma_f}{3}\tilde{\tilde{\mathbf{e}}}^T\right\}\cdot\left(-\frac{3R^2}{r^3}\varepsilon_o\tilde{\tilde{\mathbf{e}}}\right)dR$$

$$= \frac{6R^2}{r^3}\varepsilon_o\sigma_f dR = 6(1-\nu)\cdot\frac{\sigma_f^2}{E}\cdot\kappa\cdot\frac{R^2}{r^3}dR. \tag{6b}$$

The total increment of plastic work $dW_p$ then follows as

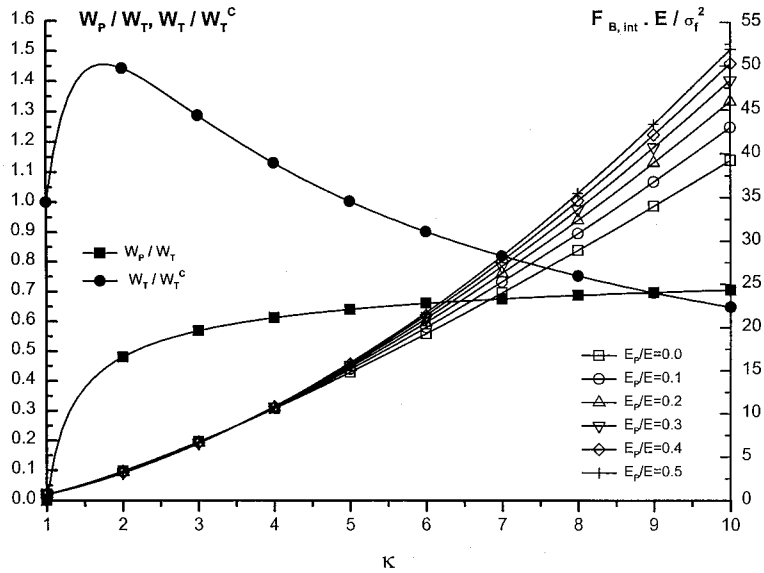



**Fig. 2 Plastic work $W_P$ in relation to total work $W_T$, total work $W_T$ in relation to total work $W_T^c$ in a purely elastic comparison material, dimensionless transformation barrier $F_{B,\mathrm{int}}\cdot(E/\sigma_f^2)$ as function of the plasticity constant $\kappa$.**

$$dW_p=\int_R^{R_p} dW_p' 4r^2\pi dr=8(1-\nu)\cdot\frac{\sigma_f^2}{E}\cdot\kappa\ln\kappa\cdot R^2\pi dR, \tag{6c}$$

and total plastic work $W_p$ done during the growth of the inclusion from the origin to a certain radius $R$ becomes

$$W_p=\frac{\sigma_f^2}{E}\cdot\frac{4R^3\pi}{3}\cdot 2(1-\nu)\kappa\ln\kappa. \tag{6d}$$

Böhm et al. [11] pointed out that Lee et al. [6] reported a different result, namely $\widetilde{W}_p=\sigma_f^2/E\cdot 4R^3\pi/3\cdot 2(1-\nu)\kappa[\ln\kappa-(1-1/\kappa)]$. They did not consider a growing inclusion and had, therefore, no plastic work in the region $0\leqslant r\leqslant R$.

We can now compare the total mechanical energy $W_T$ produced during generation of an inclusion with the radius $R$,

$$W_T=U+W_p=\frac{\sigma_f^2}{E}\cdot\frac{4R^3\pi}{3}\cdot(1-\nu)(2\kappa\ln\kappa+2\kappa-1), \tag{7}$$

with the amount of total mechanical energy $W_T^c$ which would occur in a purely elastic material. We can use the results from Fischer et al. [7] leading to

- the inner elastic zone $0\leqslant r\leqslant R(t)$:

$$\boldsymbol{\sigma}=-\frac{2}{3(1-\nu)}\varepsilon_o E\mathbf{e},\quad E\boldsymbol{\varepsilon}_e=-\frac{2(1-2\nu)}{3(1-\nu)}\varepsilon_o E\mathbf{e},\quad \boldsymbol{\varepsilon}^*=\varepsilon_o\mathbf{e}, \tag{8a}$$

$$U_T'=\frac{2}{3}\frac{(1-2\nu)}{(1-\nu)^2}\varepsilon_o^2 E; \tag{8b}$$

- the outer elastic zone $R(t)\leqslant r\leqslant\infty$:

$$\boldsymbol{\sigma}=-\frac{1}{3(1-\nu)}\varepsilon_o E\left(\frac{R}{r}\right)^3\tilde{\tilde{\mathbf{e}}},\quad E\boldsymbol{\varepsilon}_e=-\frac{(1+\nu)}{3(1-\nu)}\varepsilon_o E\left(\frac{R}{r}\right)^3\tilde{\tilde{\mathbf{e}}}, \tag{8c}$$

$$U_T'=\frac{(1+\nu)}{3(1-\nu)^2}\varepsilon_o^2 E\left(\frac{R}{r}\right)^6, \tag{8d}$$

$$W_T^c=\int_0^\infty U_T' 4r^2\pi dr=\frac{\varepsilon_o^2 E}{(1-\nu)}\cdot\frac{4R^3\pi}{3}. \tag{8e}$$

The ratios $W_T/W_T^c$

$$W_T/W_T^c=\frac{2\kappa\ln\kappa+2\kappa-1}{\kappa^2} \tag{9a}$$

and $W_p/W_T^c$

$$W_p/W_T=\frac{2\kappa\ln\kappa}{2\kappa\ln\kappa+2\kappa-1} \tag{9b}$$

are depicted in Fig. 2.

$W_T/W_T^c$ possesses a maximum of 1.456 at $\kappa=1.763$ and reaches the value 1 at $\kappa=5.03$.

For values of $\kappa>5.03$ the ratio $W_T/W_T^c$ tends to be smaller than 1, going to 0 as $\kappa$ goes to infinity. The conclusion is that depending on $\kappa$ an elastic-plastic body may consume more or less deformation energy compared with a purely elastic body.

Earmme et al. [12] offered also a solution for the stress and strain field around a spherical inclusion with fixed radius in a hardening material. Carrol [13] investigated independently some further types of hardening laws. In particular, for linear strain hardening it is interesting to note that the size of the plastic zone remains unaffected by the hardening. Taking into account the hardening factor $f=1+2(1-\nu)E_p/E$ and denoting the solution for no hardening, $f=1$, (see (3a–g)) as $\boldsymbol{\sigma}_n$, $\boldsymbol{\varepsilon}_{e,n}$, etc., the following relations can be derived at the interface:

- The inner elastic zone at $r=R(t)$

$$\boldsymbol{\sigma}=\frac{1}{f}\boldsymbol{\sigma}_n-\frac{2\sigma_f}{3f}\cdot 2E_p/E\cdot(1-\nu)\kappa\mathbf{e}, \tag{10a}$$

$$E\boldsymbol{\varepsilon}_e=\frac{E}{f}\boldsymbol{\varepsilon}_{e,n}-\frac{2\sigma_f}{3f}\cdot 2E_p/E\cdot(1-2\nu)(1-\nu)\kappa\mathbf{e}, \tag{10b}$$

$$U'=\frac{1}{f^2}\left\{U_n'+8\frac{\sigma_f^2}{3E}\cdot E_p/E\cdot(1-2\nu)\times(1-\nu)\cdot[1+\ln\kappa+E_p/E(1-\nu)\kappa]\right\}. \tag{10c}$$

- The elastic-plastic zone at $r=R(t)$

$$\boldsymbol{\sigma}=\frac{1}{f}\boldsymbol{\sigma}_n-\frac{2\sigma_f}{3f}\cdot 2E_p/E\cdot(1-\nu)\kappa\tilde{\tilde{\mathbf{e}}}, \tag{10d}$$

$$E\boldsymbol{\varepsilon}_e=\frac{E}{f}\boldsymbol{\varepsilon}_{e,n}-\frac{2\sigma_f}{3f}\cdot E_p/E\cdot(1-\nu)(1+\nu)\kappa\tilde{\tilde{\mathbf{e}}}, \tag{10e}$$

$$U'=\frac{1}{f^2}\left\{U'_n+4\frac{\sigma_f^2}{3E}\cdot E_p/E\cdot(1+\nu)(1-\nu)\cdot[\kappa+E_p/E\cdot(1-\nu)\kappa^2]\right\}. \tag{10f}$$

**2.2 Thermodynamics of Interface Motion.** If one wants to calculate the thermodynamic force on the interface $r=R(t)$, its mechanical contribution $F_M$ can be calculated as the normal component of the Eshelby energy momentum tensor, see, e.g., Fischer and Reisner [14] Eq. (26). With $[\![.]\!]$ denoting the jump of a quantity and $\langle . \rangle$ the average of a quantity at the interface $r=R(t)$, one gets

$$F_M=[\![U']\!]-\langle\boldsymbol{\sigma}^T\rangle\cdot[\![\boldsymbol{\varepsilon}]\!]. \tag{11a}$$

The jump in strain energy $[\![U']\!]$ can be reformulated with the identity $[\![\mathbf{a}^T\cdot\mathbf{b}]\!]=\langle\mathbf{a}^T\rangle\cdot[\![\mathbf{b}]\!]+[\![\mathbf{a}^T]\!]\cdot\langle\mathbf{b}\rangle$ as

$$[\![U']\!]=\frac{1}{2}\langle\boldsymbol{\sigma}^T\rangle\cdot[\![\boldsymbol{\varepsilon}_e]\!]+\frac{1}{2}[\![\boldsymbol{\sigma}^T]\!]\cdot\langle\boldsymbol{\varepsilon}_e\rangle=\langle\boldsymbol{\sigma}^T\rangle\cdot[\![\boldsymbol{\varepsilon}_e]\!] \tag{11b}$$

due to $\langle\boldsymbol{\sigma}^T\rangle[\![\boldsymbol{\varepsilon}_e]\!]=\langle\boldsymbol{\varepsilon}_e^T\rangle[\![\boldsymbol{\sigma}]\!]$.

$[\![\boldsymbol{\varepsilon}]\!]$ is $[\![\boldsymbol{\varepsilon}_e]\!]-\varepsilon_o\mathbf{e}$, since no jump in the plastic strain $\boldsymbol{\varepsilon}_p$ occurs at the interface. Inserting $[\![\boldsymbol{\varepsilon}]\!]$ and (11*b*) into (11*a*) yields

$$F_M=\varepsilon_o\langle\boldsymbol{\sigma}^T\rangle\cdot\mathbf{e}. \tag{11c}$$

If one introduces an additional load stress $\sigma_o$, $F_M$ finally yields

$$F_M=3\sigma_o\varepsilon_o-\varepsilon_o\frac{\sigma_f}{f}[(1+2\ln\kappa)+2E_p/E\cdot(1-\nu)\kappa]. \tag{11d}$$

The contribution $F_M=3\sigma_o\varepsilon_o$ may be positive or negative depending on the sign of $\sigma_o$. The second part of $F_M$ can be considered as a barrier (due to its sign),

$$F_{B,\mathrm{int}}=\frac{\sigma_f^2}{E}\cdot\frac{(1-\nu)\kappa}{f}[(1+2\ln\kappa)+2E_p/E\cdot(1-\nu)\kappa], \tag{11e}$$

due to the generation of the strain and stress by the accommodation of the transformation volume strain $\varepsilon_o\mathbf{e}$. The quantity $F_{B,\mathrm{int}}$ is depicted in Fig. 2. Only a rather weak influence of $E_p/E$ can be observed.

The total mechanical driving force $F_D$ follows by adding the jump of the chemical free energy $[\![\rho\varphi_{ch}]\!]$ to $F_M$. As a further barrier the energy necessary to rebuild the lattice, $F_L$, as well as to form a new surface, $\gamma/R$, $\gamma$ specific surface energy, must be considered for the energy balance at the interface.

Finally, a driving force $F_D$

$$F_D=[\![\rho\varphi_{ch}]\!]+3\sigma_o\varepsilon_o \tag{11f}$$

and a barrier $F_B$ with (11*e*)

$$F_B=F_{B,\mathrm{int}}+F_L+\gamma/R \tag{11g}$$

can be derived.

The interface velocity $\omega$ can then be formulated as a function

$$\omega=\phi(F_D/F_B)\quad\text{for}\quad F_D/F_B\geqslant 1,\quad\text{else}\quad 0 \tag{11h}$$

see, e.g., Rosakis and Knowles [15].

**2.3 Remark.** Recently Durban and Fleck [16] presented a solution for an expanding spherical cavity in a Drucker-Prager material with the yield condition

$$\sigma_\theta-\sigma_r+\frac{\mu}{3}\boldsymbol{\sigma}^T\cdot\mathbf{e}=\sigma_f, \tag{12}$$

where $\mu$ is a pressure sensitivity material parameter.

Surprisingly their solution for $R_p$, see (2), and $\boldsymbol{\sigma}$ cannot be reduced to the solution (3*a*)–(3*g*) for $\mu=0$. Their $\boldsymbol{\sigma}$ in the plastic zone leads to a constant stress vector for $\mu=0$ which contradicts the variation of $\boldsymbol{\sigma}$ with $r$ given in relation (3*c*).

## 3 Closing Comments

A full set of analytical expressions is given for the deformation, strain and stress state due to the dilatational growth of a spherical inclusion in an elastic-plastic material. The solutions can also be applied for a spherical cavity with the pressure $\tilde{p}=2\sigma_f/3\cdot(1+\ln\kappa)$ at the cavity surface. In addition the dissipated plastic work as well as the thermodynamic force on the transformation front are presented. The set of equations allows further metal-physical considerations with respect to the nucleation and growth of precipitations in an elastic-plastic material.

## References


[1] Roitburd, A. L., and Temkin, D. E., 1986, ''Plastic Deformation and Thermodynamic Hysteresis in Phase Transformation in Solids,'' Sov. Phys. Solid State, **28**, pp. 432–436.

[2] Roitburd, A. L., and Temkin, D. E., 1986, ''Hysteresis of a Phase Transformation in an Elastoplastic Medium,'' Sov. Phys. Dokl., **31**, pp. 414–416.

[3] Kaganova, I. M., and Roitburd, A. L., 1989, ''Effect of Plastic Deformation on the Equilibrium Shape of a New-Phase Inclusion and Thermodynamic Hysteresis,'' Sov. Phys. Solid State, **31**, pp. 545–550.

[4] Nazyrov, I. R., and Freidin, A. B., 1998, ''Phase Transformations in Deformation of Solids in a Model Problem of an Elastic Ball,'' Mech. Solids, **33**, pp. 39–56 Izevestiya AN.MTT, **5**, pp. 52–71.

[5] Vandermeer, R. A., 1990, ''Modeling Diffusional Growth During Austenite Decomposition to Ferrite in Polycrystalline Fe-C Alloys,'' Acta Metall. Mater., **38**, pp. 2461–2470.

[6] Lee, J. K., Earmme, Y. Y., Aronson, H. J., and Russel, K. C., 1980, ''Plastic Relaxation of the Transformation Strain Energy of a Misfitting Spherical Precipitate: Ideal Plastic Behavior,'' Metall. Trans. A, **11A**, pp. 1837–1847.

[7] Fischer, F. D., Berveiller, M., Tanaka, K., and Oberaigner, E. R., 1994, ''Continuum Mechanical Aspects of Phase Transformations in Solids,'' Arch. Appl. Mech., **64**, pp. 54–85.

[8] Lubliner, J., 1990, *Plasticity Theory*, Macmillan, New York.

[9] Budiansky, B., 1959, ''A Reassessment of Deformation Theories in Plasticity,'' ASME J. Appl. Mech., **81**, pp. 259–264.

[10] Franciosi, P., Lormand, G., and Fougeres, R., 1998, ''On Elastic Modeling of Inclusion Induced Microplasticity in Metallic Matrices From the Dilatating Sphere Problem,'' Int. J. Plast., **14**, pp. 1013–1032.

[11] Böhm, H. J., Fischer, F. D., and Reisner, G., 1997, ''Evaluation of Elastic Strain Energy of Spheroidal Inclusions With Uniform Volumetric and Shear Eigenstrains,'' Scr. Mater., **36**, pp. 1053–1059.

[12] Earmme, Y. Y., Johnson, W. C., and Lee, J. K., 1981, ''Plastic Relaxation of the Transformation Strain Energy of a Misfitting Spheroidal Precipitate: Linear and Power-Law Strain Hardening,'' Metall. Trans. A, **12A**, pp. 1521–1530.

[13] Carrol, M. M., 1985, ''Radial Expansion of Hollow Spheres of Elastic-Plastic Hardening Material,'' Int. J. Solids Struct., **21**, pp. 645–670.

[14] Fischer, F. D., and Reisner, G., 1998, ''A Criterion for the Martensitic Transformation of a Microregion in an Elastic-Plastic Material,'' Acta Mater., **46**, pp. 2095–2102.

[15] Rosakis, Ph., and Knowles, J. K., 1997, ''Unstable Kinetic Relations and the Dynamics of Solid-Solid Phase Transitions,'' J. Mech. Phys. Solids, **45**, pp. 2055–2081.

[16] Durban, D., and Fleck, N. A., 1997, ''Spherical Cavity Expansion in a Drucker-Prager Solid,'' ASME J. Appl. Mech., **64**, pp. 743–750.

# On Symmetrizable Systems of Second Kind

**S. Adhikari**
Department of Engineering,
University of Cambridge,
Trumpington Street,
Cambridge CB2 1PZ, England
e-mail: sa225@eng.cam.ac.uk

*We discuss under what conditions multiple-parameter asymmetric linear dynamical systems can be transformed into equivalent symmetric systems by nonsingular linear transformations. So far, in structural dynamics literature this problem has been addressed in the context of the original work by Taussky. Taussky's approach of symmetrization was based on similarity transformation. In this paper an approach is proposed to transform asymmetric systems into symmetric systems by equivalence transformation. We call Taussky's approach of symmetrization by similarity transformation ''first kind'' and proposed approach by equivalence transformation ''second kind.'' Since equivalence transformations are most general nonsingular linear transformations, conditions of symmetrizability obtained here are more ''liberal'' than the first kind and numerical calculations also become more straightforward. Several examples are provided to illustrate the new approach.* [S0021-8936(00)00504-3]

## 1 Introduction

Theory of linear dynamics of multiple parameter symmetric systems is well developed now. However, dynamical behavior of some systems encountered in practice cannot be expressed in terms of symmetric coefficient matrices or self-adjoint linear operators. Some examples are gyroscopic and circulatory systems ([1]), aircraft flutter ([2]), ship motion in sea water ([3]), contact problems ([4]), and many actively controlled systems ([5]). Few authors have considered such general asymmetric dynamical systems. Fawzy and Bishop [6] presented several relationships satisfied by the eigenvectors and ''eigenrows'' of a damped asymmetric system and also presented a method to normalize them. Caughey and Ma [5] have derived conditions under which such systems can be diagonalized by a *similarity transformation.* In a subsequent paper Ma and Caughey [7] utilized *equivalence transformation* to analyze asymmetric nonconservative systems and gave the condition under which they can be diagonalized. Recently Adhikari [8] proposed a method to obtain (complex) eigensolutions of general asymmetric nonconservative systems without converting the equations of motion into first-order form.

The above-mentioned works have certainly enhanced the power of modal analysis in dealing with asymmetric systems. However, asymmetric systems are still not as well understood as symmetric systems. For example, rightly pointed out by Kliem [9], stability investigations become substantially easier if the system matrices are symmetric. Without doubt, it would be preferable if asymmetric systems can be transformed into equivalent symmetric systems so that one can take advantage of the well-developed theories for symmetric systems to analyze them. This is the primary reason to study ''symmetrizability'' of asymmetric systems. In linear dynamics literature symmetrizability has been addressed by *similarity transformation* based on Taussky's [10] definition. We call Taussky's approach of symmetrization by similarity transformation as ''symmetrization of first kind.'' In this paper symmetrizability of a matrix is redefined in the context of *equivalence transformation.* Such symmetrization will be called ''symmetrization of second kind.''

Equivalence transformations are the most general class of nonsingular linear transformations. This motivates us to utilize equivalence transformation rather than similarity transformation for symmetrization of asymmetric systems. It will be shown that much generality can be achieved by using symmetrization of the second kind compared to the first kind. Notations and definitions of some terminologies frequently used in this paper are given in Section 2. Taussky's version of symmetrization and a brief review of the literature on its applications in structural dynamics is presented in Section 3. In Section 4 we formally define the symmetrization of the second kind. For the sake of generality, some basic results on such symmetrization are presented on complex matrices. Numerical methods are outlined to carry out such a symmetrization procedure. In view of the undamped and damped dynamical systems, simultaneous symmetrization of two and three matrices are considered in Section 5 and Section 6, respectively. Finally, Section 7 summarizes the main results of this paper and some suggestions towards further research required to successfully apply this new approach is provided. Throughout the paper suitable numerical examples are provided to illustrate the derived results.

## 2 Notations, Basic Concepts, and Definitions

By $\mathbb{R}^{N\times N}$ we mean the space of $N\times N$ real matrices and $\mathbb{C}^{N\times N}$ stands for the space of $N\times N$ complex matrices. An $N\times N$ matrix taken from either real or complex number field will be denoted by $\mathbb{F}^{N\times N}$. A matrix $\mathbf{A}$ is called positive definite if all of its eigenvalues $\lambda_i>0$ and will be denoted by $\mathbf{A}>0$. A unit matrix $\mathbf{I}\in\mathbb{R}^{N\times N}$ is a diagonal matrix with all diagonal entries equal to one. Let $\mathbf{A}\in\mathbb{C}^{N\times N}$, then we denote $\mathbf{A}^T$, $\overline{\mathbf{A}}$, $\mathbf{A}^{-1}$, $\mathbf{A}^{-T}$, and $\mathbf{A}^*$ be the transpose, complex conjugate, inverse, inverse transposed, and transposed conjugate. A matrix $\mathbf{A}$ is called *symmetric* if $\mathbf{A}=\mathbf{A}^T$, *Hermitian* if $\mathbf{A}=\mathbf{A}^*$, and *unitary* if $\mathbf{A}\mathbf{A}^*=\mathbf{I}$. If $\mathbf{A}$ is real then a Hermitian matrix is equivalent to a symmetric matrix and a unitary matrix is equivalent to a real orthogonal matrix as $\mathbf{A}\mathbf{A}^T=\mathbf{I}$. For a matrix $\mathbf{A}$, by saying $\mathbf{A}^{-1}$ exists we mean none of its eigenvalues is equal to 0 and that $\mathbf{A}$ is nonsingular. Two matrices, $\mathbf{A}$ and $\mathbf{B}$, related by $\mathbf{B}=\mathbf{V}^T\mathbf{A}\mathbf{U}$ for some nonzero $\mathbf{U}$ and $\mathbf{V}$, are called an *equivalence transformation.* When $\mathbf{V}^T=\mathbf{U}^{-1}$, the equivalence transformation is called the *similarity transformation* and we call $\mathbf{A}$ and $\mathbf{B}$ *similar.* In the event $\mathbf{V}=\mathbf{U}$ the equivalence transformation is a *congruence transformation.* Classical modal transformation in symmetric systems is an example of congruence transformation. When $\mathbf{V}^T=\mathbf{U}^{-1}=\mathbf{U}^T$, we call such transformation the *orthogonal transformation.* Further, if $\mathbf{V}^T=\mathbf{V}^{-1}$ and $\mathbf{U}^T=\mathbf{U}^{-1}$, such transformation is called the *biorthogonal transformation.* A

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Sept. 21, 1999; final revision, May 2, 2000. Associate Technical Editor: A. A. Ferri. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

 Copyright © 2000 by ASME 

matrix $\mathbf{A}$ is said to be *diagonalizable* if it is similar to a diagonal matrix and that diagonal matrix contains all the eigenvalues of $\mathbf{A}$ with proper multiplicities. A matrix is diagonalizable if and only if it has $N$ linearly independent eigenvectors.

## 3 Symmetrizability of the First Kind

The concept of symmetrizability of an asymmetric matrix was originated from an excellent result by Taussky and Zassenhaus [11] which says ''for every $\mathbf{A}\in\mathbb{F}^{N\times N}$ there is a non-singular symmetric matrix transforming $\mathbf{A}$ into its transpose.'' Based on this general result it was shown that every real square matrix can be expressed as a product of two real symmetric matrices, that is

$$\mathbf{A}=\mathbf{S}_1\mathbf{S}_2;\quad \mathbf{S}_i=\mathbf{S}_i^T\in\mathbb{R}^{N\times N}, \tag{3.1}$$

always holds. Later Taussky [10] proved that if one of the factors in this representation is positive definite then $\mathbf{A}$ can be transformed into a symmetric matrix by a similarity transformation and vice versa. She formally defined *symmetrizability* of a matrix as the following:

DEFINITION 1. *A matrix* $\mathbf{A}$ *is symmetrizable if and only if any one of the following hold:*

1. $\mathbf{A}$ is the product of two symmetric matrices, one of which is positive definite;
2. $\mathbf{A}$ is similar to a symmetric matrix;
3. $\mathbf{A}^T=\mathbf{S}^{-1}\mathbf{A}\mathbf{S}$ with $\mathbf{S}=\mathbf{S}^T>0$; and
4. $\mathbf{A}$ has real characteristic roots and a full set of characteristic vectors.

Huseyin and Liepholz [1] were possibly first to recognize the importance of Taussky's result in the context of structural dynamics. However, symmetrization was not properly exploited until a decade later when Inman's [12] paper appeared. Inman's work inspired several authors to consider the dynamics of symmetrizable asymmetric systems. Subsequent works by Ahmadian and Inman [13,14], Ahmadian and Chou [15], Shahruz and Ma [16], and Cherng and Abdelhamid [17] made significant contributions and symmetrization of the first kind is much better understood now. Next, a more general approach for symmetrization is proposed.

## 4 Symmetrizability of the Second Kind

**4.1 Definition.** Now we will introduce a new definition of symmetrizability by utilizing equivalence transformation.

DEFINITION 2. *A matrix* $\mathbf{A}$ *is symmetrizable of the second kind if and only if there exist two nonzero matrices* $\mathbf{L}$ *and* $\mathbf{R}$ *such that* $\tilde{\mathbf{A}}=\mathbf{L}^T\mathbf{A}\mathbf{R}$ *is a symmetric matrix.*

This definition of symmetrizability is quite general and valid for both real and complex matrices. It also holds Taussky's definition as a special case when $\mathbf{L}^T=\mathbf{R}^{-1}=\mathbf{R}^{-T}$. In the above definition, even $\mathbf{A}$ is real $\mathbf{L}$ and $\mathbf{R}$ might be complex in order to make $\tilde{\mathbf{A}}$ symmetric. For any $\mathbf{A}$, when $\mathbf{L}$ and $\mathbf{R}$ are real matrices we call such $\mathbf{A}$ *real symmetrizable* and *complex symmetrizable* when $\mathbf{L}$ and $\mathbf{R}$ are complex. Taussky [18] and Pommer and Kliem [19] have discussed complex symmetrizable matrices in the context of a similarity transformation. Applications of such complex symmetrizability in linear dynamical systems were given by Kliem [9].

**4.2 Basic Results.** Some basic results on symmetrizability of the second kind and a numerical method to carry out such a symmetrization procedure will be developed. In order to achieve more generality, one of our main results is presented for nonsquare complex matrices in the following theorem:

THEOREM 4.1. *Every rectangular complex matrix* $\mathbf{A}\in\mathbb{C}^{r\times s}$ *can be transformed to a real symmetric (square) matrix by an equivalence transformation.*

*Proof.* We use a result by Eckrat and Young ([20], Theorem 1) which states that *for every* $\mathbf{A}\in\mathbb{C}^{r\times s}$ *there are two unitary matrices* $\mathbf{X}\in\mathbb{C}^{r\times r}$ *and* $\mathbf{Y}\in\mathbb{C}^{s\times s}$ *such that* $\mathbf{D}=\mathbf{X}^*\mathbf{A}\mathbf{Y}$ *is a diagonal matrix with real elements, none of which are negative.* If $\mathrm{rank}(\mathbf{A})=k$, then Williamson [21] has shown that $\mathbf{D}=\left[\begin{smallmatrix}\mathbf{\Lambda} & \mathbf{0}_1\\ \mathbf{0}_2 & \mathbf{0}_3\end{smallmatrix}\right]$ where $\mathbf{\Lambda}=\mathrm{diag}(\lambda_1,\cdots,\lambda_k)$ is a real diagonal matrix and $\mathbf{0}_1$, $\mathbf{0}_2$, and $\mathbf{0}_3$ are $(k,s-k)$, $(r-k,k)$, and $(r-k,s-k)$ null matrices. Without loss of generality we select $\mathbf{L}=\overline{\mathbf{X}}\mathbf{Q}_1$ and $\mathbf{R}=\mathbf{Y}\mathbf{Q}_2$ for some nonzero $\mathbf{Q}_1\in\mathbb{R}^{r\times n}$ and $\mathbf{Q}_2\in\mathbb{R}^{s\times n}$ such that $\mathbf{Q}_1=\left[\begin{smallmatrix}\mathbf{\Theta} & \mathbf{Q}_1^1\\ \mathbf{Q}_2^1 & \mathbf{Q}_3^1\end{smallmatrix}\right]$ and $\mathbf{Q}_2=\left[\begin{smallmatrix}\mathbf{\Theta} & \mathbf{Q}_1^2\\ \mathbf{Q}_2^2 & \mathbf{Q}_3^2\end{smallmatrix}\right]$ where $\mathbf{\Theta}\in\mathbb{R}^{k\times k}$ and $\mathbf{Q}_j^i$, $i=1,2$; $j=1,2,3$ are all real matrices of proper orders. Using these, one has $\tilde{\mathbf{A}}=\mathbf{L}^T\mathbf{A}\mathbf{R}=\mathbf{Q}_1^T\mathbf{X}^*\mathbf{A}\mathbf{Y}\mathbf{Q}_2=\mathbf{Q}_1^T\mathbf{D}\mathbf{Q}_2=\left[\begin{smallmatrix}\mathbf{\Theta}^T\mathbf{\Lambda}\mathbf{\Theta} & \tilde{\mathbf{0}}_1\\ \tilde{\mathbf{0}}_2 & \mathbf{0}_3\end{smallmatrix}\right]\in\mathbb{R}^{n\times n}$ where $\tilde{\mathbf{0}}_j$; $j=1,2,3$ are all null matrices with proper orders. So $\tilde{\mathbf{A}}$ is real symmetric matrix and this completes the proof. □

*Remark 4.1.* The two matrices $\mathbf{Q}_1$ and $\mathbf{Q}_2$ utilized above can be selected in such a way that the symmetric form becomes nonsingular and has the same rank as $\mathbf{A}$. Observe that because $\mathbf{D}$ is a diagonal matrix with non-negative real elements, $\mathbf{Q}_1$ and $\mathbf{Q}_2$ can also be selected in a way such that the symmetric form becomes positive definitive. The symmetric form of $\mathbf{A}$ is, however, not unique because $\mathbf{Q}_1$ and $\mathbf{Q}_2$ can be chosen in several ways.

This result clearly demonstrates the generality of the symmetrization of the second kind compared to the first kind which cannot be applied to rectangular matrices. Of course, Theorem 4.1 is applicable for real matrices as a special case. When $\mathbf{A}$ is a real matrix we have the following interesting special result:

THEOREM 4.2. *Every real matrix* $\mathbf{A}$ *is real symmetrizable of the second kind.*

*Proof.* We use a result by Pearl ([22], Theorem 1). This gives the condition under which the unitary matrices which reduces the matrix $\mathbf{A}$ to a real diagonal form by an equivalence transformation in Eckrat and Young's theorem are real orthogonal matrices. According to this result, *for* $\mathbf{A}\in\mathbb{C}^{r\times s}$ *there are two real orthogonal matrices* $\mathbf{O}_1\in\mathbb{R}^{r\times r}$ *and* $\mathbf{O}_2\in\mathbb{R}^{s\times s}$ *such that* $\mathbf{D}=\mathbf{O}_1^*\mathbf{A}\mathbf{O}_2$ *is a diagonal matrix with real non-negative elements if and only if* $\mathbf{A}\mathbf{A}^*$ *and* $\mathbf{A}^*\mathbf{A}$ *are both real matrices.* The theorem is proved since (trivially) this condition is always satisfied when $\mathbf{A}$ is real. □

We have described two important results on symmetrization of the second kind. From a practical point of view the obvious question is, for a given matrix $\mathbf{A}$, how to calculate $\mathbf{L}$ and $\mathbf{R}$ so that $\tilde{\mathbf{A}}=\mathbf{L}^T\mathbf{A}\mathbf{R}$ is a real symmetric matrix? In answer to this question, we observe that the proof given by Eckrat and Young [20] of their theorem described previously can be used for obtaining $\mathbf{L}$ and $\mathbf{R}$. For computational purposes the following steps may be followed to transform a general complex matrix $\mathbf{A}\in\mathbb{C}^{r\times s}$ to a nonsingular real square matrix using an equivalence transformation.

1. Calculate the matrix $\mathcal{M}=\mathbf{A}\mathbf{A}^*\in\mathbb{C}^{r\times r}$. Solve the eigenvalue problem $\mathcal{M}\mathbf{x}_i=\lambda_i^2\mathbf{x}_i$, $\forall i=1,\cdots r$. Since $\mathcal{M}$ is a non-negative Hermitian its eigenvector matrix can be normalized to a unitary matrix. Suppose $\mathrm{rank}(\mathbf{A})=k$, arrange the numbering of the eigenvalues so that $\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_k>0$; $\lambda_{k+1}=\cdots=\lambda_r=0$. Denote $\mathbf{X}=[\mathbf{x}_1,\mathbf{x}_2,\cdots,\mathbf{x}_k]\in\mathbb{C}^{r\times k}$ as the ordered collection of the eigenvectors.
2. Evaluate $\mathbf{y}_i=\mathbf{A}^*\mathbf{x}_i/\lambda_i$, $\forall i=1,\cdots k$. Set $\mathbf{Y}=[\mathbf{y}_1\mathbf{y}_2,\cdots,\mathbf{y}_k]\in\mathbb{C}^{s\times k}$.
3. Consider any nonzero $\mathbf{Q}\in\mathbb{R}^{k\times k}$ and obtain $\mathbf{L}=\overline{\mathbf{X}}\mathbf{Q}\in\mathbb{C}^{r\times k}$ and $\mathbf{R}=\mathbf{Y}\mathbf{Q}\in\mathbb{C}^{s\times k}$. Finally check that $\tilde{\mathbf{A}}=\mathbf{L}^T\mathbf{A}\mathbf{R}$ is a $(k\times k)$ real nonsingular symmetric matrix.

For the sake of generality the method is proposed for rectangular complex matrices. This procedure is obviously applicable to real square matrices we normally encounter in the equations of motion of linear systems. When $\mathbf{A}$ is real, $\mathcal{M}$ becomes a real symmetric matrix and consequently $\mathbf{X}$, $\mathbf{Y}$, $\mathbf{L}$, and $\mathbf{R}$ all become real matrices. For a further special case, when $\mathbf{A}$ is a symmetrizable matrix of first kind, this procedure provides an alternative and easy way to find symmetric forms as the calculation of Taussky's factorization can be avoided.

*Example 4.1.* The procedure outlined above can be illustrated by the following rectangular complex matrix:

$$\mathbf{A}=\begin{bmatrix} 1.0+3.0i & -2.0+1.0i & 1.5-2.0i \\ -2.0-3.0i & 6.0+2.0i & 7.0+2.0i \end{bmatrix}. \tag{4.1}$$

Step 1: We calculate $\mathcal{M}=\mathbf{AA}^*=\begin{bmatrix} 21.25 & -14.50-10.00i \\ -14.50+10.00i & 106.00 \end{bmatrix}$. Solving the eigenvalue problem $\mathcal{M}\mathbf{x}_i=\lambda_i^2\mathbf{x}_i$ and arranging accordingly one has $\lambda_i^2=\{109.5150, 17.7350\}$ and $\mathbf{X}=\begin{bmatrix} -0.1611-0.1111i & 0.9807 \\ 0.9807 & 0.1611-0.1111i \end{bmatrix}$. Note that rank$(\mathbf{A})=2$.

Step 2: Calculating one obtains

$$\mathbf{Y}=\begin{bmatrix} -0.2347+0.3167i & 0.2355-0.5311i \\ 0.5824-0.1508i & -0.2890-0.4677i \\ 0.6541-0.2341i & 0.5643+0.2045i \end{bmatrix}.$$

Step 3: Now arbitrarily select the matrix $\mathbf{Q}=\begin{bmatrix} 0.1021 & 0.2605 \\ 1.4109 & 0.3519 \end{bmatrix}$ so that $\mathbf{L}=\overline{\mathbf{X}}\mathbf{Q}=\begin{bmatrix} 1.3672+0.0113i & 0.3031+0.0289i \\ 0.3274+0.1568i & 0.3121+0.0391i \end{bmatrix}$ and

$$\mathbf{R}=\mathbf{YQ}=\begin{bmatrix} 0.3083-0.7170i & 0.0218-0.1044i \\ -0.3483-0.6752i & 0.0500-0.2039i \\ 0.8630+0.2647i & 0.3690+0.0110i \end{bmatrix}.$$

Using this $\mathbf{L}$ and $\mathbf{R}$ one has $\tilde{\mathbf{A}}=\mathbf{L}^T\mathbf{AR}=\begin{bmatrix} 8.4923 & 2.3692 \\ 2.3692 & 1.2315 \end{bmatrix}$, which is a real nonsingular symmetric matrix.

*Example 4.2.* Consider a real asymmetric matrix

$$\mathbf{A}=\begin{bmatrix} 1.0 & -2.0 & 1.5 \\ 12.0 & 6.0 & 7.0 \\ -2.0 & 4.0 & 9.0 \end{bmatrix}. \tag{4.2}$$

Observe that rank$(\mathbf{A})=3$. Selecting the matrix

$$\mathbf{Q}=\begin{bmatrix} 0.0645 & -0.1647 & 0.1436 \\ -0.8923 & 0.2226 & 0.6155 \\ 0.1611 & 0.2298 & 0.3708 \end{bmatrix} \tag{4.3}$$

and following the procedure described previously we obtain

$$\mathbf{L}=\begin{bmatrix} 0.1735 & 0.2198 & 0.3701 \\ 0.3901 & -0.2469 & -0.1170 \\ -0.8026 & 0.1421 & 0.6215 \end{bmatrix}$$

and (4.4)

$$\mathbf{R}=\begin{bmatrix} 0.7299 & -0.2298 & -0.2942 \\ -0.2638 & -0.2377 & -0.1620 \\ -0.4733 & 0.1419 & 0.6513 \end{bmatrix}$$

and both are real matrices. From these one has

$$\tilde{\mathbf{A}}=\mathbf{L}^T\mathbf{AR}=\begin{bmatrix} 7.0392 & -1.7960 & -4.4600 \\ -1.7960 & 1.0004 & 1.0316 \\ -4.4600 & 1.0316 & 3.9722 \end{bmatrix}$$

a real nonsingular symmetric matrix. Interestingly, note that eig$(\mathbf{A})=\{12.8914, 1.5543+5.0507i, 1.5543-5.0507i\}$, i.e., $\mathbf{A}$ does not satisfy Taussky's condition of symmetrizability (Definition 1). Thus $\mathbf{A}$ in (4.2) is a real symmetrizable matrix of the second kind but not a real symmetrizable of the first kind. This illustrates generality of the proposed approach of symmetrization compared to the conventional approach.

Because the coefficient matrices in the equations of motion of linear vibrations are real we next consider only real square matrices.

## 5 Simultaneous Symmetrization of Two Matrices

Equations of motion of a linear undamped (nongyroscopic) system can be expressed by

$$\mathbf{A}\ddot{\mathbf{x}}(t)+\mathbf{B}\mathbf{x}(t)=\mathbf{0}, \quad t>0, \tag{5.1}$$

where $\mathbf{A}\in\mathbb{R}^{N\times N}$ and $\mathbf{B}\in\mathbb{R}^{N\times N}$ and $\mathbf{x}(t)\in\mathbb{R}^N$ is the vector of the *generalized coordinates.* Were $\mathbf{A}$ and $\mathbf{B}$ symmetric and positive definite matrices they would, respectively, represent the mass and stiffness matrices. Here, however, no such restrictions are imposed on $\mathbf{A}$ and $\mathbf{B}$. Huseyin [23] has shown that the traditional modal analysis, originally developed for symmetric systems, can be extended to solve Eq. (5.1). There exist *right eigenvectors* or *normal modes* and *left eigenvectors* or *adjoint modes* which uncouple the equations of motion into $N$ single-degree-of-freedom oscillators. Although system (5.1) can always be solved using normal modes and adjoint modes, it would be more useful if it can be transformed to a symmetric system. We consider the following two cases to investigate symmetrizability of system (5.1).

**5.1 Case 1: A and B Have Real Roots and A is Nonsingular.** This is the most common case we encounter in practice. If system (5.1) is ''truly'' undamped then $\mathbf{A}$ or $\mathbf{B}$ does not possess any complex roots because, otherwise a stable system will not have periodic motion for an infinitely long time but vibration (in some degrees-of-freedom) will decay due to the complex nature of the eigenvalues. As $\mathbf{A}$ is nonsingular, rewriting Eq. (5.1) one has

$$\mathbf{I}\ddot{\mathbf{x}}+\mathbf{E}\mathbf{x}=\mathbf{0} \tag{5.2}$$

where $\mathbf{E}=\mathbf{A}^{-1}\mathbf{B}\in\mathbb{R}^{N\times N}$ has real roots as both $\mathbf{A}$ and $\mathbf{B}$ have real roots. Because this matrix satisfies Condition 4 of Taussky's definition of symmetrizability, system (5.2) can be transformed to a symmetric system by a similarity transformation. From this discussion we have the following interesting result:

THEOREM 5.1. *All asymmetric undamped systems are similar to symmetric undamped systems.*

Practical implementation of Taussky's factorization (3.1) is not straightforward (see Ahmadian and Chou, [15], Shahruz and Ma [16], and Cherng and Abdelhamid [17]). For this reason it is required to develop efficient numerical methods for finding associated symmetric form(s) of system (5.1). Recall that as system (5.1) is symmetrizable of the first kind it is also symmetrizable of the second kind. We use the latter approach to avoid the calculation of Taussky's factorization.

It has been mentioned that there always exist two matrices $\mathbf{U}\in\mathbb{R}^{N\times N}$ (normal modal matrix) and $\mathbf{V}\in\mathbb{R}^{N\times N}$ (adjoint modal matrix) such that $\mathbf{V}^T\mathbf{AU}$ and $\mathbf{V}^T\mathbf{BU}$ are both real diagonal matrices. Numerical methods for obtaining $\mathbf{U}$ and $\mathbf{V}$ are well developed (see Huseyin [23] and Ma and Caughey [7] for further discussions). Select $\mathbf{L}=\mathbf{VQ}$ and $\mathbf{R}=\mathbf{UQ}$ for some nonzero $\mathbf{Q}\in\mathbb{R}^{N\times N}$. Clearly $\mathbf{L}^T\mathbf{AR}$ and $\mathbf{L}^T\mathbf{BR}$ are both symmetric matrices now. This approach is much easier for finding symmetric forms than the approach via Taussky's factorization. Also note that symmetric forms are nonunique since one can select the matrix $\mathbf{Q}$ in many ways. The following example demonstrates this procedure.

*Example 5.1.* Suppose the coefficient matrices of an undamped system of the form (5.1) are given by

$$\mathbf{A}=\begin{bmatrix} 0.5740 & 1.3858 & 1.3858 \\ 0.7070 & 0.7070 & -0.7070 \\ 0.4620 & -0.1914 & -0.1914 \end{bmatrix}$$

and

$$\mathbf{B}=\begin{bmatrix} 1.3748 & 10.9440 & 25.2975 \\ 1.2625 & 2.8770 & -17.4195 \\ 0.7455 & -4.1244 & 0.8625 \end{bmatrix}. \tag{5.3}$$

Numerical values of $\mathbf{A}$ and $\mathbf{B}$ are taken from Ma and Caughey [7]. It may be verified that $\mathbf{A}$, $\mathbf{B}$ and consequently $\mathbf{E}=\mathbf{A}^{-1}\mathbf{B}$ are symmetrizable matrices of first kind.

Solving the right and left eigenvalue problems $\mathbf{BU}=\mathbf{\Lambda AU}$ and $\mathbf{V}^T\mathbf{B}=\mathbf{\Lambda V}^T\mathbf{A}$ we obtain the undamped modal matrices

$$\mathbf{U}=\begin{bmatrix} 0.9996 & -0.5434 & 0.3951 \\ -0.0271 & 0.8391 & -0.4850 \\ -0.0044 & -0.0253 & 0.7802 \end{bmatrix}$$

and

$$\mathbf{V}=\begin{bmatrix} 0.2868 & 0.4247 & 0.4617 \\ 0.4677 & 0.2613 & -0.7124 \\ 0.8361 & -0.8668 & 0.5285 \end{bmatrix} \tag{5.4}$$

and the natural frequencies squared $\mathrm{diag}(\mathbf{\Lambda})=\{1.8241,\ 9.5561,\ 23.5486\}$. Selecting the matrix $\mathbf{Q}$ given by (4.3) one obtains

$$\mathbf{L}=\mathbf{VQ}=\begin{bmatrix} -0.2860 & 0.1534 & 0.4737 \\ -0.3177 & -0.1826 & -0.0362 \\ 0.9126 & -0.2092 & -0.2176 \end{bmatrix}$$

and

$$\mathbf{R}=\mathbf{UQ}=\begin{bmatrix} 0.6131 & -0.1948 & -0.0445 \\ -0.8286 & 0.0798 & 0.3328 \\ 0.1480 & 0.1744 & 0.2731 \end{bmatrix}. \tag{5.5}$$

Using these matrices we have

$$\tilde{\mathbf{A}}=\mathbf{L}^T\mathbf{AR}=\begin{bmatrix} 0.6282 & -0.1303 & -0.3608 \\ -0.1303 & 0.1032 & 0.1514 \\ -0.3608 & 0.1514 & 0.4150 \end{bmatrix}$$

and

$$\tilde{\mathbf{B}}=\mathbf{L}^T\mathbf{BR}=\begin{bmatrix} 6.2653 & -0.7600 & -2.8429 \\ -0.7600 & 1.3947 & 2.5574 \\ -2.8429 & 2.5574 & 5.3631 \end{bmatrix}. \tag{5.6}$$

Above is a symmetric form of the asymmetric system (5.3). Note that symmetric form of this system can also be obtained alternatively by calculating Taussky's factorization of $\mathbf{E}$.

**5.2 Case 2: A and B Are General Matrices.** We have shown that when $\mathbf{A}$ and $\mathbf{B}$ have real roots, system (5.1) has an equivalent symmetric form via real linear transformations. However, if these matrices have complex roots, the matrix $\mathbf{E}$ in Eq. (5.2) in general does not satisfy Taussky's definition of symmetrizability. In that case, can system (5.1) be transformed to a real symmetric system using a real linear transformation? Our answer is the following:

LEMMA 5.2. *Linear undamped system* (5.1) *is real symmetrizable if* $\mathbf{A}^T\mathbf{B}$ *and* $\mathbf{BA}^T$ *are symmetric matrices.*

*Proof.* According to Thompson [24] ''there exist two real orthogonal matrices $\mathbf{O}_1\in\mathbb{R}^{N\times N}$ and $\mathbf{O}_2\in\mathbb{R}^{N\times N}$ such that square matrices $\mathbf{O}_1^T\mathbf{AO}_2=\mathbf{\Lambda}_a$ and $\mathbf{O}_1^T\mathbf{BO}_2=\mathbf{\Lambda}_b$ are diagonal matrices if and only if $\mathbf{A}^T\mathbf{B}$ and $\mathbf{BA}^T$ are symmetric matrices.'' Now construct $\mathbf{L}=\mathbf{O}_1\mathbf{Q}$ and $\mathbf{R}=\mathbf{O}_2\mathbf{Q}$ for some nonzero $\mathbf{Q}\in\mathbb{R}^{N\times N}$. Using these we have $\mathbf{L}^T\mathbf{AR}=\mathbf{Q}^T\mathbf{\Lambda}_a\mathbf{Q}$ and $\mathbf{L}^T\mathbf{BR}=\mathbf{Q}^T\mathbf{\Lambda}_b\mathbf{Q}$ are both symmetric matrices. This completes the proof. □

This result provides only a *sufficient* condition for simultaneous real symmetrizability of the general matrices $\mathbf{A}$ and $\mathbf{B}$. Lack of it does not necessarily preclude existence of a real symmetric form of system (5.1). The following examples illustrates this fact.

*Example 5.2.* Consider an undamped system with $\mathbf{A}$ same as Example 4.2 and

$$\mathbf{B}=\begin{bmatrix} 1.2044 & -5.4425 & 2.7013 \\ 2.1007 & 0.8393 & 0.1894 \\ -1.8393 & 0.8953 & 2.5087 \end{bmatrix}. \tag{5.7}$$

One may verify that $\mathbf{A}$ and $\mathbf{B}$ are not symmetrizable of the first kind. Observe that the matrices

$$\mathbf{A}^T\mathbf{B}=\begin{bmatrix} 30.0910 & 2.8380 & -0.0427 \\ 2.8380 & 19.5018 & 5.7688 \\ -0.0427 & 5.7688 & 27.9563 \end{bmatrix}$$

and

$$\mathbf{BA}^T=\begin{bmatrix} 16.1414 & 0.7063 & 0.1331 \\ 0.7063 & 31.5697 & 0.8607 \\ 0.1331 & 0.8607 & 29.8380 \end{bmatrix}$$

are both symmetric. Since the system matrices satisfy the condition of Lemma 5.2 the transforming matrices $\mathbf{L}$ and $\mathbf{R}$ (computed before in Eq. (4.4)) which symmetries $\mathbf{A}$ also symmetries $\mathbf{B}$ as

$$\mathbf{L}^T\mathbf{BR}=\begin{bmatrix} 2.8766 & -0.4670 & -1.4786 \\ -0.4670 & 0.5501 & 0.9468 \\ -1.4786 & 0.9468 & 2.1818 \end{bmatrix}$$

is a symmetric matrix.

*Example 5.3.* Consider another undamped system with $\mathbf{A}$ in the above example and

$$\mathbf{B}=\begin{bmatrix} 0.1955 & 1.8857 & -3.2199 \\ -2.1312 & -0.2236 & 0.3609 \\ 1.0378 & 1.9501 & -1.5606 \end{bmatrix}. \tag{5.8}$$

Observe that the matrices

$$\mathbf{A}^T\mathbf{B}=\begin{bmatrix} -27.4545 & -4.6975 & 4.2325 \\ -9.0269 & 2.6874 & 2.3629 \\ -5.2845 & 18.8142 & -16.3493 \end{bmatrix}$$

and

$$\mathbf{BA}^T=\begin{bmatrix} -8.4058 & -8.8789 & -21.8278 \\ -1.1426 & -24.3894 & 6.6164 \\ -5.2033 & 13.2300 & -8.3212 \end{bmatrix}$$

are not symmetric. However, the transforming matrices $\mathbf{L}$ and $\mathbf{R}$ obtained in Eq. (4.2) and also the symmetries $\mathbf{B}$ as

$$\mathbf{L}^T\mathbf{BR}=\begin{bmatrix} -1.2355 & 0.8082 & 1.2380 \\ 0.8082 & -0.4868 & -0.9953 \\ 1.2380 & -0.9953 & -2.0335 \end{bmatrix}$$

is a symmetric matrix. So the condition outlined in Lemma 5.2 is only sufficient.

## 6 Simultaneous Symmetrization of Three Matrices

The equations of motion describing free vibration of a viscously damped linear system can be expressed by

$$\mathbf{A}\ddot{\mathbf{x}}(t)+\mathbf{C}\ddot{\mathbf{x}}(t)+\mathbf{B}\mathbf{x}(t)=\mathbf{0},\quad t>0. \tag{6.1}$$

We assume $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ are $N\times N$ arrays of real numbers but are otherwise general. Were these matrices symmetric and positive definite, $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ would, respectively, be the mass, stiffness, and viscous damping matrices. The equations of motion are now characterized by *three* real matrices and this brings an additional complication in the system dynamics. It is required to find a nonsingular linear transformation which simultaneously symmetries $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$. Unlike the undamped case, where in the absence of

the **C** matrix conditions for obtaining such a transformation were easy to meet (Theorem 5.1), admissible forms of these matrices gets restricted here. We will investigate what forms of restrictions should be imposed on the system matrices so that the equations of motion can be transformed into a symmetric form. The following two cases are considered here.

**6.1 Case 1: A is Nonsingular and All the Matrices Have Real Roots.** When **A** is nonsingular, Eq. (6.1) can be rewritten as

$$\mathbf{I}\ddot{\mathbf{x}}(t)+\mathbf{F}\dot{\mathbf{x}}(t)+\mathbf{E}\mathbf{x}(t)=\mathbf{0}. \tag{6.2}$$

where $\mathbf{E}=\mathbf{A}^{-1}\mathbf{B}\in\mathbb{R}^{N\times N}$ and $\mathbf{F}=\mathbf{A}^{-1}\mathbf{C}\in\mathbb{R}^{N\times N}$. Since **A**, **B**, and **C** have real roots **E** and **F** also have real roots. Thus **E** and **F** are individually symmetrizable of the first kind. But this does not imply that system (6.2) is also symmetrizable, unless **I**, **E**, and **F** are simultaneously symmetrizable. A basic result in this regard is the following:

THEOREM 6.1. *All diagonalizable damped asymmetric systems are similar to symmetric systems.*

It is well known that system (6.2) is diagonalizable by a similarity transformation if and only if **E** and **F** commutes in multiplication, i.e., $\mathbf{EF}=\mathbf{FE}$. When this condition is satisfied the system is symmetrizable of the first kind and hence also symmetrizable of the second kind. Following the approach outlined before in Section 5.1, symmetric form(s) may be obtained by utilizing equivalence transformation to avoid the calculation of Taussky's factorization. A result similar to Theorem 6.1 was also established by Inman ([12], Theorem 1) and Shahruz and Ma ([16], Theorem 3.1). Note that the opposite statement of the above theorem, i.e., ''a damped asymmetric system similar to a symmetric system is diagonalizable'' is in general not true. Thus, diagonalizability of a system is only a sufficient condition for symmetrizability and rather restrictive. The following is a more stronger and liberal result:

THEOREM 6.2. (*Inman* [12], *Theorem* 2) *system* (6.2) *is symmetrizable if and only if* **E** *and* **F** *have a common symmetric positive definite factor.*

From Taussky's factorization (3.1) one can write $\mathbf{E}=\mathbf{S}_1\mathbf{S}_2$, $\mathbf{S}_1=\mathbf{S}_1^T>0$, $\mathbf{S}_2=\mathbf{S}_2^T$ and $\mathbf{F}=\mathbf{T}_1\mathbf{T}_2$, $\mathbf{T}_1=\mathbf{T}_1^T>0$, $\mathbf{T}_2=\mathbf{T}_2^T$. According to the above theorem system (6.2) is symmetrizable if and only if $\mathbf{S}_1=\mathbf{T}_1$. If this condition is satisfied Eq. (6.2) can be brought into symmetric form using a similarity transformation involving a *symmetric* matrix. For practical purposes this condition is difficult to check. A more convenient result is the following:

THEOREM 6.3. (*Kliem* [9], *Theorem 1*) *system* (6.2) *is real symmetrizable if and only if there exist modal matrices* $\mathbf{U}_E\in\mathbb{R}^{N\times N}$ *and* $\mathbf{U}_F\in\mathbb{R}^{N\times N}$ (*consisting of full eigenvector sets of* **E** *and* **F**) *such that* $\mathbf{U}_F^{-1}\mathbf{U}_E$ *is orthogonal, or equivalently* $\mathbf{U}_F\mathbf{U}_F^T=\mathbf{U}_E\mathbf{U}_E^T$.

Further discussions on this theorem can found in Pommer and Kliem [19]. When the condition of the above theorem is satisfied, Eq. (6.2) can be brought into symmetric form by a similarity transformation. In this case the symmetrizing matrix may be any nonzero matrix, unlike a symmetric matrix utilized in Theorem 6.2. For this reason Kliem's [9] condition is more liberal and holds Inman's [12] condition as a special case when the symmetrizing similarity transformation itself is a symmetric matrix. Thus Theorem 6.2, and consequently results obtained by Ahmadian and Chou ([15], Theorem 2) and Shahruz and Ma ([16], Theorem 4.1) based on this theorem are only *sufficient* conditions for symmetrization of the first kind of system (6.1).

Now, symmetrizability of system (6.1) by means of the equivalence transformation will be introduced. Our main result is the following:

THEOREM 6.4. *System* (6.1) *is real symmetrizable of the second kind if and only if there exist* $\mathbf{U}\in\mathbb{R}^{N\times N}$ *and* $\mathbf{V}\in\mathbb{R}^{N\times N}$, *the matrices of undamped right and left eigenvectors, such that* $\mathbf{V}^T\mathbf{CU}$ *is symmetric.*

*Proof.* Because $\mathbf{U}\in\mathbb{R}^{N\times N}$ and $\mathbf{V}\in\mathbb{R}^{N\times N}$ are the matrices of right and left eigenvectors of the generalized eigenvalue problem involving **A** and **B**, $\mathbf{V}^T\mathbf{AU}=\mathbf{\Lambda}_a$ and $\mathbf{V}^T\mathbf{BU}=\mathbf{\Lambda}_b$ are both real diagonal matrices. Select $\mathbf{L}=\mathbf{VQ}$ and $\mathbf{R}=\mathbf{UQ}$, where $\mathbf{Q}\in\mathbb{R}^{N\times N}$ is nonzero. Clearly $\mathbf{L}^T\mathbf{AR}=\mathbf{Q}^T\mathbf{\Lambda}_a\mathbf{Q}$ and $\mathbf{L}^T\mathbf{BR}=\mathbf{Q}^T\mathbf{\Lambda}_b\mathbf{Q}$ are both real symmetric matrices. Using this transformation, the second term of Eq. (6.1) is now $\mathbf{L}^T\mathbf{CR}=\mathbf{Q}^T(\mathbf{V}^T\mathbf{CU})\mathbf{Q}$. Thus the transformed system is symmetric if and only if $\mathbf{V}^T\mathbf{CU}$ is a symmetric matrix. □

Since the equivalence transformations are more general than similarity transformations it is expected that the condition of symmetrizability given by this theorem is less restrictive than the three previously established results discussed before. The following example illustrates this fact:

*Example 6.1.* Suppose, for a damped dynamic system, **A** and **B** are as in Example 5.1 and

$$\mathbf{C}=\begin{bmatrix} 4.1429 & 9.4510 & 9.0014 \\ 4.6465 & 3.8206 & -3.2816 \\ 1.2006 & 0.5460 & -1.5001 \end{bmatrix}. \tag{6.3}$$

Transforming the system in the form of Eq. (6.2) one may verify that both **E** and **F** are individually symmetrizable of the first kind. However, none of the conditions of simultaneous symmetrizability outlined in Theorems 6.1–6.3 is satisfied by **E** and **F.** Now, using the modal matrices **U** and **V** calculated before in Eq. (5.4) we have

$$\mathbf{V}^T\mathbf{CU}=\begin{bmatrix} 4.2300 & 1.7900 & -0.8400 \\ 1.7900 & 26500 & 1.8900 \\ -0.8400 & 1.8900 & 3.2100 \end{bmatrix}$$

is a symmetric matrix, i.e., the condition of Theorem 6.4 is satisfied. The transformation matrices **L** and **R** given by Eq. (5.5) which symmetries **A** and **B** also symmetries **C** as

$$\mathbf{L}^T\mathbf{CR}=\begin{bmatrix} 1.4439 & -0.4735 & -1.8602 \\ -0.4735 & 0.5412 & 0.8591 \\ -1.8602 & 0.8591 & 2.6222 \end{bmatrix}$$

is a symmetric matrix. Thus, the system under consideration is real symmetrizable of the second kind but not the first kind.

**6.2 Case 2: A, B, and C are General Real Matrices.** When the system matrices are general, some or all of them may have complex roots so that they do not satisfy Tausskkey's condition of symmetrizability. For such systems, in general Kliem's result (Theorem 6.3) gives the condition of *complex symmetrizability* of the first kind and Theorem 6.4 gives the condition of complex symmetrizability of the second kind. A *sufficient* condition for real symmetrizability of such systems is the following:

LEMMA 6.5. *System* (6.1) *is real symmetrizable of second kind if there exist* **U** *and* **V**, *the matrices of undamped right and left eigenvectors, such that* $\mathbf{A}^T\mathbf{B}$, $\mathbf{BA}^T$ and $\mathbf{V}^T\mathbf{CU}$ *are all symmetric matrices.*

This lemma can be proved easily following the results of Lemma 5.2 and Theorem 6.4. The example considered below illustrates this result:

*Example 6.2.* Suppose, for a damped linear system, **A** and **B** are given by Example 5.2 and

$$\mathbf{C}=\begin{bmatrix} 4.3160 & -2.5771 & -1.4626 \\ 2.7122 & 1.8365 & -0.1999 \\ 1.3827 & -2.5631 & 4.3419 \end{bmatrix}. \tag{6.4}$$

One may easily verify that none of the system matrices are individually real symmetrizable of the first kind. We compute the undamped modal matrices

$$\mathbf{U}=\begin{bmatrix} 0.6498 & 0.7383 & 0.1807 \\ 0.4364 & -0.1676 & -0.8840 \\ 0.6224 & -0.6533 & 0.4311 \end{bmatrix}$$

and

$$\mathbf{V}=\begin{bmatrix} 0.0445 & -0.0108 & 0.990 \\ 0.9245 & -0.3784 & -0.0453 \\ 0.3785 & 0.9256 & -0.0068 \end{bmatrix}$$

and notice that

$$\mathbf{V}^T\mathbf{C}\mathbf{U}=\begin{bmatrix} 3.2290 & 1.3664 & 0.6412 \\ 1.3664 & -2.0229 & 4.4962 \\ 0.6412 & 4.4962 & 2.5404 \end{bmatrix}$$

is a symmetric matrix. So all the conditions outlined in Lemma 6.5 are satisfied. Using the matrix $\mathbf{Q}$ given by (4.3) one has

$$\mathbf{L}=\mathbf{V}\mathbf{Q}=\begin{bmatrix} 0.1735 & 0.2198 & 0.3701 \\ 0.3901 & -0.2469 & -0.1170 \\ -0.8026 & 0.1421 & 0.6215 \end{bmatrix}$$

and

$$\mathbf{R}=\mathbf{U}\mathbf{Q}=\begin{bmatrix} -0.5878 & 0.0988 & 0.6148 \\ 0.0353 & -0.3123 & -0.3683 \\ 0.6926 & -0.1489 & -0.1529 \end{bmatrix}.$$

These real matrices transform the system to a real symmetric form.

## 7 Concluding Discussions and Further Research

A method of symmetrization of asymmetric linear dynamical systems by means of equivalence transformation is introduced. This new method of symmetrization is called ''symmetrization of the second kind'' and the existing symmetrization method based on the similarity transformation is called ''symmetrization of first kind.'' Because the equivalence transformations are most general linear transformations, symmetrization of the second kind holds that of the first kind as a special case and offers a generalized method to obtain symmetric forms of finite dimensional asymmetric linear systems.

Several results are provided on symmetrization of the second kind. We have discussed real and complex symmetrizability based on the nature of the symmetrizing matrices. It was proved that every rectangular complex matrix is complex symmetrizable of the second kind to a real square matrix. It is possible to perform the symmetrization in such a way that the resulting symmetric form becomes positive definite. As a special case it was shown that all real matrices are real symmetrizable of the second kind. These results are more general than symmetrizability of the first kind which cannot be applied to rectangular matrices. An easy numerical method is presented to calculate symmetric forms of general matrices. This method is much easier than calculating the symmetric form of the first kind using Taussky's factorization.

In view of the undamped and damped dynamical systems, we have discussed simultaneous symmetrizability of two and three real square matrices. It was shown that (truly) undamped asymmetric systems are always real symmetrizable. The condition of real symmetrizability of general undamped systems is derived. We also have given a condition for real symmetrizability of the second kind for damped systems. Our result is more liberal than the existing results and also much easier to check from a numerical point view.

Further research is required to successfully apply the concept of symmetrizability of the second kind in asymmetric linear dynamical systems. By an example it was shown that the condition of symmetrizability of a damped system given by Theorem 6.4 is less restrictive than the existing results. A mathematical proof to establish this fact will be useful. Moreover, this condition is spectral (i.e., an eigensolution calculation is required), a nonspectral condition which is more easy to check needs to be developed. Some of our results (Lemma 5.2 and Lemma 6.5) provide only sufficient conditions for real symmetrizability of systems with general coefficient matrices. More stronger results are required in these directions. Further studies regarding stability of the symmetrizing transformations and the robustness of them with respect to perturbations in the entries of the coefficient matrices are also worth pursuing.


## Acknowledgments

Funding provided by the Nehru Memorial Trust, London and the Cambridge Commonwealth Trust is greatly acknowledged.


## References


[1] Huseyin, K., and Liepholz, H. H. E., 1973, ''Divergence Instability of Multiple Parameter Circulatory Systems,'' Q. Appl. Math., **31**, pp. 185–197.

[2] Fawzy, I., and Bishop, R. E. D., 1977, ''On the Nature of Resonance in Nonconservative Systems,'' J. Sound Vib., **55**, No. 4, pp. 475–485.

[3] Bishop, R. E. D., and Price, W. G., 1979, ''An Investigation Into the Linear Theory of Ship Response to Waves,'' J. Sound Vib., **62**, No. 3, pp. 353–363.

[4] Soom, A., and Kim, C., 1983, ''Roughness-Induced Dynamic Loading at Dry Boundary-Lubricated Sliding Contacts,'' ASME J. Lubr. Technol., **105**, pp. 514–517.

[5] Caughey, T. K., and Ma, F., 1993, ''Complex Modes and Solvability of Nonclassical Linear Systems,'' ASME Trans. J. Appl. Mech., **60**, pp. 26–28.

[6] Fawzy, I., and Bishop, R. E. D., 1976, ''On the Dynamics of Linear Nonconservative Systems,'' Proc. R. Soc. London, Ser. A, **A352**, pp. 25–40.

[7] Ma, F., and Caughey, T. K., 1995, ''Analysis of Linear Conservative Vibrations,'' ASME J. Appl. Mech., **62**, pp. 685–691.

[8] Adhikari, S., 1999, ''Modal Analysis of Linear Asymmetric Nonconservative Systems,'' J. Eng. Mech., **125**, No. 12, pp. 1372–1379.

[9] Kliem, W. R., 1992, ''Symmetrizable Systems in Mechanics and Control Theory,'' ASME Trans. J. Appl. Mech., **69**, pp. 454–456.

[10] Taussky, O., 1968, ''Positive Definite Matrices and Their Role in the Study of the Characteristic Roots of General Matrices,'' Adv. Math., **2**, pp. 175–186.

[11] Taussky, O., and Zassenhaus, H., 1959, ''On the Similarity Transformation Between a Matrix and Its Transpose,'' Pac. J. Math., **9**, pp. 893–896.

[12] Inman, D. J., 1983, ''Dynamics of Asymmetric Nonconservative Systems,'' ASME Trans. J. Appl. Mech., **50**, pp. 199–203.

[13] Ahmadian, M., and Inman, D. J., 1984, ''Classical Normal Modes in Asymmetric Nonconservative Dynamic Systems,'' AIAA J., **22**, No. 7, pp. 1012–1015.

[14] Ahmadian, M., and Inman, D. J., 1984, ''On the Nature of Eigenvalues of General Nonconservative Systems,'' ASME Trans. J. Appl. Mech., **51**, pp. 193–194.

[15] Ahmadian, M., and Chou, S.-H., 1987, ''A New Method for Finding Symmetric Form of Asymmetric Finite-Dimensional Dynamic Systems,'' ASME Trans. J. Appl. Mech., **54**, pp. 700–705.

[16] Shahruz, S. M., and Ma, F., 1989, ''On Symmetrizability of Asymmetric Nonconservative Systems,'' ASME Trans. J. Appl. Mech., **56**, pp. 474–476.

[17] Cherng, A., and Abdelhamid, M. K., 1993, ''On the Symmetrization of Asymmetric Finite Dimensional Linear Dynamic Systems,'' ASME J. Vibr. Acoust., **115**, pp. 417–421.

[18] Taussky, O., 1972, ''The Role of Symmetric Matrices in the Study of General Matrices,'' Linear Algebr. Appl., **5**, pp. 147–154.

[19] Pommer, C., and Kliem, W. R., 1987, ''Simultaneously Normalizable Matrices,'' Linear Algebr. Appl., **94**, pp. 113–125.

[20] Eckrat, C., and Young, G., 1939, ''A Principal Axis Transformation for Non-Hermitian Matrices,'' Bull. Am. Math. Soc., **45**, pp. 118–121.

[21] Williamson, J., 1939, ''Note on a Principal Axis Transformation for Non-Hermitian Matrices,'' Bull. Am. Math. Soc., **45**, pp. 920–922.

[22] Pearl, M. H., 1967, ''A Decomposition Theorem for Matrices,'' Can. J. Math., **19**, pp. 344–349.

[23] Huseyin, K., 1978, *Vibration and Stability of Multiple Parameter Systems*, Sijthoff and Noordhoff, Alpen aan den Rijn, The Netherlands.

[24] Thompson, R. C., 1969, ''On Pearl's Paper 'A Decomposition Theorem for Matrices,''' Canadian Mathematical Bulletin **12**, pp. 805–808.

**S. Hao**
Senior Research Associate
e-mail: suhao@tam.mech.nwu.edu

**W. K. Liu**[2]
Professor of Mechanical Engineering
Fellow ASME
e-mail: w-liu@nwu.edu

**D. Qian**
Research Assistant
e-mail: dqian@tam.mech.nwu.edu

Department of Mechanical Engineering,
Northwestern University,
2145 Sheridan Road,
Evanston, IL 60208

# Localization-Induced Band and Cohesive Model[1]

*A localization-induced cohesive model has been proposed for shear band evolution, crack growth, and fracture. Strain gradient theory has been applied to establish the criterion of the onset of localization and the governing equation in the post-bifurcation stage. Analytical solutions in one-dimensional case are used to establish the ''traction-separation'' law, in which strain gradient and material intrinsic length scale present strong effects. In addition, the solution predicts a finite width for the localization-induced band. It is observed that a larger length scale contributes to the growth of a larger width of localization region and separation for softening materials. The proposed model provides a procedure to establish the fracture toughness analytically since the material length scale is taken into account. From the traction-separation analysis, it is found that damage decreases separation, whereas an increase in material length scale increases the opening displacement; however, the traction-normalized opening displacement curves (with respect to the material length scale) are identical. Based on the methodology of multiple scale analysis in meshfree method, a computational approach has been proposed to enrich the one-dimensional traction-separation law to define fracture.*
[S0021-8936(00)01104-1]

## 1 Introduction

Fracture and localization are multiple scale physical processes. The onset and the subsequent development of strain localization during material de-cohesion process focuses energy into a very small region. This highlights the length scale and micromechanisms of the phenomena. In computational implementation, various constitutive laws or the relations derived based on continuum mechanics have been applied for simulation. For instance, the cohesive model ([1,2]) has been widely used in many aspects of fracture mechanics. In the localization analysis, numerous nonlocal constitutive models are proposed to achieve shear bands independent of the computational grid size. Under dynamic loading with heat conduction, the adiabatic shear band has been investigated in Wright and Walter [3] and DiLellio and Olmstead [4].

In this paper, localization analysis is conducted based on the strain gradient theory proposed by Fleck, Muller, Ashby, and Hutchinson [5] and Gao, Huang, Nix, and Hutchinson [6]. The corresponding ''traction-separation'' relation is derived. Using this relation we propose a localization-induced band and cohesive model. The effect of the material intrinsic length scale is included in this model. A multiple scale numerical procedure based on reproducing kernel particle method (RKPM) is implemented in the solution. Several interesting results are reported. It is concluded that material intrinsic length scale has a significant effect on the localization analysis. The proposed multiple scale model, which is based on the mechanism-based strain gradient theory, resolves the intrinsic length scale regardless of the discretization. Finally, the authors also wish to point out that the proposed multiple scale approach can be applied to any strain gradient model, although an additional mountain of work is needed.

The paper is organized as follows: A brief review of the model and the establishment of the governing equation are presented in Sections 2 and 3, respectively. Section 4 and 5 give the details on the localization analysis. The cohesive model based on strain gradient plasticity is established in Section 6. After the demonstration of the proposed multiple scale numerical approach in Section 7, conclusions are made in Section 8.

## 2 Review of Strain Gradient Plasticity

**2.1 Classical Deformation Plasticity.** The classical plasticity model states that the plastic strain rate, $\dot{\boldsymbol{\varepsilon}}^p$, is proportional to the gradient of the plastic potential, $F$, with respect to the stress tensor $\boldsymbol{\sigma}$:

$$\dot{\boldsymbol{\varepsilon}}^p = \dot{\lambda}\frac{\partial F}{\partial \boldsymbol{\sigma}} \tag{2.1}$$

where $\dot{\lambda}$ is the flow factor; $\dot{\boldsymbol{\varepsilon}}$ is the strain rate tensor and its components are related to the velocity field $\mathbf{v}$ by

$$\dot{\boldsymbol{\varepsilon}} = \frac{1}{2}[\mathbf{v}\nabla + \nabla\mathbf{v}] \quad \text{or} \quad \dot{\varepsilon}_{ij} = \frac{1}{2}(v_{i,j} + v_{j,i}). \tag{2.2}$$

The superposed dot represents the time derivative and the superscript ''$p$'' denotes the plastic part of a quantity. The normality rule applies, $F = F(\bar{\sigma})$, and the flow stress $\bar{\sigma}$ is defined by

$$\bar{\sigma} = \sigma_Y f(\bar{\varepsilon}); \quad \bar{\varepsilon} = \left[\frac{2}{3}\varepsilon_{ij}\varepsilon_{ij}\right]^{1/2} \tag{2.3}$$

where $\sigma_Y$ is the yield stress; $\bar{\varepsilon}$ is the equivalent strain; $f(\bar{\varepsilon})$ is calibrated from the uniaxial stress-strain relation, i.e., $\bar{\sigma}/\sigma_Y$ versus $\bar{\varepsilon}/\varepsilon_Y$; $\varepsilon_Y$ is the yield strain; and $\boldsymbol{\sigma}$ can be expressed as the sum of the deviatoric stress tensor $\boldsymbol{\sigma}^d$ and the mean stress $p\mathbf{I}$, that is

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}^d + p\mathbf{I}; \quad p = \frac{\sigma_{ii}}{3}. \tag{2.4}$$

$\mathbf{I}$ is the identity tensor; repeated indices imply summation.

Assuming small strain, the elastic portion of strain rate, $\dot{\boldsymbol{\varepsilon}}^e$, is given by

$$\dot{\boldsymbol{\varepsilon}}^e = \frac{1+\nu}{E}\dot{\boldsymbol{\sigma}}^d + \frac{1-2\nu}{E}\dot{p}\mathbf{I} \quad \text{or} \quad \dot{\boldsymbol{\varepsilon}}^e = \mathbf{L}_e : \dot{\boldsymbol{\sigma}} \tag{2.5}$$

[1]This is an extended version of the paper presented at the Multi-scale Analysis symposium of the ASME Mechanics and Materials Conference, June 27–30, 1999, Blacksburg, VA.

[2]To whom correspondence should be addressed.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received and accepted by the ASME Applied Mechanics Division, March 7, 2000. Associate Technical Editor: B. M. Moran. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

 Copyright © 2000 by ASME 

$\mathbf{L}_e$ is the elastic compliance tensor; $\nu$ and $E$ are the Poisson's ratio and Young's modulus, respectively.

Combining Eq. (2.1) and Eq. (2.5), the constitutive law can be written as

$$\dot{\boldsymbol{\sigma}}=\mathbf{C}:\dot{\boldsymbol{\varepsilon}};\quad \dot{\boldsymbol{\varepsilon}}=\dot{\boldsymbol{\varepsilon}}^{e}+\dot{\boldsymbol{\varepsilon}}^{p};\quad \mathbf{C}=\mathbf{L}_e^{-1}-\frac{\left(\mathbf{L}_e^{-1}:\dfrac{\partial F}{\partial\boldsymbol{\sigma}}\right)\left(\dfrac{\partial F}{\partial\boldsymbol{\sigma}}:\mathbf{L}_e^{-1}\right)}{f'(\overline{\varepsilon})+\dfrac{\partial F}{\partial\boldsymbol{\sigma}}:\mathbf{L}_e^{-1}:\dfrac{\partial F}{\partial\boldsymbol{\sigma}}}. \tag{2.6}$$

**2.2 Rate Form of Strain Gradient Plasticity.** The rate form of a strain gradient tensor is defined by

$$\dot{\boldsymbol{\eta}}=\nabla\nabla\mathbf{v}\ \text{ or }\ \dot{\eta}_{ijk}=v_{k,ij}=\frac{\partial^2 v_k}{\partial x_i\partial x_j}. \tag{2.7}$$

The strain gradient theory employed in this paper has the following properties:

- A material intrinsic length scale $l$, which relates to the density of the geometrically necessary dislocation for a single-slip system of a single crystal, is defined as ([7])

$$l=3\left[\frac{\alpha\mu}{\sigma_Y}\right]^2 b \tag{2.8}$$

  where $b$ is Burger's vector, $\mu$ denotes the shear modulus, and $\alpha$ is a constant.
- At a microscale level, the flow stress is governed by the dislocation motion ([6])

$$\overline{\sigma}=\sigma_Y\sqrt{f^2(\overline{\varepsilon})+l\,\overline{\eta}}\ \text{ and }\ \overline{\eta}=\sqrt{\frac{1}{4}\,\eta_{ijk}\eta_{ijk}}. \tag{2.9}$$

  Equation (2.9) is referred to as the Taylor's hardening relation.
- In our study, we discovered that the effect from the couple stress introduced by Gao et al. is relatively small. Hence, the effect from the couple stress is ignored in this paper.
- Incompressibility and small strain theory are assumed.

With the above assumptions, the rate form of the strain gradient plasticity can be obtained by the time differentiation of the deformation plasticity relation,

$$\boldsymbol{\sigma}^d=\frac{2}{3}\frac{\boldsymbol{\varepsilon}}{\overline{\varepsilon}}\overline{\sigma} \tag{2.10}$$

with the flow stress given by Eq. (2.9):

$$\dot{\boldsymbol{\sigma}}^d=\overline{\mathbf{C}}^{\sigma\varepsilon}:\dot{\boldsymbol{\varepsilon}}+\mathbf{C}^{\sigma\eta}\vdots\,\dot{\boldsymbol{\eta}}. \tag{2.11}$$

The components of $\mathbf{C}^{\sigma\varepsilon}$ and $\mathbf{C}^{\sigma\eta}$ are

$$\overline{C}^{\sigma\varepsilon}_{ijkl}=\frac{2\overline{\sigma}}{3\overline{\varepsilon}}\left[\delta_{ik}\delta_{jl}+\frac{2}{3}\,\frac{f(\overline{\varepsilon})f'(\overline{\varepsilon})\varepsilon_{ij}\varepsilon_{kl}}{\overline{\varepsilon}(f^2(\overline{\varepsilon})+l\,\overline{\eta})}-\frac{2\varepsilon_{ij}\varepsilon_{kl}}{3\overline{\varepsilon}^2}\right] \tag{2.12}$$

and

$$C^{\sigma\eta}_{ijklm}=\frac{\eta_{klm}}{4\overline{\eta}}\,\frac{2\overline{\sigma}}{3\overline{\varepsilon}}\,\frac{\varepsilon_{ij}l}{2(f^2(\overline{\varepsilon})+l\,\overline{\eta})}. \tag{2.13}$$

The traction rate along a natural boundary is given as

$$\dot{\mathbf{t}}=\mathbf{n}\cdot(\dot{\boldsymbol{\sigma}}^d+\overline{p}\mathbf{I}). \tag{2.14}$$

**2.3 Damage.** To incorporate damage, we employ a simple damage law such that the yield stress in Eq. (2.9) is replaced by

$$\sigma_Y=\sigma_{Y0}(1-D) \tag{2.15}$$

where $\sigma_{Y0}$ is the yield strength without damage. The linear damage evolution law by Tvergaard [8] is applied:

$$\dot{D}=\kappa\dot{\overline{\varepsilon}}^p \tag{2.16}$$

where $\kappa$ is a material constant, $\dot{\overline{\varepsilon}}^p$ is the rate of nonlinear strain invariant, which is defined by $\overline{\varepsilon}^p=\overline{\varepsilon}-\overline{\varepsilon}_Y$. With the inclusion of the above damage, the deformation plasticity is still governed by Eq. (2.10), and the rate form of strain gradient plasticity with damage becomes

$$\dot{\boldsymbol{\sigma}}^d=\overline{\mathbf{C}}^{\sigma\varepsilon}:\dot{\boldsymbol{\varepsilon}}+\mathbf{C}^{\sigma\eta}\vdots\,\dot{\boldsymbol{\eta}}+\overline{\mathbf{C}}^{\sigma\varepsilon^p}:\dot{\boldsymbol{\varepsilon}}^p \tag{2.17a}$$

or

$$\dot{\boldsymbol{\sigma}}^d=\mathbf{C}^{\sigma\varepsilon}:\dot{\boldsymbol{\varepsilon}}+\mathbf{C}^{\sigma\eta}\vdots\,\dot{\boldsymbol{\eta}}. \tag{2.17b}$$

In Eq. (2.17*a*) the components of the additional coefficient tensor $\overline{\mathbf{C}}^{\sigma\varepsilon^p}$ are given by

$$\overline{C}^{\sigma\varepsilon^p}_{ijkl}=-\frac{4\overline{\sigma}}{9\overline{\varepsilon}^2}\,\frac{\sigma_{Y0}}{\sigma_Y}\,\kappa\varepsilon_{ij}\varepsilon^p_{kl}. \tag{2.18}$$

Comparing Eq. (2.17*a*) with Eq. (2.11), it is found that the first two terms are identical to while the third term is a result of the damage law.

## 3 Localization for Strain Gradient Plasticity

**3.1 Preliminaries and Assumptions.** Consider the boundary value problem for a homogeneous body $\Lambda$ as shown in Fig. 1(*a*); it is assumed that a strain localization width $w$ takes place along $\Gamma$ with normal vector $\mathbf{n}$ and tangent unit vector $\mathbf{m}$ defined

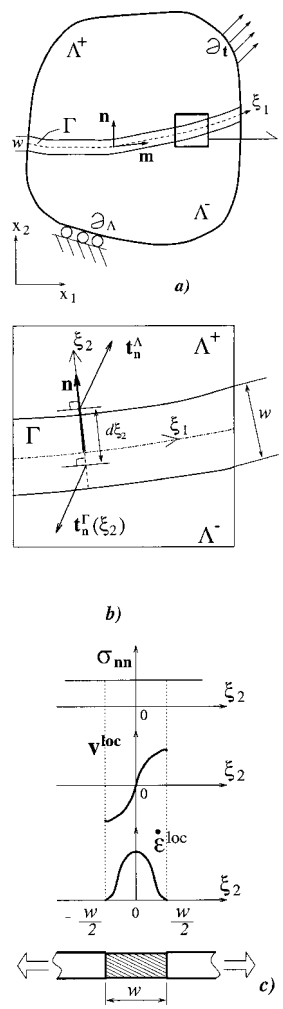



**Fig. 1 Kinematic representation of a localization zone**

with respect to the centerline of $\Gamma$. As shown in Fig. 1(*b*), $\xi_1$ and $\xi_2$ are the coordinates in $\mathbf{m}$ and $\mathbf{n}$ directions, respectively. The following assumptions are made in the analysis:

1 For the localization zone shown in Fig. 1, it is assumed that the velocity field can be separated into two parts ([9])

$$\mathbf{v}=\mathbf{v}^{\text{hom}}+\mathbf{v}^{\text{loc}} \tag{3.1}$$

where the superscript ''hom'' refers to the homogeneously deformed solid outside the localization zone. Superscript ''loc'' refers to the additional deformation due to the localization.

2 The localized velocity $\mathbf{v}^{\text{loc}}$ is assumed to be independent of $\xi_1$, i.e., $\partial\mathbf{v}^{\text{loc}}/\partial\xi_1=0$ ([9]).

3 The gradient of the width of the localization zone $w$ respects to the local coordinate $\xi_1$ is assumed to be small, i.e., $\partial w/\partial\xi_1 \ll 1$.

4 The gradient of the shear stress rate along $\xi_1$ is also small, in other words, $\partial\sigma_{\alpha\beta}/\partial\xi_1 \ll 1$ when $\alpha\neq\beta$.

5 The gradient of the pressure rate is small along $\xi_2$, i.e., $\partial\dot{p}/\partial\xi_2 \ll 1$.

6 Inertial effect is only considered at the onset of localization. It is ignored in the post-bifurcation analysis.

From the postulations introduced above, the strain and strain gradient tensors can be expressed as

$$\dot{\boldsymbol{\varepsilon}}=\dot{\boldsymbol{\varepsilon}}^{\text{hom}}+\dot{\boldsymbol{\varepsilon}}^{\text{loc}} \qquad \dot{\boldsymbol{\eta}}=\dot{\boldsymbol{\eta}}^{\text{hom}}+\dot{\boldsymbol{\eta}}^{\text{loc}}. \tag{3.2}$$

The corresponding components can be shown as

$$\dot{\boldsymbol{\varepsilon}}^{\text{loc}}=\frac{1}{2}(\mathbf{gn}+\mathbf{ng}) \quad \text{or} \quad \dot{\varepsilon}_{ij}^{\text{loc}}=\frac{1}{2}\left(\frac{\partial v_i^{\text{loc}}}{\partial\xi_2}\frac{\partial\xi_2}{\partial x_j}+\frac{\partial v_j^{\text{loc}}}{\partial\xi_2}\frac{\partial\xi_2}{\partial x_i}\right) \tag{3.3}$$

$$\dot{\boldsymbol{\eta}}^{\text{loc}}=\mathbf{qnn}+\overline{\mathbf{g}}\mathbf{mm} \quad \text{or} \quad \dot{\eta}_{ijk}^{\text{loc}}=\frac{\partial^2 v_k^{\text{loc}}}{\partial\xi_2^2}\frac{\partial\xi_2}{\partial x_i}\frac{\partial\xi_2}{\partial x_j}+\vartheta\frac{\partial v_k^{\text{loc}}}{\partial\xi_2}\frac{\partial\xi_1}{\partial x_i}\frac{\partial\xi_1}{\partial x_j} \tag{3.4}$$

where

$$\mathbf{g}=\frac{\partial\mathbf{v}^{\text{loc}}}{\partial\xi_2}; \quad \mathbf{q}=\frac{\partial^2\mathbf{v}^{\text{loc}}}{\partial\xi_2^2}; \quad \overline{\mathbf{g}}=\vartheta\mathbf{g};$$
$$\vartheta=\left[\left(\frac{\partial^2 x_1}{\partial\xi_1^2}\right)^2+\left(\frac{\partial^2 x_2}{\partial\xi_1^2}\right)^2\right]^{-1/2}. \tag{3.5}$$

**3.2 Governing Equations.** With reference to the free-body diagram shown in Fig. 1(*b*), the equilibrium equation can be obtained by applying conservation of linear momentum along the normal direction. In order to determine the relationship between the time scale and the material length scale, the inertia effect is included in the onset of localization analysis to give

$$\dot{\mathbf{t}}_n^{\Lambda}+\dot{\mathbf{t}}_n^{\Gamma}=\int_{\Gamma}\rho\ddot{\mathbf{v}}d\xi_1. \tag{3.6}$$

From the eigenanalysis given in Section 4, it can be deduced that the solution is proportional to $e^{\sqrt{(\lambda_n/\rho)\sigma_Y}t}$. The $n$th eigenvalue $\lambda_n$ is to be defined in Eq. (4.9*a*). It has the following properties:

$$\lambda_n\propto\left(\frac{w}{l^3}\right). \tag{3.7}$$

Since $w$ usually has the same order of magnitude as $l$, therefore

$$\lambda_n\propto\left(\frac{1}{l^2}\right). \tag{3.8}$$

From Eq. (3.8), one can estimate that $\lambda_n$ is in the order of $10^{12}/\text{m}^2$ considering $l$ to be around 1 $\mu$m. According to this estimate, let $\sigma_Y=100$ MPa and $\rho=8000$ kg/m$^3$, it can be further concluded that the decaying ratio of the response $\sqrt{(\lambda_n/\rho)\sigma_Y}$ is approximately in the order of $10^8$/sec. This indicates that the acceleration cannot be ignored. In the post-bifurcation stage, we consider only the static case and the acceleration term is neglected.

$$\dot{\mathbf{t}}_n^{\Lambda}=\dot{\mathbf{t}}_n^{\Gamma} \tag{3.9}$$

is the governing equation for post-bifurcation. However, for fast transient problems such as high velocity impact and penetration, the inertial effect should be included.

According to Eqs. (3.2)–(5), the governing equation for $\mathbf{v}^{\text{loc}}$ in a localization zone can be obtained from the fact that strain and strain gradient are functions of $\mathbf{v}^{\text{loc}}$. Using the traction rate (2.14) and the rate form of gradient plasticity, we derive the governing equations for the following two cases:

*3.2.1 Onset of Localization.* The equilibrium equation Eq. (3.6) becomes

$$\begin{aligned}&(\mathbf{n}\cdot\mathbf{C}^{\sigma\varepsilon}|_{\Gamma}\cdot\mathbf{n})\cdot\mathbf{g}+\mathbf{n}\cdot\mathbf{C}^{\sigma\eta}|_{\Gamma}\vdots(\overline{\mathbf{g}}\mathbf{mm})+\mathbf{n}\cdot\mathbf{C}^{\sigma\eta}|_{\Gamma}\vdots(\mathbf{qnn})\\&\quad+\mathbf{n}\cdot(\mathbf{C}^{\sigma\varepsilon}|_{\Gamma}-\mathbf{C}^{\sigma\varepsilon}|_{\Lambda}):\dot{\boldsymbol{\varepsilon}}^{\text{hom}}+\mathbf{n}\cdot(\mathbf{C}^{\sigma\eta}|_{\Gamma}-\mathbf{C}^{\sigma\eta}|_{\Lambda})\vdots\dot{\boldsymbol{\eta}}^{\text{hom}}\\&\quad=\int_{\Gamma}\rho\ddot{\mathbf{v}}d\xi_1.\end{aligned} \tag{3.10}$$

The equilibrium equation (Eq. (3.10)) can be rewritten in terms of $\mathbf{v}^{\text{loc}}$ by expressing $\mathbf{g}$ and $\mathbf{q}$ in terms of the first and second derivatives of $\mathbf{v}^{\text{loc}}$. The governing equation of localization, which is a second-order partial differential equation, can be shown to be

$$\begin{aligned}&\begin{bmatrix}\overline{A}_{11} & \overline{A}_{12}\\ \overline{A}_{21} & \overline{A}_{22}\end{bmatrix}\begin{Bmatrix}\dfrac{\partial^2 v_1^{\text{loc}}}{\partial\xi_2^2}\\ \dfrac{\partial^2 v_2^{\text{loc}}}{\partial\xi_2^2}\end{Bmatrix}+\begin{bmatrix}\overline{B}_{11} & \overline{B}_{12}\\ \overline{B}_{21} & \overline{B}_{22}\end{bmatrix}\begin{Bmatrix}\dfrac{\partial v_1^{\text{loc}}}{\partial\xi_2}\\ \dfrac{\partial v_2^{\text{loc}}}{\partial\xi_2}\end{Bmatrix}\\&\quad=\begin{Bmatrix}C_{11}\\ C_{22}\end{Bmatrix}+\begin{Bmatrix}D_{11}\\ D_{22}\end{Bmatrix}.\end{aligned} \tag{3.11}$$

The $\overline{A}_{ij}$ and $\overline{B}_{ij}$ coefficients are functions of the components of $\mathbf{C}^{\sigma\varepsilon}$ and $\mathbf{C}^{\sigma\eta}$. $C_{11}$ and $C_{22}$ represent the inertia effect. $D_{11}$ and $D_{22}$ are constants.

*Remark*: (1) If the couple stress is kept in the formulation, the governing equations for onset of localization and post-bifurcation can also be obtained following the usual conservation of linear and angular momentum. (2) For hyperelasticity in which the material length scale $l$ is set to be zero, the above equation reduces to the governing equation obtained by Hill [10]. If the effects of strain gradient and inertia are ignored and $\mathbf{C}^{\sigma\varepsilon}|_{\Gamma}=\mathbf{C}^{\sigma\varepsilon}|_{\Lambda}$ in the close neighborhood of the localization zone, Eq. (3.11) degenerates to $\mathbf{Q}\cdot\mathbf{g}=\mathbf{0}$ where $\mathbf{Q}=\mathbf{n}\cdot\mathbf{C}^{\sigma\varepsilon}\cdot\mathbf{n}$ is the *acoustic* tensor defined in Rice [9] and Rudnicki and Rice [11].

*3.2.2 Post-bifurcation of Localization.* Because of the strain gradient effect, the $\mathbf{C}^{\sigma\varepsilon}$ and $\mathbf{C}^{\sigma\eta}$ matrices in the localization zone are different from the ones in the homogeneous region. Hence, the conservation of linear momentum requires

$$\begin{aligned}&[\mathbf{n}\cdot(\mathbf{C}^{\sigma\varepsilon}|_{\Gamma}-\mathbf{C}^{\sigma\varepsilon}|_{\Lambda})\cdot\mathbf{n}]\cdot\mathbf{g}+\mathbf{n}\cdot\mathbf{C}^{\sigma\eta}|_{\Gamma}\vdots(\overline{\mathbf{g}}\mathbf{mm})+\mathbf{n}\cdot\mathbf{C}^{\sigma\eta}|_{\Gamma}\vdots(\mathbf{qnn})\\&\quad+\mathbf{n}\cdot(\mathbf{C}^{\sigma\varepsilon}|_{\Gamma}-\mathbf{C}^{\sigma\varepsilon}|_{\Lambda}):\dot{\boldsymbol{\varepsilon}}^{\text{hom}}+\mathbf{n}\cdot(\mathbf{C}^{\sigma\eta}|_{\Gamma}-\mathbf{C}^{\sigma\eta}|_{\Lambda})\vdots\dot{\boldsymbol{\eta}}^{\text{hom}}=\mathbf{0}.\end{aligned} \tag{3.12}$$

Equation (3.12) governs the subsequent deformation at post-bifurcation stage. They are used as the governing equations in the proposed localization-induced cohesive model. Similarly, Eq. (3.12) can be written as a second-order partial differential equation as

$$\begin{bmatrix} \bar{A}_{11} & \bar{A}_{12} \\ \bar{A}_{21} & \bar{A}_{22} \end{bmatrix} \begin{Bmatrix} \dfrac{\partial^2 v_1^{\text{loc}}}{\partial \xi_2^2} \\ \dfrac{\partial^2 v_2^{\text{loc}}}{\partial \xi_2^2} \end{Bmatrix} + \begin{bmatrix} \bar{B}_{11} & \bar{B}_{12} \\ \bar{B}_{21} & \bar{B}_{22} \end{bmatrix} \begin{Bmatrix} \dfrac{\partial v_1^{\text{loc}}}{\partial \xi_2} \\ \dfrac{\partial v_2^{\text{loc}}}{\partial \xi_2} \end{Bmatrix} = \begin{Bmatrix} C_{11} \\ C_{22} \end{Bmatrix}. \tag{3.13}$$

In the post-bifurcation stage, $C_{11}$ and $C_{22}$ components are also functions of material response matrix coefficients.

## 4 One-Dimensional Case

**4.1 Boundary Conditions.** For the one-dimensional rod shown in Fig. 1(*c*), the only nonzero velocity component is $v(\xi_2)$. In the subsequent analysis, we let $\xi_2 = x_1$ because of the one-dimensional simplification. The localization analysis is studied by applying a velocity $\bar{v}$ at both ends of the bar in the opposite direction so that

$$v\left(+\frac{L}{2}\right) = \bar{v} \tag{4.1a}$$

and

$$v\left(-\frac{L}{2}\right) = -\bar{v} \tag{4.1b}$$

where $L$ is the length of the rod.

As shown in Fig. 1(*c*), the nonzero of strain rate component at the right boundary of the localization zone (i.e., $x_1 = w/2$) is

$$\left.\frac{dv}{dx_1}\right|_{x_1=+w/2} = \frac{dv^{\text{hom}}}{dx_1}. \tag{4.2}$$

From symmetry, we only consider half of the model. Due to the fact that the symmetric structure is under antisymmetric loading, $v^{\text{loc}}$ should also be antisymmetric. Therefore

$$v^{\text{loc}}|_{x_1=0} = 0. \tag{4.3}$$

**4.2 Onset of Localization Solution.** Right before onset of localization, a uniformly distributed strain $\bar{\varepsilon}_0^{\text{hom}}$ exists over a one-dimensional bar with a corresponding equivalent stress $\bar{\sigma} = \sigma_Y$. At or after the onset of localization, each material point undergoes either localized deformation or elastic unloading. The $f(\bar{\varepsilon})$ in Eq. (2.9) can be given as a piecewise linear function:

$$f(\bar{\varepsilon}) = \begin{cases} \dfrac{E}{\sigma_Y}\left[\bar{\varepsilon} - \left(\bar{\varepsilon}_0^{\text{hom}} - \dfrac{\sigma_Y}{E}\right)\right] & \bar{\varepsilon}_0^{\text{hom}} - \dfrac{\sigma_Y}{E} < \bar{\varepsilon} < \bar{\varepsilon}_0^{\text{hom}} \\ 1 + \dfrac{E_t}{\sigma_Y}(\bar{\varepsilon} - \bar{\varepsilon}_0^{\text{hom}}) & \bar{\varepsilon} \geqslant \bar{\varepsilon}_0^{\text{hom}} \end{cases} \tag{4.4a}$$

and

$$f'(\bar{\varepsilon}) = \begin{cases} \dfrac{E}{\sigma_Y} & \bar{\varepsilon}_0^{\text{hom}} - \dfrac{\sigma_Y}{E} < \bar{\varepsilon} < \bar{\varepsilon}_0^{\text{hom}} & E_t > 0 \text{ hardening;} \\ & & E_t = 0 \text{ perfect plasticity;} \\ \dfrac{E_t}{\sigma_Y} & \bar{\varepsilon} \geqslant \bar{\varepsilon}_0^{\text{hom}} & E_t < 0 \text{ softening;} \end{cases} \tag{4.4b}$$

where $E_t$ is the tangent modulus of strain softening/hardening without localization.

Before softening starts, $f(\bar{\varepsilon}) = f(\bar{\varepsilon}^{\text{hom}})$, the first equation of Eq. (3.13) becomes

$$A_{11}\frac{d^2 v^{\text{loc}}}{dx_1^2} + B_{11}\frac{dv^{\text{loc}}}{dx_1} = C_{11} + D_{11}. \tag{4.5}$$

According to the constitutive law given by Eq. (2.10), $A_{11}$ is found to be the material length; $B_{11}$ is a linear function of the hardening-softening constitutive law, $C_{11}$ represents the inertia effect and $D_{11}$ is a constant given below, i.e.,

$$A_{11} = l; \tag{4.6a}$$

$$B_{11} = 2\sqrt{2} f(\bar{\varepsilon}) f'(\bar{\varepsilon}); \tag{4.6b}$$

$$C_{11} = \frac{3\sqrt{2}}{\sigma_Y}\int_{-w/2}^{w/2} \rho \ddot{v}\, d\xi_1; \tag{4.6c}$$

$$D_{11} = 2\sqrt{2} f(\bar{\varepsilon})(E - E_t)\dot{\varepsilon}^{\text{hom}}. \tag{4.6d}$$

By taking the partial derivative with respect to $x_1$, Eq. (4.5) becomes

$$l\frac{\partial^3 v^{\text{loc}}}{\partial x_1^3} + 2\sqrt{2}E_t\frac{\partial^2 v^{\text{loc}}}{\partial x_1^2} = \bar{\rho}\frac{\partial^2 v^{\text{loc}}}{\partial t^2} \quad \text{with } \bar{\rho} = \frac{3\sqrt{2}}{\sigma_Y}\rho. \tag{4.7}$$

Introducing the symmetric Eq. (4.1) and antisymmetric Eq. (4.3) boundary conditions, the solution to Eq. (4.7) can be put in the following form:

$$v^{\text{loc}}(x_1) = \sum_{n=1}^{+\infty} A_n(t) e^{\varphi_n |x_1|} \sin(\sqrt{3}\psi_n x_1) \quad \text{for } -\frac{w}{2} \leqslant x \leqslant \frac{w}{2}. \tag{4.8}$$

Assuming a form of $A_n(t)$ in terms of the eigenvalue $\lambda_n$ as

$$A_n(t) = (b_{n1} e^{-i\sqrt{\lambda_n/\rho}t} + b_{n2} e^{i\sqrt{\lambda_n/\rho}t}), \tag{4.9a}$$

then one can write

$$\varphi_n = -\frac{1}{2}(z_{1n} + z_{2n}) + z_{3n}, \tag{4.9b}$$

$$\psi_n = \frac{1}{2}(z_{1n} - z_{2n}). \tag{4.9c}$$

The coefficients in Eqs. (4.9*a*)–(4.9*c*) are

$$z_{1n} = [h_{1n} + \sqrt{h_{1n}^2 - h_{2n}^2}]^{1/3}, \tag{4.10a}$$

$$z_{2n} = [h_{1n} - \sqrt{h_{1n}^2 - h_{2n}^2}]^{1/3}, \tag{4.10b}$$

$$z_{3n} = -\frac{2\sqrt{2}}{3l}\frac{E_t}{\sigma_Y}, \tag{4.10c}$$

$$h_{1n} = \frac{\lambda_n}{2l} - h_{2n}, \tag{4.10d}$$

$$h_{2n} = \frac{16\sqrt{2}E_t^3}{27\sigma_Y^3 l^3}. \tag{4.10e}$$

The $n$th eigenvalue $\psi_n$ can be determined from the remaining boundary condition and the use of Eq. (4.2) leads to the following equation:

$$\tan\left(\sqrt{3}\psi_n\frac{w}{2}\right) + \frac{d|x_1|}{dx_1}\cdot\frac{\sqrt{3}\psi_n}{\varphi_n} = 0. \tag{4.11}$$

In the above, $\psi_n$ and $\varphi_n$ can be shown to be functions of $\lambda_n$. Hence, Eq. (4.11) is an eigenvalue problem, which involves the unknown $\lambda_n$ and the localization width $w$. In order to solve this eigenvalue problem, we propose a minimum second-order work principle for the determination of $w$. This is addressed in Section 4.2.1. With the assumption that the width $w$ is predetermined, $\lambda_n$ can be shown to be

$$\lambda_n = l\left(\omega_n^3 + \frac{512E_t^6}{729\sigma_Y^6 l^6 \psi_n^3}\right) + \frac{32\sqrt{2}E_t^3}{27\sigma_Y^3 l^2} \tag{4.12}$$

with

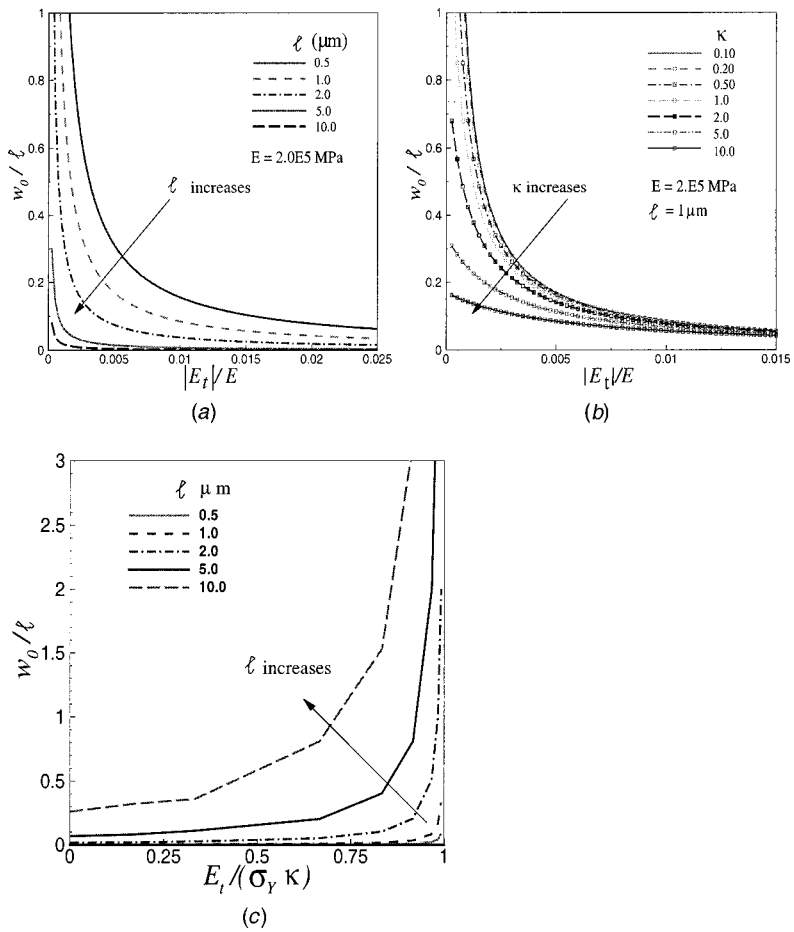



**Fig. 2 (*a*) Initial localization width as a function of material length; (*b*) initial localization width as a function of damage (softening); (*c*) initial localization width as a function of damage (hardening)**

$$\omega_n=\left(\frac{2n\pi}{\sqrt{3}w}+\frac{\sqrt{2}E_t}{3l\sigma_Y}+\sqrt{\frac{4n^2\pi^2}{3w^2}+\frac{4\sqrt{2}n\pi E_t}{3\sqrt{3}lw\sigma_Y}+\frac{10E_t^2}{9l^2\sigma_Y^2}}\right). \tag{4.13}$$

Combining Eqs. (4.12) with (4.13), one can conclude that time scale

$$\lambda_n\propto\left(\frac{w}{l^3}\right) \tag{4.14a}$$

or

$$\lambda_n\propto\frac{1}{l^2}. \tag{4.14b}$$

*4.2.1 Minimum Principle for the Determination of Localization Width.* The localization width $w$ is determined by minimizing the following second-order work term:

$$\dot{\Theta}=\int_{-w/2}^{w/2}\left(\dot{\sigma}^d\dot{\varepsilon}^{\mathrm{loc}}-\frac{1}{2}\rho\dot{v}^{\mathrm{loc}}\cdot\dot{v}^{\mathrm{loc}}\right)dx_1. \tag{4.15}$$

In Eq. (4.15), $\dot{\sigma}^d$ is computed from Eq. (2.17); $\dot{\varepsilon}^{\mathrm{loc}}$ can be obtained by taking the derivative with respect to $x_1$ on Eq. (4.8); and $\dot{v}^{\mathrm{loc}}$ is constructed by taking the derivative with respect to $t$ on Eq. (4.8).

Since $w$ is the only undetermined variable in the kinematic admissable velocity field defined in Eq. (4.15), the variation of $\dot{\Theta}$ becomes

$$\delta\dot{\Theta}=\frac{\partial\dot{\Theta}}{\partial v^{\mathrm{loc}}}\frac{\partial v^{\mathrm{loc}}}{\partial w}\delta w \tag{4.16a}$$

or

$$\frac{\partial\dot{\Theta}}{\partial v^{\mathrm{loc}}}\frac{\partial v^{\mathrm{loc}}}{\partial w}=0. \tag{4.16b}$$

Equation (4.16*b*) gives one extra equation for solving the width $w$, which should be solved with the $n$ sets of equations from Eq. (4.11).

Depicted in Figs. 2(*a*–*c*) are the initial localization width calculated according to the energy minimum principle introduced in the previous section. The following observations can be drawn based on Figs. 2(*a*) to 2(*c*):

- The initial localization width is inversely proportional to the tangent modulus ratio $|E_t|/E$ for softening material.
- The initial localization width increases as the material length scale increases in the case of softening, while the opposite trend is observed in hardening. However, when the initial localization width is normalized with respect to the material length scale, it decreases for softening and increases for hardening when the material length scale is increased.
- The initial localization width dies off quickly with $|E_t|/E$ for the case of softening.

• For hardening, larger value of $|E_t|/E$ gives larger initial width when $E_t/(\sigma_Y\kappa)\leqslant 1$. When $E_t/(\sigma_Y\kappa)>1$, no localization occurs.

• The initial localization width decreases rapidly when the value of the damage parameter $\kappa$ is increased.

• For the case of perfect plasticity, the initial width obtained is infinite when damage is not considered. This validates the argument that damage must be included.

• When sufficient damage is included in the model, finite width of localization is obtained even for the case of perfect plasticity.

## 5 Post-bifurcation Solution Using Deformation Gradient Plasticity

Due to the presence of strain gradient, the homogeneous strain, $\varepsilon^{\text{hom}}$ is further separated into two parts, one is due to the gradient effect ($\varepsilon^g$) and the other is a constant strain ($\varepsilon^h$). That is

$$\varepsilon=\varepsilon^h+\varepsilon^g+\varepsilon^{\text{loc}} \tag{5.1a}$$

and

$$\eta=\eta^h+\eta^g+\eta^{\text{loc}} \tag{5.1b}$$

with

$$\varepsilon^h=\frac{P_{11}}{E} \tag{5.2}$$

where $P_{11}$ is the prescribed traction at both ends of the bar. Another reason for this decomposition is because we need to define $\varepsilon^g$ such that at $x_1=\pm w/2$

$$(\varepsilon^{\text{loc}}+\varepsilon^g)|_{x_1=w/2}=0. \tag{5.3}$$

This equation is compatible to the boundary condition (Eq. (4.2)):

$$\left|\left(\frac{dv^h}{dx_1}+\frac{dv^g}{dx_1}+\frac{dv^{\text{loc}}}{dx_1}\right)\right|_{x_1=w/2}=\left|\frac{dv^h}{dx_1}\right|_{x_1=w/2} \tag{5.4a}$$

or

$$\left|\left(\frac{dv^g}{dx_1}+\frac{dv^{\text{loc}}}{dx_1}\right)\right|_{x_1=w/2}=0. \tag{5.4b}$$

To further simplify the solution procedure, $\varepsilon^g$ is assumed to be a constant with the value of $\bar{\varepsilon}_B^g$ within the localization band (i.e., $-w/2<x_1<w/2$). However, it is a function of $x_1$ for $|x_1|\geqslant w/2$.

Since $\varepsilon^h$ is given, by traction continuity and the strain decomposition (Eqs. (5.1*a*), (5.1*b*)), the solution for $\varepsilon^{\text{loc}}$ can be obtained in two steps:

Step 1. General solution to $\varepsilon^{\text{loc}}$.

It is noted that the strain gradient effects vanish when $x_1$ is sufficiently away from the localization region, i.e., $\varepsilon^g,\eta^g\Rightarrow 0$ when $x_1\Rightarrow\infty$ and $\varepsilon^{\text{loc}}=\eta^{\text{loc}}=0$ at $x_1=L/2$. Therefore, the traction evaluated at $x_1=L/2$ is

$$|t|_{x_1=L/2}=\sigma_Y f(\bar{\varepsilon}^h+\bar{\varepsilon}^g). \tag{5.5}$$

We equate the traction inside the localization region, that is

$$|t|_{x_1<w/2}=\sigma_Y\sqrt{f^2(\bar{\varepsilon}^{\text{loc}}+\bar{\varepsilon}_B^g+\bar{\varepsilon}^h)+l\,\bar{\eta}^{\text{loc}}} \tag{5.6}$$

with the traction evaluated at $x_1=L/2$ to yield

$$\sigma_Y f(\bar{\varepsilon}^h)=\sigma_Y\sqrt{f^2(\bar{\varepsilon}^{\text{loc}}+\bar{\varepsilon}_B^g+\bar{\varepsilon}^h)+l\,\bar{\eta}^{\text{loc}}} \tag{5.7}$$

Given that $\bar{\varepsilon}=\varepsilon$ in one dimension, Eq. (5.7) is further simplified by rearranging terms to give a nonlinear differential equation and

$$\frac{d\varepsilon^{\text{loc}}}{dx_1}=\sqrt{2}\,\frac{f^2(\varepsilon^h+\varepsilon_B^g)-(\bar{\alpha}+\bar{\beta}\varepsilon^{\text{loc}})^2(1-\kappa\varepsilon^{\text{loc}})^2}{l(1-\kappa\varepsilon^{\text{loc}})^2} \tag{5.8}$$

or

$$dx_1=\frac{l(1-\kappa\varepsilon^{\text{loc}})^2}{\sqrt{2}[f^2(\varepsilon^h+\varepsilon_B^g)-(\bar{\alpha}+\bar{\beta}\varepsilon^{\text{loc}})^2(1-\kappa\varepsilon^{\text{loc}})^2]}\,d\varepsilon^{\text{loc}} \tag{5.9}$$

with

$$\bar{\alpha}=(\varepsilon^h-\varepsilon_Y)\frac{E_t}{\sigma_Y}, \tag{5.10a}$$

$$\bar{\beta}=\frac{E_t}{\sigma_Y}, \tag{5.10b}$$

The solution to Eq. (5.9) can be given as $x_1$ versus $\varepsilon^{\text{loc}}$, i.e.,

$$x_1=-\frac{l}{4\sqrt{2}\beta^2}\ln\left[\frac{(\varepsilon^{\text{loc}}+g_3-g_2)^{(1-2b/g_2)+g_4}}{(\varepsilon^{\text{loc}}+g_3+g_2)^{(1-2b/g_2)-g_4}}\cdot\frac{(\varepsilon^{\text{loc}}+g_3-g_1)^{(1+2b/g_1)-g_4}}{(\varepsilon^{\text{loc}}+g_3+g_1)^{(1+2b/g_1)+g_4}}\right]+C_0 \tag{5.11}$$

with

$$g_1=\sqrt{\left(\frac{1}{2\kappa}+\frac{\bar{\alpha}}{2\bar{\beta}}\right)^2+\frac{f(\varepsilon^h+\varepsilon_B^g)}{\bar{\beta}\kappa}} \tag{5.12a}$$

$$g_2=\sqrt{\left(\frac{1}{2\kappa}+\frac{\bar{\alpha}}{2\bar{\beta}}\right)^2-\frac{f(\varepsilon^h+\varepsilon_B^g)}{\bar{\beta}\kappa}} \tag{5.12b}$$

$$g_3=\frac{\bar{\alpha}}{2\bar{\beta}}-\frac{1}{2\kappa} \tag{5.12c}$$

$$g_4=\frac{4b\bar{\beta}\kappa}{\bar{\alpha}\kappa+\bar{\beta}} \tag{5.12d}$$

$$b=\frac{\bar{\beta}\kappa}{f(\varepsilon^h+\varepsilon_B^g)}\left(\frac{1}{2\kappa}+\frac{\bar{\alpha}}{2\bar{\beta}}\right)^2. \tag{5.12e}$$

Step 2. Solution to $\varepsilon^g$ outside the localization region.

The traction evaluated right outside the localization zone, i.e., at $|x_1|>w/2$ is

$$|t|_{|x_1|>w/2}=\sigma_Y\sqrt{f^2(\bar{\varepsilon}^g+\bar{\varepsilon}^h)+l\,\bar{\eta}^g} \tag{5.13}$$

because $\bar{\varepsilon}^{\text{loc}}=\bar{\eta}^{\text{loc}}=0$ for $L/2>|x_1|>w/2$.

By equating the traction given above with traction at $x_1=L/2$, the following traction equilibrium equation is obtained:

$$\sigma_Y\sqrt{f^2(\bar{\varepsilon}^g+\bar{\varepsilon}^h)+l\,\bar{\eta}^g}=\sigma_Y f(\bar{\varepsilon}^h). \tag{5.14}$$

Following the same procedure as in Step 1,

$$\frac{d\varepsilon^g}{dx_1}=\frac{\sqrt{2}}{l}[\alpha^2-(\alpha+\beta\varepsilon^g)^2]$$

or

$$\frac{l}{\sqrt{2}[\alpha^2-(\alpha+\beta\varepsilon^g)^2]}\,d\varepsilon^g=dx_1. \tag{5.15}$$

with constants

$$\alpha=f(\varepsilon^h), \tag{5.16a}$$

$$\beta=-\frac{E}{\sigma_Y}. \tag{5.16b}$$

The solution for Eq. (5.15) is given as

$$\varepsilon^g=-\frac{\alpha}{\beta}+\frac{\alpha}{\beta}\tanh\left(\beta f(\varepsilon^h)\frac{x_1}{l}+C_1\right). \tag{5.17}$$

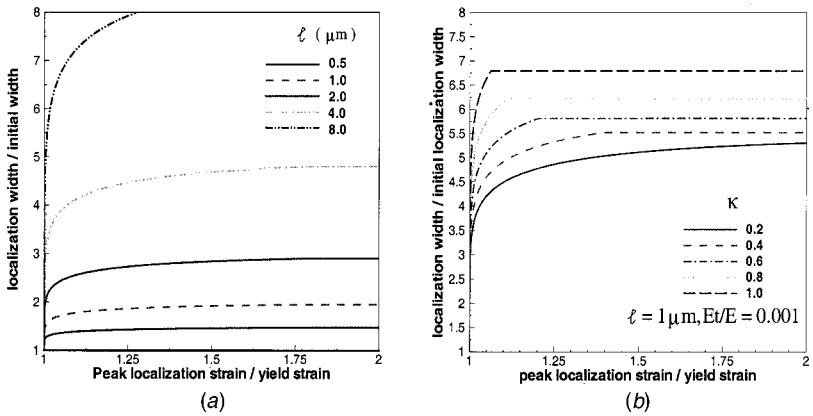



**Fig. 3 (*a*) Evolution of localization width for different material length scale; (*b*) evolution of localization width for different damage constants**

Equation (5.3) gives only one condition for the determination of the two constants ($C_0$ and $C_1$). An additional condition, which is based on the energy transferred from the outside to the localization region, is needed. It is given by

$$\int_{-L/2}^{-w/2}\int_0^{\varepsilon^g}\sigma d\varepsilon^g dx_1=\int_{-w/2}^{0}\int_0^{\varepsilon^{\mathrm{loc}}}\sigma d\varepsilon^{\mathrm{loc}}dx_1. \tag{5.18}$$

At the first step right after the onset of localization, the initial width solution $w_0$ can be used directly in Eq. (5.3) as the known variables to determine $C_0$ and $C_1$. For the remaining steps in the post-bifurcation stage, $w$ is solved based on Eq. (5.3) in which $C_0$ and $C_1$ have been solved from Step 1.

We have established the onset and post-bifurcation solutions based on a strong coupling between the strain gradient theory that has a length scale ($l$) and the continuum scale solution. In a sense, we have applied a multiple scale approach to the problem. The evolution of the localization width in post-bifurcation stage is plotted in Figs. 3(*a*) and 3(*b*). From these figures one can draw the following conclusions:

- For the same peak strain, i.e., $\varepsilon(x_1=0)$, a larger value of material length results in a larger width of the localization zone.
- As the damage constant $\kappa$ becomes larger, the width of the localization zone gets larger.
- The evolution of the width tends to converge to an almost constant value after the peak strain value reached twice the yield strain.

The major features of the analysis can also be summarized as follows:

1 Material length scale is embedded within a multiscale framework.
2 In the proposed model, finite width of localization zone can be predicted and we have established the relation with the material length scale.

## 6 Localization-Induced Cohesive Model

To consider the effect of stress triaxiality, we assume that in the radial direction, a uniformly distributed stress proportional to the axial stress is applied on the rod; i.e.,

$$\sigma_{22}=\sigma_{33}=k_{tr}\sigma_{11}\quad k_{tr}<1 \tag{6.1}$$

where $k_{tr}$ is an arbitrary constant representing the effect of stress triaxiality. From Eq. (6.1) it can be further given that

$$\sigma_{11}-\sigma_{22}=(1-k_{tr})\sigma_{11}=\sigma_{11}^d-\sigma_{22}^d. \tag{6.2}$$

The uniaxial stress can then be expressed as

$$\sigma_{11}=\frac{1}{1-k_{tr}}[\sigma_{11}^d-\sigma_{22}^d]. \tag{6.3}$$

In Eq. (6.3) $\sigma^d$ is computed by integrating the rate from of the constitutive equation (Eq. (2.17)). Equation (6.3) represents the relation between the axial stress and the transverse stress. Based on the solution to $v^{\mathrm{loc}}$, the traction $\Sigma$ over the localization (cohesive) region can be calculated as follows:

$$\Sigma=A\sigma_{11} \tag{6.4}$$

in which $A$ is the cross-sectional area of the bar and chosen to be unity in the computation.

The localization-induced separation $\Delta$ is defined to be the separation between the two ends of the localization width:

$$\Delta=u^{\mathrm{loc}}\left(\frac{w}{2}\right)-u^{\mathrm{loc}}\left(-\frac{w}{2}\right) \tag{6.5}$$

with

$$u^{\mathrm{loc}}(x_1)=\int_{-w/2}^{x_1}\varepsilon^{\mathrm{loc}}(y)dy \tag{6.6a}$$

or

$$\begin{aligned}u^{\mathrm{loc}}(x_1)=&-\frac{l}{4\bar{\beta}^2}\Bigg\{(1-2b)\ln((\varepsilon^{\mathrm{loc}}+g_3)^2-g_2^2)+(1+2b)\\&\times\left[2(\varepsilon^{\mathrm{loc}}+g_3)+g_1\ln\frac{(\varepsilon^{\mathrm{loc}}+g_3-g_1)}{(\varepsilon^{\mathrm{loc}}+g_3+g_1)}\right]\\&+g_4\left[g_1\ln\left(\frac{\varepsilon^{\mathrm{loc}}+g_3-g_1}{\varepsilon^{\mathrm{loc}}+g_3+g_1}\right)-g_2\ln\left(\frac{\varepsilon^{\mathrm{loc}}+g_3-g_2}{\varepsilon^{\mathrm{loc}}+g_3+g_2}\right)\right]\Bigg\}.\end{aligned} \tag{6.6b}$$

The computed traction-separation law are plotted in Figs. 4(*a*) to 4(*d*) with varying material softening properties, material length scales, damage evolution parameter, and the level of hydrostatic stress. The following observations are emphasized:

- All these figures present an S-type traction-separation curve.
- Stronger softening leads to smaller separations.
- Larger material intrinsic length scales result in larger separations, however, the traction-normalized opening displacement curves (with respect to the material length scale) are identical.
- Larger damage constant produces less separation.
- For the same separation, higher traction is produced due to a larger triaxiality (hydrostatic pressure) constant.

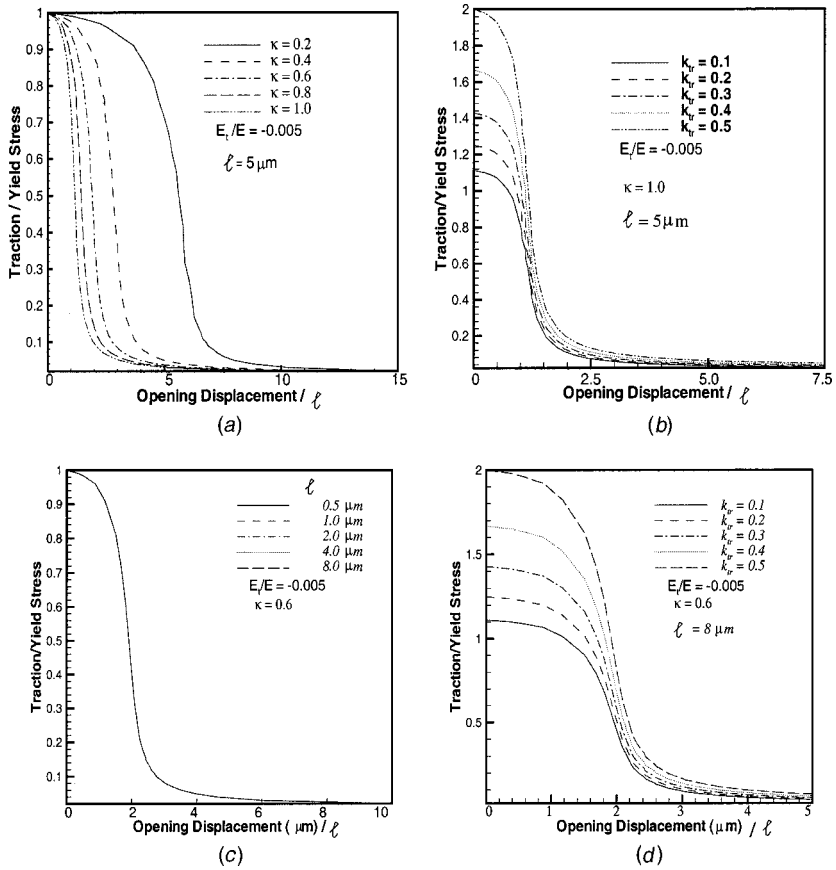



**Fig. 4 (*a*) Mode I traction-separation law with different damage constants; (*b*) corresponding traction-separation law under varying hydrostatic pressure; (*c*) mode I traction-separation law with different material length; (*d*) corresponding traction-separation law under varying hydrostatic pressure**

In fact, during the separation process of a material there are two opposite tendencies: strain softening accelerates the localization and decreases the stress; and the strain gradient which elevates the yielding strength and suppresses localization strain. In the terminology of strain gradient theory, more geometrically necessary dislocation is required for yielding and it homogenizes the localization strain. Apparently, the latter is significant right after the onset of localization, and a slight drop of traction causes a relatively large amount of separations at this stage. This also provides an explanation for the rapid increasing of localization width at the beginning of the post-bifurcation, as shown in Figs. 4(*a*) and 4(*c*).

It is noted that empirical formula based on curve fitting for the relation between normalized traction $\Sigma/\sigma_{Y0}$ and normalized opening displacement $\bar{d}$ (with respect to material length scale $l$) can be given for different cases. For instance, when $\kappa=0.6$, $E_t/E=-0.005$, we have

$$\frac{\Sigma}{\sigma_{Y0}}=\frac{(-0.5398Y^2+1.5867Y-0.0466)}{(1-k_{tr})} \tag{6.7}$$

$$\text{with}\quad Y=\exp[-(\bar{d})^{2.2}\times 10^6]\quad k_{tr}<1. \tag{6.8}$$

The major difference between the proposed multiscale cohesive model and the phenomenological cohesive model can be summarized as follows:

1 The traction-separation law is derived from a physically based strain gradient theory while the classical approach employs a phenomenological traction-separation law.
2 Material length scale is embedded in the proposed model and has a significant effect at small scale. The classical cohesive model has no length scale in it.
3 Since the model is mechanism-based, fewer parameters need to be determined. While in the classical approach a curve-fitting procedure has to be taken.
4 The multiscale framework of the proposed model makes it applicable for multiscale analysis.

## 7 Application to Multiscale Analysis

In this paper, the multiple scale enrichment procedure proposed in ([12,13]) has been applied to embed the Traction-Separation law, derived in the previous section, into meshfree or finite element code. This method is briefly explained as follows:

In general, for an unknown field variable such as velocity $\mathbf{v}$, a computed $\mathbf{v}$ can be denoted as $P\mathbf{v}$, the projection of $\mathbf{v}$ obtained by a projection operator $P$, where $P$ represents a numerical procedure. Assume there is a two-scale decompositions of $\mathbf{v}$:

$$\mathbf{v}=\mathbf{v}^{\text{hom}}+\mathbf{v}^{\text{loc}}. \tag{7.1}$$

The multiple scale expansion of $\mathbf{v}$ using the operator $P$ is

$$\mathbf{v}=\overline{\mathbf{v}}_{\text{low}}+\overline{\mathbf{v}}_{\text{high}}+\overline{\mathbf{v}}_{\text{bridging}} \tag{7.2}$$

where

$$\overline{\mathbf{v}}_{\text{low}}=P\mathbf{v}_{\text{low}},\quad \overline{\mathbf{v}}_{\text{high}}=\overline{\mathbf{v}}_{\text{high}},\quad \overline{\mathbf{v}}_{\text{bridging}}=-P\mathbf{v}_{\text{high}}. \tag{7.3}$$

A one-dimensional bar with length of 0.1 mm or 1 mm is investigated. An imperfection area is prescribed in the middle sec-

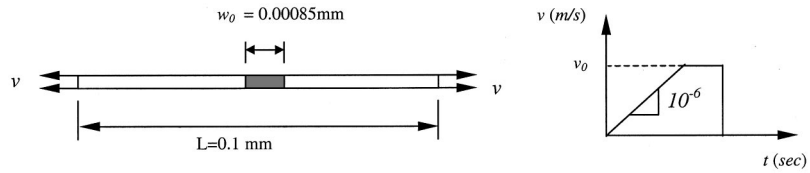



**Fig. 5 A rod under tension**

tion with length of 0.003 mm. The material constants chosen are: yield stress $\sigma_{y0}=100$ MPa, Young's modulus $E=2\times10^5$ MPa, softening modulus $E_t=0.05E$, and damage constant $\kappa=0$. An imperfection of the yield stress $(0.99\,\sigma_Y)$ is employed in the initial localization region as shown in Fig. 5. Before $\varepsilon^h$ reaches $0.98\varepsilon_y$, the magnitude of velocity $v$ is controlled as $10^{-6}t$. After that it is kept as a constant until localization occurs. Right at the onset of the localization, the magnitude of the velocity at the boundary is dropped to zero.

In the present work, the RKPM (reproduced kernel particle method) ([14–16]), a recently developed meshfree method ([17]), has been employed. Four examples ranging from fine to coarse distributions of particles have been computed and shown on Fig. 6. The particle numbers that are used in the above four cases are given in Table 1. Two stress points are assigned in between two neighbor particles.

In the problem, $\mathbf{u}^{\mathrm{loc}}$ is taken from Eqs. (6.6) as the exact solution, the bridging term serves to couple the analytical solution with the RKPM. It is expected that the bridging term vanishes when a coarse discretization is prescribed, while it dominates the solution in the case of fine discretization. This approach allows one to use a numerical procedure with a very coarse mesh to compute localization solution with considerable accuracy by embedding a localized exact solution.

In Fig. 6, the domain of interest is one length scale with 420, 22, 10, 4 stress points for cases *a*, *b*, *c*, *d*, respectively. Hence, the particle numbers in this region are correspondingly 211, 12, 6, and 3 for the four cases. As can be seen, the simulated strain distributions from fine to coarse discretization are given from Fig. 6(*a*) to (*d*). In Fig. 6(*a*) the low-scale strain represents the localization phenomenon with sufficient accuracy. Since the high-scale solution becomes redundant, the bridging scale solution is significant. In this case the high-scale and bridging-scale strains will essentially cancel each other and the combination of the three scales leads to the solution with higher accuracy. As can be seen in the transition case (Figs. 6(*b*) and 6(*c*)), the effect of the bridging scale solution decreases as more coarse discretization is prescribed. Figure 6(*d*) shows that both the low scale solution and bridging scale solution can not catch the localization phenomenon due to the coarse mesh. The phenomenon is captured by the high-scale solution. These four examples have proven that the compu-

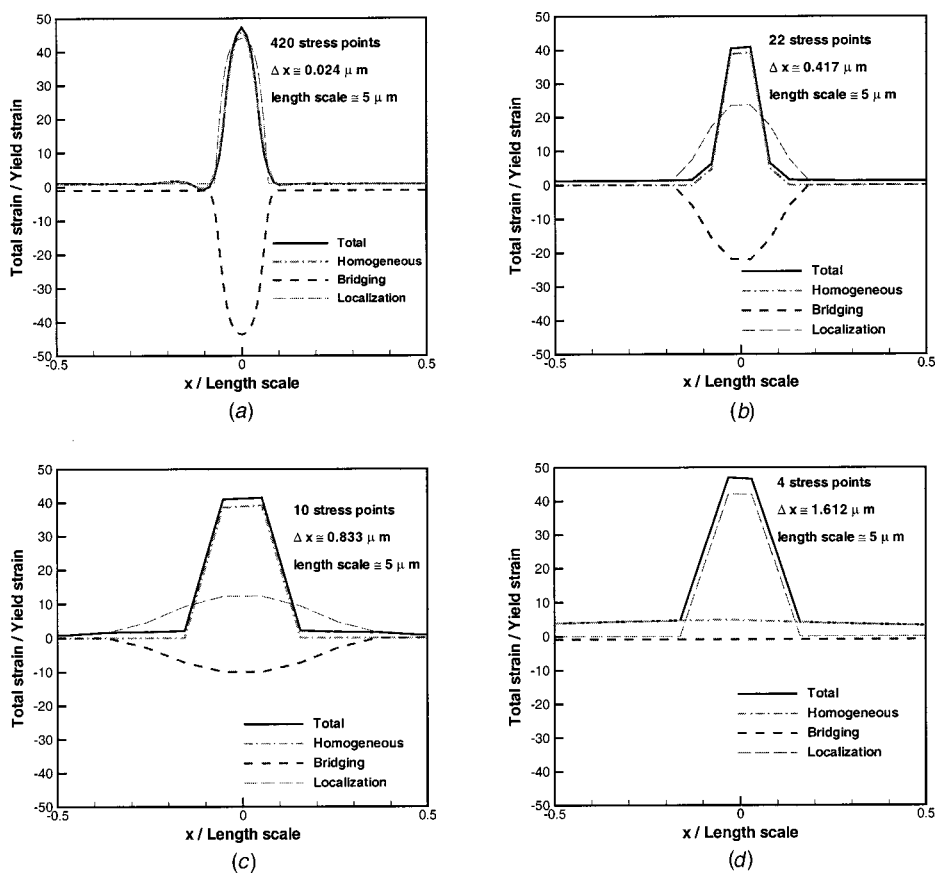



**Fig. 6 Comparison of solutions among different discretizations**

**Table 1 Discretization for different cases**

| | Total Number of Particles | Number of Particles in One Length Scale | Length Scale | $L$ | $\Delta x$ |
|---|---|---|---|---|---|
| Case a | 4221 | 211 | 5 $\mu$m | 100 $\mu$m | 0.024 $\mu$m |
| Case b | 241 | 12 | 5 $\mu$m | 100 $\mu$m | 0.417 $\mu$m |
| Case c | 121 | 6 | 5 $\mu$m | 100 $\mu$m | 0.833 $\mu$m |
| Case d | 621 | 3 | 5 $\mu$m | 1000 $\mu$m | 1.612 $\mu$m |

tation procedure can reproduce the localization-induced cohesive model by using a coarse distribution of particles.

## 8 Conclusion

A methodology is proposed to establish the boundary value problem at the onset of localization and post-bifurcation for materials obeying strain gradient theory in conjunction with damage. To compare with conventional plasticity, significant difference can be seen for both onset and post-bifurcation localization deformation. Based on the analytical solutions obtained from one-dimensional tension, the ''traction-separation law'' for cohesive model is derived. Some important observation is given as follows:

- The evolution of the localized band width is a strong function of the material length scale. Larger material length scale contributes to larger localization zone and separation.
- Damage must be considered and faster damage evolution will produce a smaller localization band.
- Initial localization width is inversely proportional to the magnitude of softening modulus, while the trend is opposite for the hardening case.

The significance of the proposed methodology and results can be summarized as follows:

1 From the one-dimensional solution the localization zone size/shear band width has been obtained explicitly. It has the same order as material intrinsic length scale and is dependent upon common mechanical properties.

2 The derived ''traction-separation law'' can be considered as a ''cohesive law'' since it determines the energy release rate of a given material.

3 The corresponding numerical procedure is proposed based on multiple scale analysis. In conjunction with the ''traction-separation law'' and the solution of shear band width, the proposed numerical procedure demonstrates the ability to capture localized deformation with coarse discretization.

4 Potentially, the combination of proposed theoretical framework and numerical procedure provides a multiple scale methodology to determine the material fracture toughness from the constitutive law with multiphysics, or inversely, to calibrate constitutive law based on experimental results.

5 The shear band solution and cohesive model due to shear can similarly derived to yield the same expressions given in this paper, however, with a different set of constants.

## Acknowledgment

The support of this research by Army Research Office (ARO) and National Science Foundation (NSF) to Northwestern University is gratefully acknowledged. Sponsored in part by the Army High Performance Computing Research Center (AHPCRC) under the auspices of the Department of the Army Research Laboratory. The content does not necessarily reflect the position or the policy of the government, and no official endorsement should be inferred.

The authors would like to thank Dr. Patrick Klein of Sandia National Lab for the helpful discussions and comments.

## References

[1] Needleman, A., 1987, ''A Continuum Model for Void Nucleation by Inclusion Debonding,'' ASME J. Appl. Mech., **54**, pp. 525–531.
[2] Tvergaard, V., and Hutchinson, J. W., 1992, ''The Relation Between Crack Growth Resistance and Fracture Process Parameters in Elastic-Plastic Solids,'' J. Mech. Phys. Solids, **40**, pp. 1377–1397.
[3] Wright, T. W., and Walter, J. W., 1987, ''On Stress Collapse in Adiabadic Shear Bands,'' J. Mech. Phys. Solids, **35**, pp. 701–720.
[4] DiLellio, J. A., and Olmstead, W. E., 1997, ''Temporal Evolution of Shear Band Thickness,'' J. Mech. Phys. Solids, **45**, pp. 345–359.
[5] Fleck, N. A., Muller, G. M., Ashby, M. F., and Hutchinson, J. W., 1994, ''Strain Gradient Plasticity: Theory and Experiment,'' Acta Metall. Mater., **42**, pp. 475–487.
[6] Gao, H., Huang, Y., Nix, W. D., and Hutchinson, J. W., 1999, ''Mechanism-Based Strain Gradient Plasticity—I. Theory,'' J. Mech. Phys. Solids, **47**, pp. 1239–1263.
[7] Nix, W. D., and Gao, H., 1998, ''Indentation Size Effects in Crystalline Materials: A Law for Strain Gradient Plasticity,'' J. Mech. Phys. Solids, **46**, pp. 411–425.
[8] Tvergaard, V., 1982, ''On Localization in Ductile Materials Containing Spherical Voids,'' Int. J. Fract., **18**, pp. 237–252.
[9] Rice, J. R., 1976, ''The Localization of Plastic Deformation,'' *Theoretical and Applied Mechanics*, pp. 207–220.
[10] Hill, R., 1962, ''Acceleration Waves in Solid,'' J. Mech. Phys. Solids, **10**, pp. 1–16.
[11] Rudnicki, J. W., and Rice, J. R., 1975, ''Conditions for the Localization of Deformation in Pressure-Sensitive Dialatant Materials,'' J. Mech. Phys. Solids, **23**, pp. 371–394.
[12] Liu, W. K., Uras, R. A., and Chen, Y., 1997, ''Enrichment of the Finite Element Method With the Reproducing Kernel Particle Method,'' ASME J. Appl. Mech., **64**, pp. 861–870.
[13] Liu, W. K., Zhang, Y. F., and Ramirez, M. R., 1991, ''Multiple Scale Finite Element Methods,'' Int. J. Numer. Methods Eng., **32**, pp. 960–990.
[14] Liu, W. K., Jun, S., and Zhang, Y. F., 1995, ''Reproducing Kernel Particle Method,'' Int. J. Numer. Methods Fluids, **20**, pp. 1080–1106.
[15] Liu, W. K., Chen, Y., Uras, R. A., and Chang, C. T., 1996, ''Generalized Multiple Scale Reproducing Kernel Particle Methods,'' Comput. Methods Appl. Mech. Eng., **139**, pp. 91–158.
[16] Liu, W. K., Chen, Y., Jun, S., Chen, J. S., Belytschko, T., Pan, C., Uras, R. A., and Chang, C. T., 1996, ''Overview and Applications of the Reproducing Kernel Particle Methods,'' Archives of Computational Methods in Engineering, State of Art Reviews, **3**, pp. 3–80.
[17] Liu, W. K., Chen, Y., Chang, C. T., and Belytschko, T., 1996, ''Advances in Multiple Scale Kernel Particle Methods,'' Computational Mechanics, **18**, No. 2, pp. 73–111.

# Bending and Extension of Thin-Walled Composite Beams of Open Cross-Sectional Geometry

**O. Rand**
Associate Professor,
Faculty of Aerospace Engineering,
Technion–Israel Institute of Technology,
Haifa 32000, Israel

*A theoretical model for bending and extension of composite beams of open thin-walled geometry is devised. The model is based on a detailed description of the out-of-plane warping and may handle generic open cross-sectional geometries and arbitrary layup configurations. In view of the potential of composite beams to produce useful structural couplings, the analysis is focused on the ability to predict the coupling mechanisms, and draw lines of similarities and differences between such beams and similar thin-walled beams of closed cross sections.* [S0021-8936(00)01204-6]

## Introduction

The structural behavior of isotropic (and mostly metal) thin-walled structures has attracted considerable attention during the last decades. Thin-walled structures play a major role in modern engineering and in particular in aerospace engineering applications (see, e.g., [1,2]). The analysis of thin-walled beams (that originally emerged from analyses of long tubes) is an important subset of the above huge range of thin-walled structures (see, e.g., [3,4]), and beams of open cross-sectional geometry constitute an important ingredient of this subset.

In recent years, the increasing interest in composite beams seems to become a clear trend. Currently, composite beams are of special interest in many advanced engineering applications, among which aircraft wings and helicopter blades are typical examples (see, e.g., [5–10]).

The primary attraction of thin-walled beams is due to the intrinsic high ''strength-to-weight'' ratio that such beams may offer. Similarly, the initial attraction of composite beams was originated from the additional improvement of the above specific strength and the improved fatigue characteristics that composites offer; however, the potential of composite materials to supply *structural (elastic) coupling* which is a unique feature of directional composites is much more promising. Such couplings emerge from the stress/strain coupling characteristics at the material level, and are relatively simple to implement in composite beam structures. In general, the structural couplings in beams may be classified into two categories: The extension/twist coupling and the bending/twist coupling. While in the former case one is interested in the coupling between the axial strain and the elastic twist, the latter coupling deals with the coupling between the lateral bending and the elastic twist. In general, extension/twist coupling is obtained in beams with antisymmetric layup, while bending/twist coupling is obtained in beams with symmetric layup configuration. A simple symmetric and antisymmetric layup for open thin-walled cross section is presented in Fig. 1. Similar phenomena are observed in beams of solid cross section where a simple example is a uniform beam made of two laminae. In this case, opposite layup angles create extension/twist coupling while identical layup angles create bending/twist coupling. In many engineering applications and mainly in aeronautical applications, the twist angle plays a major role (e.g., it directly influences the angle of attach in aircraft wings and helicopter blades), and therefore, the above elastic couplings are of primary importance. If modeled and tailored properly, such couplings may significantly contribute to the structural response of the beam and enable adequate tailoring of the beam structure to comply with specific requirements (see, e.g., [7,11]).

The analysis of thin-walled composite beams is fundamentally different from a similar analysis of isotropic structures. This is due to the fact that realistic composite thin-walled beams are typically produced in a way that results in material laminae which are parallel to the local wall direction (e.g., the ''filament-winding'' process). Such a change in the material directions dramatically influences the level of complexity of the required analysis, and forces a development of different analytical methodologies.

Compared with closed thin-walled geometries, it is well known that beams of open cross sections produce out-of-plane warping deformation which is very large and of different nature. Thus, since warping is known to play a major role in the determination of the coupling effects (see, e.g., [9]), it is expected that composite beams of open cross section will provide substantially different coupling mechanisms and magnitudes compared with similar beams of closed cross-section.

Due to the above described characteristics of thin-walled composite beams, the literature contains a vast theoretical and experimental research effort in this area most of which is focused on closed cross sections. It is beyond the scope of the present paper to present a thorough review of it. Yet, the reader is referred to ([12–17]), while only few of the relevant models are discussed here as representative analyses.

A Vlasov-type theory for fiber-reinforced beams with thin-walled open cross sections has been presented in ([12]). The kinematics is based on three spanwise functions (two transverse displacements and a rotation angle of each cross section), while the out-of-plane warping is linearly proportional to the twist. The model also includes a nonlinear set of equations which is used for bifurcation and stability analyses. References ([3,14]) focus on composite beams of ''*I*'' cross sections with elastic couplings. The model consists of torsional-related out-of-plane warping and transverse shear. Warping distribution is based on a Vlasov-type warping function which is a function of the distribution of the normal distance to the rotation center. The equations of equilibrium are derived by the principle of virtual work. A simplified theory for composite closed and open thin-walled beams is presented in ([14]) where the out-of-plane warping is expressed using a (geometry-dependent) predetermined shape function, the magnitude of which is the twist. The resulting model is subsequently based on four equations for each cross section.

The analysis presented in this paper offers, in addition to the general formulation, some closed-form solutions for simple cases

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Nov. 11, 1999; final revision, Jan. 14, 2000. Associate Technical Editor: M.-J. Pindera. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS.

 **Copyright © 2000 by ASME** 

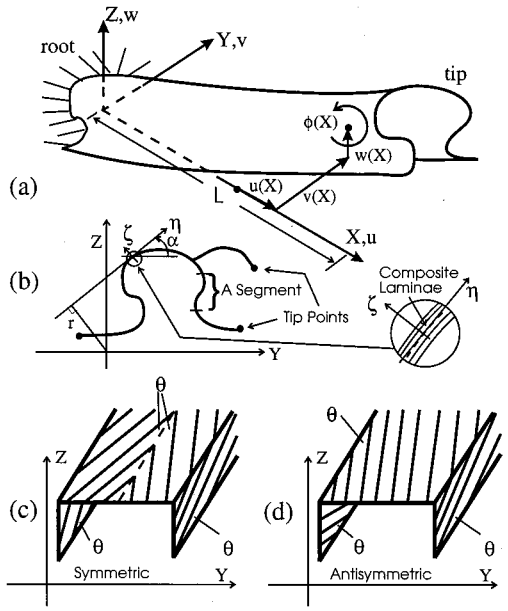



**Fig. 1 A scheme of a thin-walled composite beam of open cross-sectional geometry. (*a*) The cross-sectional deformation components (the out-of-plane warping, $\psi(x,\eta)$, is not shown); (*b*) a generic open cross-sectional geometry; (*c*), (*d*) symmetric and antisymmetric layups (with respect to the $X-Z$ plane).**

that clarify the structural mechanisms in thin-walled open composite beams. The distinguishing feature of the present analysis is the lack of preliminary assumptions regarding the shape of the out-of-plane warping distribution and its origin (i.e., torsional-related or shear-related out-of-plane deformation). Therefore, in view of the important role of the out-of-plane warping in the determination of the beam structural characteristics, the present theoretical model is capable of providing a proper insight into the structural elastic coupling effects and their influence on the overall structural response. It should be emphasized that the present model deals with the effect of extension and bending only. Hence, for a complete picture, a solution of the beam response to a torsional moment should be superimposed on the present results.

## Kinematics

Figure 1(*a*) presents a uniform thin-walled beam the axis of which is straight and parallel (before deformation) to the $X-Y-Z$ system of coordinates, while the location of the $Y-Z$ system origin over the cross section is arbitrary. The beam is defined in the domain between its ''root'' end area and its ''tip'' end area and its length is denoted $L$. The beam is acted upon a distribution of forces per unit length $f_x(X)$, $f_y(X)$, and $f_z(X)$, and a distribution of moments per unit length, $q_x(X)$, $q_y(x)$, and $q_z(X)$. In addition, concentrated forces $F_x^i$ $(i=1,...,N_x^F)$, $F_y^i$ $(i=1,...,N_y^F)$ and $F_z^i$ $(i=1,...,N_z^F)$ are applied at the spanwise locations $X_x^{Fi}, X_y^{Fi}, X_z^{Fi}$, respectively. Similarly, concentrated moments $Q_x^i$, $(i=1,...,N_x^Q)$, $Q_y^i$ $(i=1,...,N_y^Q)$ and $Q_z^i$ $(i=1,...,N_z^Q)$ are applied at the spanwise locations $X_x^{Qi}, X_y^{Qi}, X_z^{Qi}$, respectively. In the above definitions, the subscripts $(\ )_x$, $(\ )_y$, and $(\ )_z$ indicate quantities related to the $X$, $Y$, and $Z$ directions, respectively. Note that present modeling deals with slender beams, and therefore the exact distribution of the loads at each cross sections is immaterial as only their equivalent components along the beam $X$-axis (i.e., the $Y-Z$ origin at each cross section) are of interest. As already discussed, the above $q_x$ and $Q_x^i$ torsional moments are assumed to vanish throughout the present analysis as they should be taken into account within a separate analysis.

Figure 1(*b*) presents a scheme of an ''open'' thin-walled cross section. The notion ''open'' stands for the case of a *thin-walled simply connected domain* (in thin-walled terminology it is a cross-sectional geometry where no ''closed loops'' exist). Such a cross section may include additional branches (e.g., ''I'' geometry). As shown, a local system of coordinates $\xi-\eta-\zeta$ is located at each point along the cross-sectional walls so that $\xi$ is parallel to the $X$-axis, $\eta$ coincides with the local tangent to the wall, and $\zeta$ is perpendicular to $\xi$ and $\eta$. The angle between the $Y$ and the $\eta$-axes is denoted $\alpha$. The $\zeta-\eta$ origin is located at the wall middle plane. The normal (perpendicular) distance from the $Y-Z$ origin to the tangent line at the point under discussion is denoted $r$ and is given by $r=Z\cos\alpha-Y\sin\alpha$.

The bending kinematics of the beam is based on the fact that the beam is a slender structure. The deformation is expressed by the cross-sectional displacements $u(X)$, $v(X)$, and $w(X)$ in the $X$, $Y$, and $Z$ directions, respectively, and a twist angle, $\phi(X)$ about the $X$-axis—see Fig. 1(*a*). These components of the deformation are functions of $X$ only, and therefore, they represent ''rigid'' deformation of each cross section that contains no warping. Subsequently, an additional out-of-plane warping function is superimposed (in the $X$ direction) upon the above mentioned displacements. This warping function is denoted $\psi$, and is assumed to be of zero average value over the cross-sectional area (i.e., $\int\int_A \psi dA=0$). $\psi$ is a function of $X$, $Y$, and $Z$. The in-plane warping components (i.e., the distortion of the cross-sectional shape) are neglected.

As shown in ([9]), the above kinematics yields the following axial strain, $\varepsilon_{\xi\xi}$ (perpendicular to the $\zeta-\eta$ plane), and shear strain, $\gamma_{\xi\eta}$ (in the $\xi-\eta$ plane):

$$\varepsilon_{\xi\xi}=u_{,x}-Yv_{,xx}-Zw_{,xx}+\psi_{,x} \tag{1a}$$

$$\gamma_{\xi\eta}=-r\phi_{,x}+\psi_{,\eta} \tag{1b}$$

The role of the shear strain, $\gamma_{\xi\zeta}$ (perpendicular to the $\xi-\eta$ plane) is limited and should be discussed separately in the context of the beam response to a torsional moment.

The cross section is assumed to be constructed out of orthotropic laminae which are parallel to the $\xi-\eta$ plane. In the general case, the principal axes of this laminae do not coincide with the $\xi$ (or $X$) direction and create an angle $\theta$ with this direction—see Figs. (1*a*) and (1*b*). Hence, an ''unbalanced'' layup is obtained, and the corresponding constitutive relations may be put as

$$\{\sigma\}=[C']\{\varepsilon\} \tag{2}$$

where $[C']$ is a 6×6 stiffness matrix, the elastic moduli of which are functions of the material properties and the ply angle relative to the $\xi$-axis, and $\{\sigma\}$ and $\{\varepsilon\}$ are the stress and strain vectors. As indicated above, except for the out-of-plane warping, the cross section is assumed to remain rigid in its own plane. However, as discussed in ([9]), more realistic results are obtained by assuming that the normal stresses $\sigma_{yy}, \sigma_{zz}$ and the shear stress $\tau_{\eta\zeta}$ are negligible. This assumption is sometime referred to as a first-order approximation to the in-plane warping effects (see, e.g., discussion in [18]). Thus, such a reduction of the constitutive relations of Eqs. (2) to the present case yields

$$\begin{Bmatrix}\sigma_{\xi\xi}\\ \tau_{\xi\zeta}\\ \tau_{\xi\eta}\end{Bmatrix}=\begin{bmatrix}C_{11} & 0 & C_{16}\\ 0 & C_{55} & 0\\ C_{16} & 0 & C_{66}\end{bmatrix}\begin{Bmatrix}\varepsilon_{\xi\xi}\\ \gamma_{\xi\zeta}\\ \gamma_{\xi\eta}\end{Bmatrix}. \tag{3}$$

## The Governing Equations

For the sake of clarity and computational efficiency, the following discussion is broken down into sequential steps although a global (''one-step'') presentation of the theory is also possible.

First, the analysis will be derived for a general location of the origin of the $Y-Z$ system (see Fig. 1(*b*)). In addition, the twist-induced shear deformation is assumed to be embedded in a temporary warping function $\tilde{\psi}$ which is defined as $\tilde{\psi}_{,\eta}=\psi_{,\eta}-r\phi_{,x}$, and therefore, $\phi=0$ is temporarily assumed (see Eq. (1*b*)). Later on, the evaluation of $\phi_{,x}$ for this case will be discussed.

To avoid dealing with the exact mode of application of the distributed loads, the present analysis assumes that all distributed loads are applied via a generic distribution of the body force components, $B^{\xi}$ and $B^{\eta}$ in the $X(=\xi)$ and $\eta$ directions, respectively. $B^{\xi}$ is used to introduce the components $f_x$, $q_y$, and $q_z$, while $B^{\eta}$ is used to introduce the components $f_y$ and $f_z$. Subsequently, the following type of body force distribution in the $X$ direction is adopted:

$$B^{\xi}(X,Y,Z)=B_0(X)+YB_y(X)+ZB_z(X) \tag{4}$$

The above $B_0$, $B_y$, and $B_z$ coefficients are obtained from their integration over the beam's cross section:

$$\begin{bmatrix} A & A_y & A_z \\ A_y & I_{yy} & I_{yz} \\ A_z & I_{yz} & I_{zz} \end{bmatrix} \begin{Bmatrix} B_0 \\ B_y \\ B_z \end{Bmatrix} = \begin{Bmatrix} f_x \\ -q_z \\ q_y \end{Bmatrix} \tag{5}$$

while

$$[A,A_y,A_z,I_{yy},I_{yz},I_{zz}]=\int\int_A [1,Y,Z,Y^2,YZ,Z^2]dA. \tag{6}$$

As far as the transverse distributed loads $f_y$ and $f_z$ are concerned, the exact distribution of the $B^{\eta}$ component of the body force is immaterial as long as $\int\int_A B^{\eta}\cos\alpha dA=f_y$ and $\int\int_A B^{\eta}\sin\alpha dA=f_z$ since it will only influence $q_x$.

The present beam analysis is based on an integral form of the equilibrium equations in the $\xi$ and $\eta$ directions and on a differential form of the equilibrium equation in the $\xi$ direction. The above equilibrium equations are

$$\sigma_{\xi\xi,\xi}+\tau_{\xi\eta,\eta}+\underline{\tau_{\xi\zeta,\zeta}}+B^{\xi}=0 \tag{7a}$$

$$\underline{\sigma_{\eta\eta,\eta}}+\tau_{\xi\eta,\xi}+\underline{\tau_{\eta\zeta,\zeta}}+B^{\eta}=0 \tag{7b}$$

where according to the present assumptions, the underlined terms are ignored. Multiplying Eq. (7*b*) by $\cos\alpha$ and $\sin\alpha$, separately, and integrating the resulting equations and Eq. (7*a*) once over the cross-sectional area and then with respect to $X$, yields the following three integral equations:

$$[R_x,R_y,R_z]=\int\int_A [\sigma_{\xi\xi},\tau_{\xi\eta}\cos\alpha,\tau_{\xi\eta}\sin\alpha]dA \tag{8}$$

where $R_x$, $R_y$, and $R_z$ are the resultant force components that act at each cross section. These loads are determined by integrating the distributed forces ($f_x(X)$, $f_y(X)$, and $f_z(X)$) and the concentrated force ($F_x^i$, $F_y^i$, and $F_z^i$) along the beam. For the case of a cantilever beam, these integrations may be written as

$$R_{\alpha}(X)=\int_X^L f_{\alpha}(X')dX'+\sum_{i=1}^{N_{\alpha}^F} F_{\alpha}^i l^i(X) \tag{9}$$

where $\alpha=x,y,z$ and

$$l^i(X)=\begin{cases} 0 & \text{for } X_{\alpha}^{F^i}\leqslant X \\ 1 & \text{for } X_{\alpha}^{Fi}>X \end{cases}. \tag{10}$$

Other combinations of boundary conditions may be treated in a similar way.

In addition to Eq. (8) that constitute three integral equations for each cross section, and since the axial out-of-plane warping is a local quantity (i.e., a function of both $\xi$ and $\eta$), the local differential equation in the $\xi$ direction given by Eqs. (7*a*) should be explicitly imposed as well, while the distribution of the body load $B^{\xi}(X,Y,Z)$ is given by Eqs. (4), (5).

The discussion of the boundary conditions for the above differential equation will be focused on the most common case where the cross section consists of segments of constant elastic moduli, and a segment is defined by any finite length over the cross sectional walls (see Fig. 1(*b*)). Hence, at the cross-sectional free tips, the boundary condition requires $\tau_{\xi\eta}=0$. This condition implies that

$$\frac{d\tilde{\psi}}{d\eta}=-\frac{C_{16}}{C_{66}}\varepsilon_{\xi\xi}, \tag{11}$$

at these locations. Note that the above condition together with the $\tau_{\xi\zeta}=0$ assumption yield a traction-free outer surface over the entire beam since as previously mentioned, the exact application of the distributed loads is not taken into account by the present beam analysis, and all distributed loads are assumed to be applied through the above-mentioned body force distributions. In addition, between two segments (of different elastic moduli), the shear stress should remain continuous. Therefore, the shear strain ($\tilde{\psi}_{,\eta}$) exhibits a step change which is obtained from the following equation:

$$(C_{16}\varepsilon_{\xi\xi}+C_{66}\tilde{\psi}_{,\eta})|_{\eta_B^-}=(C_{16}\varepsilon_{\xi\xi}+C_{66}\tilde{\psi}_{,\eta})|_{\eta_B^+} \tag{12}$$

where $\eta=\eta_B$ represents the interface point between segments and $\eta_B^-$ and $\eta_B^+$ are two locations just below and above this location, respectively. A connection point of three or more segments yields a similar condition.

It should also be emphasized that the coexistence of an integral and a differential equations in the $X$ direction (see Eqs. (7*a*), (8)) does not create any redundancy. This is due to the fact that the unknown deformation component $u$ and $\psi$ are also integral (i.e., constant over the cross section) and differential (i.e., of a local nature over the cross section with zero average) components of the deformation in $X$ direction.

The above equations and boundary conditions may be put together in a linear system of equations that should be solved in each cross section. For that purpose, the equations are first expressed in terms of displacements. Based on Eqs. (1*a*), (1*b*), and (3) one may write

$$\sigma_{\xi\xi}=C_{11}(u_{,x}-Yv_{,xx}-Zw_{,xx}+\psi_{,x})+C_{16}(\tilde{\psi}_{,\eta}) \tag{13a}$$

$$\tau_{\xi\eta}=C_{16}(u_{,x}-Yv_{,xx}-Zw_{,xx}+\psi_{,x})+C_{66}(\tilde{\psi}_{,\eta}). \tag{13b}$$

To simplify the following discussion, one should conceptually consider a discrete model for the distribution of the warping $\tilde{\psi}$ along the wall which may be denoted $\tilde{\psi}(1)...\tilde{\psi}(N)$. Subsequently, the above system of equations may be put as

$$[S]\begin{Bmatrix} u_{,x} \\ v_{,xxx} \\ w_{,xxx} \\ \tilde{\psi}_{(1)} \\ \vdots \\ \tilde{\psi}_{(N)} \end{Bmatrix} = \begin{Bmatrix} R_x-\tilde{R}_x \\ R_y-\tilde{R}_y \\ R_z-\tilde{R}_z \\ B_0-B_yY(1)-B_zZ(1)-\tilde{B}(1) \\ \vdots \\ B_0-B_yY(N)-B_zZ(N)-\tilde{B}(N) \end{Bmatrix} \tag{14}$$

where $\tilde{R}_x$, $\tilde{R}_y$, and $\tilde{R}_z$ are functions of $v_{,xx}$, $w_{,xx}$, $\psi_{,x}(\eta)$, and $\tilde{B}(i)$ are functions of $u_{,xx}$, $v_{,xx}$, $w_{,xx}$, $\phi_{,x}$, $\phi_{,xx}$, $\psi_{,x\eta}(i)$, $\psi_{,xx}(i)$. Note that the boundary conditions of Eqs. (11), (12) are embedded in the above system and so the condition of $\int\int_A\psi dA=0$. Initially, the terms $\tilde{R}_x$, $\tilde{R}_y$, $\tilde{R}_z$, and $\tilde{B}(i)$ are set to zero, and are further updated during the iterative process that will be discussed later on.

So far, the analysis assumed that the location of the $Y$–$Z$ system at each cross section is arbitrary. However, to comply with the previously discussed assumptions of vanishing torsional moment, Eq. (14) is reconstructed and solved for different locations of the $Y$–$Z$ system at each cross section, in order to find the location that will provide zero resultant moment at each cross section, namely

$$\int \int_A \tau_{\xi\eta} r dA = 0. \tag{15}$$

In the isotropic case, the location which is obtained by fulfilling Eq. (15) is known as the ''shear center.'' In composite beams, this location varies along the beam (in uniform beams as well) since the r.h.s. of Eq. (14) is a function of the spanwise coordinate *(X)*.

It is now possible to ''perturb'' the results obtained at the previous stage by adding arbitrary modifications to $Y$ and $Z$ that will be denoted $\Delta Y$ and $\Delta Z$, respectively. These modifications are required in order to locate the origin of all cross sections along the same line, namely, on a common $X$-axis that may be arbitrary selected as the ''beam longitudinal axis.'' A practical selection are the $\Delta Y$ and $\Delta Z$ values that will restore the original $X-Y-Z$ system as the global reference system. To cancel out the changes which are induced by the above $\Delta Y$ and $\Delta Z$ modifications, two additional steps need to be carried out. First, the torsional moment which is induced by the external loads at the modified system location should be determined and saved for further determination of the beam behavior under torsion (it will be added to the torsional moments $Q_x^i$ and $q_x$). In addition, in order not to induce any changes in the axial or shear strains (see Eqs. (13*a*) and (13*b*)), the above modifications require the introduction of an additional axial deformation $\Delta u_{,x}$ which is given by

$$\Delta u_{,x} = \Delta Y v_{,xx} + \Delta Z w_{,xx}. \tag{16}$$

At this stage, one may introduce the twist angle, $\phi_{,x}$, into the formulation. Again, this is the *bending-induced twist angle* while the *torsion-induced twist angle* should be determined separately. The determination of this component of the twist angle is carried out by the determination the warping function, $\psi$, using $\psi_{,\eta} = \tilde{\psi}_{,\eta} + r\phi_{,x}$ while the value of $\phi_{,x}$ is selected as the one that minimizes the linear component of the warping distribution, namely

$$\frac{\partial}{\partial \phi_{,x}}\left[\int \int_A (\psi Y)^2 + \int \int_A (\psi Z)^2\right] = 0. \tag{17}$$

The above criteria for selecting $\phi_{,x}$ emerge from three-dimensional or boundary conditions considerations at the beam root and/or tip ends. For example, for a complete warping restraint of a clamped beam, say at the root (where $v_{,x}=0$ and $w_{,x}=0$), the axial displacement given by $\bar{u} = u - Y v_{,x} - Z w_{,x} + \psi$ (see Eq. (1*a*)) dictates the requirement for the warping to be orthogonal to $Y$ and $Z$. Otherwise, nonvanishing value of $v_{,x}$ and $w_{,x}$ will be required to minimize the axial displacement distribution at the root. Note that such a fulfillment of Eq. (17) does not provide a complete axial displacement restraint that should include additional distribution of axial stresses as well. However, the above requirement ensures that the additional axial stress distribution will produce no axial resultant and no transverse bending moments, and its influence will be of a local nature only.

To carry out the above minimization operation, $\psi$ should be first determined using

$$\psi = \tilde{\psi} + \phi_{,x} \int r(\eta) d\eta + \text{const.} \tag{18}$$

where the constant term is selected so that $\int\int_A \psi dA = 0$.

To this end, the above-discussed steps yield the values of $u_{,x}$, $v_{,xxx}$, $w_{,xxx}$, $\phi_{,x}$, and $\psi(\eta)$ at each cross section. These values may be integrated and differentiated to give the distribution of $u_{,xx}(x)$, $u(x)$, $v_{,xx}(x)$, $v_{,x}(x)$, $v(x)$, $w_{,xx}(x)$, $w_{,x}(x)$, $w(x)$, $\psi_{,xx}(x,\eta)$, $\psi_{,x}(x,\eta)$, $\psi(x,\eta)$, $\phi_{,xx}(x)$, $\phi_{,x}(x)$, and $\phi(x)$. To carry out the above integrations, eight boundary conditions for $u$, $v_{,xx}$, $v_{,x}$, $v$, $w_{,xx}$, $w_{,x}$, $w$, and $\phi$ are required. In the present case of a cantilever beam, the values of $v_{,xx}$ and $w_{,xx}$ at the tip end of the beam are obtained by equating the resultant moments there to the corresponding internal moments, namely

$$[M_y^t, M_z^t] = \int \int_A \sigma_{\xi\xi}[Z, -Y] dA. \tag{19}$$

Similar relations are obtained between concentrated moments at a point along the beam $(Q_y^i, Q_z^i)$ and the corresponding changes (''jumps'') in the $v_{,xx}$ and $w_{,xx}$ values at that point. The remaining six geometrical boundary values are applied at the root:

$$u = v = v_{,x} = w = w_{,x} = \phi = 0. \tag{20}$$

The above values are used to update the r.h.s. of the system of equations at each cross section (see Eq. (14)). The solution is then returned to a sequential cross-sectional analyses until convergence is achieved.

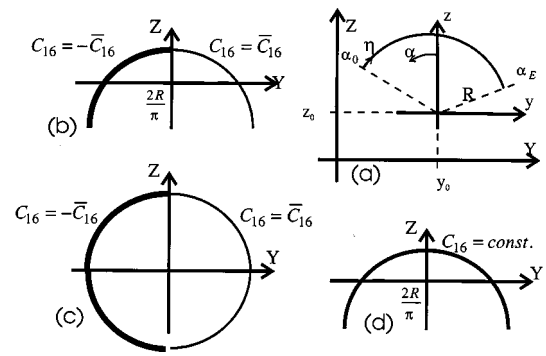



**Fig. 2 (*a*) A scheme of a circular thin-walled cross section; (*b*) a ''half-tube'' cross section of nonuniform elastic moduli (''symmetric'' layup); (*c*) an open circular cross section of nonuniform elastic moduli (''symmetric layup); (*d*) a ''half-tube'' cross section of uniform elastic moduli (''antisymmetric'' layup)**

## The Elastic Couplings

To analytically demonstrate the above theoretical model, a special case of a composite beam of open circular cross section undergoing bending moment and axial load is examined. Figure 2*(a)* presents the general notation.

**The Bending-Twist Coupling.** A closed-form solution that demonstrates the bending-twist coupling is presented in this section. As a simplified model, the ''half-tube'' cross section of Fig. 2(*b*) is used. To create the desired bending-twist coupling, the ply angle of the left-hand side of the cross section (i.e., the arc for $Y<0$ and $\alpha>0$), is assumed to be of equal magnitude but of opposite sign to the ply angle of the right hand side of the cross section (i.e., for $Y>0$ and $\alpha<0$). Similar to Fig. 1(*c*), such an arrangement creates a ''symmetric'' configurations where all elastic moduli are constant along the circular wall except for $C_{16}$ which takes the value $C_{16}=\bar{C}_{16}$ for $Y>0$ and, $C_{16}=-\bar{C}_{16}$ for $Y<0$. The above characteristics of the elastic moduli become clearer in light of the symmetric variation (with respect to the ply angle) of $C_{11}, C_{66}$, and $C_{55}$ and the antisymmetric variation of the $C_{16}$ (see, e.g., [19]). Hence, $\bar{C}_{16}$ in this case is the absolute value of $C_{16}$ obtained for the ply angle under consideration.

The beam is assumed to undergo a tip moment $M_z^t$ and therefore, a uniform resultant bending moment $M_z = M_z^t$ is induced at each cross section. Subsequently, it is assumed that $u(x) = w(x) = 0$, and that $v_{,xx}$ is constant. To obtain the desired solution, $\tilde{\psi}_{,\eta}$ is assumed to be equal to $-\bar{a}Y$ and $+\bar{a}Y$ over the left and right

sides of the cross section, respectively, where $\bar{a}$ is a constant. All of the above initial assumptions will be shown to be valid in what follows. Subsequently, Eq. (13*b*) shows that

$$\tau_{\xi\eta}=\begin{cases}-\bar{C}_{16}(-Yv_{,xx})-C_{66}\bar{a}Y & \text{for } Y<0\\ \bar{C}_{16}(-Yv_{,xx})+C_{66}\bar{a}Y & \text{for } Y\geqslant 0\end{cases}. \tag{21}$$

Hence, to assure that $\tau_{\xi\eta}$ will vanish along the *entire* cross section, one should require

$$\bar{a}=\frac{\bar{C}_{16}}{C_{66}}v_{,xx}. \tag{22}$$

Which also satisfies Eqs. (8), (11), (12), and (15). Subsequently, from Eqs. (13*a*) and (22) one may write

$$\sigma_{\xi\xi}=-\left(C_{11}-\frac{C_{16}^2}{C_{66}}\right)Yv_{,xx} \tag{23}$$

which satisfies Eq. (8) and with the aid of Eq. (19) shows that $M_y=0$ and

$$M_z=v_{,xx}\left(C_{11}-\frac{C_{16}^2}{C_{66}}\right)I_{yy} \tag{24}$$

where $I_{yy}$ is defined by Eq. (6). As shown, $C_{11}-C_{16}^2/C_{66}$ becomes the effective bending modulus, and Eq. (24) may be used for the determination of $v_{,xx}$ per given $M_z$. The warping is determined using Eq. (18) as

$$\psi_{,\eta}=\pm\bar{a}Y+R\phi_{,x}\left(1-\frac{2}{\pi}\cos\alpha\right) \tag{25}$$

where the upper and lower signs represent the $Y\geqslant 0$ and $Y<0$ regions, respectively. By substituting $Y=-R\sin\alpha$, Eq. (25) is integrated and with the aid of the requirement $\int\int_A\psi dA=0$, $\psi$ takes the form

$$\frac{\psi}{R^2v_{,xx}}=\frac{\bar{C}_{16}}{C_{66}}\left[\mp\cos\alpha\pm 1+\bar{b}\left(\frac{2}{\pi}\sin\alpha-\alpha\right)\right] \tag{26}$$

where $\bar{b}$ is defined as $\phi_{,x}/\bar{a}$ so that it reflects the geometrical coupling factor that determines the amount of twist per unit bending curvature since according to the above definition:

$$\frac{\phi_{,x}}{v_{,xx}}=\bar{b}\frac{\bar{C}_{16}}{C_{66}}. \tag{27}$$

Application of Eq. (17) shows that $\bar{b}$ is given by

$$\bar{b}=\frac{\int\int_A(\mp\cos\alpha\pm 1)YdA}{\int\int_A\left(\alpha-\frac{2}{\pi}\sin\alpha\right)YdA} \tag{28}$$

where all integrals containing $Z$ vanish and were therefore omitted. Carrying out the integrations for the present configuration yields $\bar{b}=-1$. Thus, a negative twist is induced by a lateral bending curvature.

Figure 3 presents the above elastic coupling as a function of the layup angle for a typical graphite/epoxy orthotropic laminae (in this case all laminae are identical and oriented at the same angle). As shown in ([19]), the elastic properties may be expressed in terms of engineering constants, which in the present case are given by: $E_{11}=130.\times 10^9\text{N/m}^2$, $E_{22}=E_{33}=12.\times 10^9\text{N/m}^2$, $G_{12}=G_{13}=6.\times 10^9\text{N/m}^2$, $G_{23}=4.\times 10^9\text{N/m}^2$, $\nu_{12}=\nu_{13}=0.3$, $\nu_{23}=0.5$. The evaluations of the $C'_{ij}$ elastic moduli as functions of the above engineering constants may also be found in ([19]).

Equation (26) shows that out-of-plane warping appears for nonzero ply angle and that the warping at the cross-sectional free tips, namely at the $\alpha=\pm 90$ deg locations is equal in magnitude but opposite in signs.

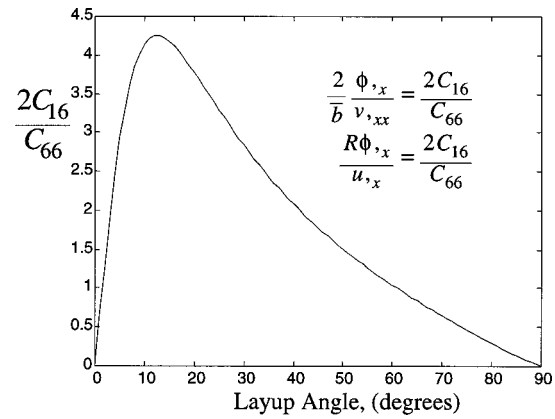



**Fig. 3 The induced twist ($\phi_{,x}$) due to bending ($v_{,xx}$) and extension ($u_{,x}$) as functions of the layup angle for a typical graphite/epoxy orthotropic laminae**

It is interesting to compare the above coupling magnitude, i.e., $\phi_{,x}/v_{,xx}=-\bar{C}_{16}/C_{66}$, with the magnitude obtained for a complete circular open cross section such as the one presented in Fig. 2(*c*). For such an open cross section, the $Y-Z$ system is located at the circle center, and Eqs. (21)–(28) are still valid, however, one should use $r=R$ and ignore the term $2/\pi\cos\alpha$ in Eq. (25) and the terms $2/\pi\sin\alpha$ in Eqs. (26), (28) (and clearly integrate over the entire circular cross section). The above calculation shows that $\bar{b}=-2/\pi$ in this case. The same value is obtained for a closed symmetric circular thin-walled cross section (see [10]). Thus, it may be concluded that a ''half-tube'' open cross section is producing a bending-twist coupling magnitude which is 57 percent larger than a symmetric (either open or closed) complete circular cross section. As an additional reference, it should be mentioned that ([9]) shows that the coupling magnitude in solid symmetric homogeneous cross sections may be expressed by Eq. (27) with $\bar{b}=-1/2$.

**The Extension-Twist Coupling.** To demonstrate the extension-twist coupling the ''half-tube'' cross section shown in Fig. 2(*d*) is adopted. The elastic moduli are assumed to be constant (i.e., not functions of $\eta$ and/or $\zeta$), and therefore, this layup configuration is generally referred to as ''antisymmetric'' (see e.g., Fig. 1(*d*)), and is expected to produce an extension-twist coupling. To demonstrate this coupling the following closed-form solution has been developed for a beam that undergoes a tip axial load (i.e., a uniform axial resultant load, $R_x(X)=F_x$).

By neglecting the warping derivatives in the $X$ direction due to the anticipated uniform longitudinal behavior and by assuming $v(x)=w(x)=0$, $u_{,xx}=0$ and $u_{,x}=\text{const.}$ (as will be proved to be valid later on), Eq. (11) shows that

$$\frac{d\tilde{\psi}}{d\eta}=-\frac{C_{16}}{C_{66}}u_{,x}. \tag{29}$$

Then, by substituting Eq. (29) in Eq. (13*a*) and using Eq. (8), the axial strain becomes

$$u_{,x}=\frac{1}{\left(C_{11}-\frac{C_{16}^2}{C_{66}}\right)}\frac{F_x}{A} \tag{30}$$

which shows that indeed $u_{,xx}=0$ and that similar to Eq. (24), $C_{11}-C_{16}^2/C_{66}$ is the effective axial modulus in this case. Hence, the solution of Eq. (14) for each cross section is represented by Eqs. (29), (30). Note that Eq. (15) is satisfied regardless of the values of $y_0$ and $z_0$. However, the boundary conditions of Eq. (19) shows that to ensure $v_{,xx}=0$ and $w_{,xx}=0$ at the tip end of the

beam (which is essential to support the above $v(x)=w(x)=0$ assumption), one should select $y_0=0$ and $z_0=-2R/\pi$. Such a selection locates the $Y-Z$ origin at the cross-sectional area center (see Fig. 2(*d*)). Otherwise, the above-obtained axial stress distribution (which was found to be uniform over the cross section) will produce nonvanishing transverse tip moments. For this cross section, $r=R(1-2/\pi\cos\alpha)$, and subsequently, in order to determine $\phi_{,x}$ using Eq. (17), $\psi_{,\eta}$ takes the form

$$\psi_{,\eta}=-\frac{C_{16}}{C_{66}}u_{,x}+R\phi_{,x}\left(1-\frac{2}{\pi}\cos\alpha\right). \tag{31}$$

Integrating the above distribution yields

$$\psi=\frac{\alpha}{\alpha_{,\eta}}\left(-\frac{C_{16}}{C_{66}}u_{,x}+R\phi_{,x}\right)-\frac{2R}{\pi\alpha_{,\eta}}\phi_{,x}\sin\alpha+\text{const.} \tag{32}$$

The requirement $\int\int_A\psi dA=0$ shows that the above constant vanishes, and Eq. (17) is satisfied provided that

$$R\phi_{,x}\left(1-\frac{2}{\pi}\frac{\int\int_A Y\sin\alpha dA}{\int\int_A Y\alpha dA}\right)=\frac{C_{16}}{C_{66}}u_{,x} \tag{33}$$

where the integrals containing the $Z$ coordinate vanish due to symmetry considerations. Carrying out the involved integrations shows that

$$\phi_{,x}=\frac{2}{R}\frac{C_{16}}{C_{66}}u_{,x} \tag{34}$$

As shown, the ratio of $C_{16}/C_{66}$ is the parameter that derives the variation with respect to the layup angle, similar to the bending-twist case (see Fig. 3). Substituting Eq. (34) in Eq. (32) yields:

$$\frac{\psi}{R}=\frac{C_{16}}{C_{66}}u_{,x}\left(\frac{4}{\pi}\sin\alpha-\alpha\right) \tag{35}$$

which shows that the warping magnitude is derived by $C_{16}/C_{66}u_{,x}$. Similar to the pure isotropic case, the case of zero ply angle contains no warping. As the ply angle is increased, a significant warping distribution which is of the order of the uniform axial extension $u_{,x}$ appears. Note that $u_{,x}$ itself is usually growing with the layup angle since the axial stiffness is dramatically softened due to the high $E_{11}/E_{22}$ ratio in this case.

Comparison of Eq. (34) with the expression of $\phi_{,x}=(p/2A)(C_{16}/C_{66})u_{,x}$ obtained in the case of a single-cell closed thin-walled cross section (where $A_m$ is the area enclosed by the median line and $p$ is its circumference, see Ref. [9], shows that these two extension-twist coupling expressions are of the same order.

## Concluding Remarks

Theoretical model for the structural response of composite thin-walled beams has been derived. The main feature of the present analysis is its capability to include a detailed out-of-plane warping distribution regardless of its origin. The analysis presents a generic treatment of the beam behavior under transverse bending and axial loading, and the discussion includes comparisons with similar beams of closed cross sections. Closed-form solutions for simple cases are discussed and provide a clear insight into both the bending-twist and the extension-twist coupling mechanisms. Some of the specific findings of the present analysis are

(i) In the general case, all ''beam'' degrees-of-freedom (i.e., the axial and two transverse displacements and the twist angle) are fully coupled in composite thin-walled open beams.

(ii) The elastic coupling characteristics are strongly dependent on the out-of-plane warping which plays a key role in their analytic prediction. Prescribed warping distribution may result in complete different coupling characteristics.

(iii) In thin-walled open beams, the cross-sectional location where applied bending loads induce no moment (which is referred to as ''shear center'' in isotropic beams) is a function of the loading, spanwise location and boundary conditions.

(iv) Open cross sections provide larger bending-twist coupling magnitudes than similar closed cross sections.

(v) The amount of twist per unit axial strain which is obtained due to axial load in open cross sections is of the same order as for similar closed cross sections.

## References


[1] Hoff, N. J., ''Thin Shells in Aerospace Structures,'' Astron. Aeronaut., 1967 **5**, pp. 26–45.

[2] Bull, J. W., 1988 ''Finite Element Analysis of Thin-Walled Structures,'' Elsevier, Amsterdam, The Netherlands.

[3] Vlasov, V. Z., 1961, ''Thin-Walled Elastic Beams,'' National Science Foundation, U.S. Department of Commerce, Washington, D.C., and the Israel Program for Scientific Translations, Jerusalem (translated from Russian).

[4] Gjelsvik, A., 1981, ''The Theory of Thin Walled Bars,'' John Wiley and Sons, New York.

[5] Librescu, L., and Song, O., 1991, ''Behavior of Thin-Walled Beams Made of Advanced Composite Materials and Incorporating Non-Classical Effects,'' Appl. Mech. Rev., **44**, No. 11/2, pp. s174–180.

[6] Berdichevsky, V., Armanios, E., and Badir, A., 1992, ''Theory of Anisotropic Thin-Walled Closed-Cross-Section Beams,'' Composites Eng., **2**, No.5–7, pp. 411–432.

[7] Bauchau, O. A., and Chiang, W., 1993, ''Dynamic Analysis of Rotor Flexbeams Based on Nonlinear Anisotropic Shell Models,'' J. Am. Helicopter Soc., **38**, No. 1, pp. 55–61.

[8] Armanios, E. A., and Badir, A. M., 1995, ''Free Vibration Analysis of Anisotropic Thin-Walled Closed-Section Beams,'' AIAA J., **33** No. 10, pp. 1905–1910.

[9] Rand, O., 1998, ''Fundamental Closed-Form Solutions for Solid and Thin-Walled Composite Beams Including a Complete Out-of-Plane Warping Model,'' Int. J. Solids Struct., **35**, No. 21, pp. 2775–2793.

[10] Rand, O., 1998, ''Similarities Between Solid and Thin-Walled Composite Beams by Analytic Approach,'' J. Aircr., **35**, No. 4, pp. 604–615.

[11] Ganguli, R., and Chopra, I., 1997, ''Aeroelastic Tailoring of Composite Couplings and Blade Geometry of a Helicopter Rotor Using Optimization Methods,'' J. Am. Helicopter Soc., **42**, No. 3, pp. 218–228.

[12] Bauld, N. R., and Tzeng, L. S., 1984, ''A Vlasov Theory for Fiber-Reinforced Beams With Thin-Walled Open Cross-Sections,'' Int. J. Solids Struct., **20**, No. 3, pp. 277–297.

[13] Chandra, R., and Chopra, I., 1991, ''Experimental and Theoretical Analysis of Composite I—Beam with Elastic Couplings,'' AIAA J., **29**, No. 12, pp. 2197 2206.

[14] Wu, X. X., and Sun, C. T., 1992, ''Simplified Theory for Composite Thin-Walled Beams,'' AIAA J. **30**, No. 12, pp. 2945–2951.

[15] Park, M. S., and Byung, C. L., 1996, ''Prediction of Bending Collapse Behaviors of Thin-Walled Open Section Beams,'' Thin-Walled Struct., **25**, No. 3, pp. 185–206.

[16] Eisenberger, M., 1997, ''Torsional Vibrations of Open and Variable Cross-Section Bars,'' Thin-Walled Struct., **28**, No. 3, pp. 269–278.

[17] QuanFeng, W., 1997, ''Lateral Buckling of Thin-Walled Open Members With Shear Lag Using Optimization Techniques,'' Int. J. Solids Struct., **34**, No. 11, pp. 1343–1352.

[18] Kaiser C., and Francescatti, D., 1996 ''Theoretical and Experimental Analysis of Composite Beams With Elastic Couplings,'' 20th Congress of the International Council of the Aeronautical Sciences, ICAS-96-5.4.4, Sorrento, Napoli, Italy. Sept. 8–13

[19] Ochoa, O. O., and Reddy, J. N., 1992, ''Finite Element Analysis of Composite Laminates,'' Kluwer, Dordrecht, The Netherlands.

# Use of Graded Finite Elements to Model the Behavior of Nonhomogeneous Materials

**M. H. Santare**
Mem. ASME

**J. Lambros**[1]
Mem. ASME

Department of Mechanical Engineering,
University of Delaware,
Newark, NJ 19716

*A finite element with a spatially varying material property field is formulated and compared to a conventional, homogeneous element for solving boundary value problems involving continuously nonhomogeneous materials. The particular element studied is a two-dimensional plane stress element with linear interpolation and an exponential material property gradient. However, the main results are applicable to other types of elements and property gradients. Exact solutions for a finite rectangular plate subjected to either uniform displacement or traction either perpendicular or parallel to the property gradient are used as the basis for comparison. The results show that for identical meshes with equal number of degrees-of-freedom, the graded elements give more accurate local stress values than conventional elements in some boundary value problems, while in other problems the reverse is true.* [S0021-8936(00)01504-X]

## 1 Introduction

In recent years, there has been growing interest in materials with mechanical properties that are continuous functions of spatial variables. These nonhomogeneous materials arise naturally from manufacturing processes which produce thermal and chemical gradients and from biological synthesis processes. In addition, functionally graded materials (FGM's) are materials engineered to exhibit property gradients for the purpose of tailoring their performance characteristics ([1]).

A significant amount of research has been invested in better understanding of the mechanics of these materials. Erdogan and his co-workers have analytically solved a number of fracture problems involving FGM's (see, for example, [2–4]). Gu and Asaro [5,6] analytically studied the stress fields associated with semi-infinite cracks and calculated the resulting initial crack deflection angle. Because these solutions are analytical they are constrained by certain simplifying assumptions. First, the geometries are simple two-dimensional domains containing single, straight cracks. In addition, the functional form of the nonhomogeneity is always an exponential, which can be written in the general form

$$E(x,y)=E_0 \exp(\beta x+\gamma y) \tag{1}$$

where $E(x,y)$ is the elastic modulus at the coordinate $(x,y)$ and $E_0$, $\beta$ and $\gamma$ are material constants ([3]). In order to overcome these limitations for the solution of more general boundary value problems, a numerical solution method is needed.

Gu et al. [7] used finite elements to study the strain energy release rate of fracture in an FGM. By constructing a highly refined, focused mesh around the crack tip, they were able to numerically verify the relationship between the J-integral and the strain energy release rate for a static crack. In that study, Gu et al. [7] mention that they achieve the material gradient by assigning corresponding material properties at the Gauss integration points. A similar technique for assigning material properties is employed in this paper for the development of the graded elements. Anlas et al. [8] looked at the convergence of standard finite element solution techniques to evaluate stress intensity factors in FGM's. They were able to show that a relatively coarse mesh and conventional, homogeneous finite elements can be used in many situations. An important aspect of the work by Anlas et al. [8] is that they used a uniform, rectangular mesh. This simple geometry lends itself to a straightforward discretization of a rectilinear material property field. The element rows are aligned with the gradient direction and each row of homogeneous elements is assigned the material properties for the midline of the row. This produces a step-wise constant approximation to a continuous material property field such as the one described in Eq. (1). This same approximation was used in Li et al. [9] in solving for strain energy release rates in cracked FGM's. Although the work of Anlas et al. [8] shows that a piecewise constant approximation for the material property gradient can be used to provide fairly accurate global energy calculations, the local stresses and displacements within each element are affected. In other applications, such as failure analysis or crack growth predictions, these local values may be of critical importance.

A finite element that admits its own material property gradient would provide a more accurate solution at the size scale of the elements themselves. The formulation of such an element is commonly mentioned in references on the derivation of finite element stiffness matrices (see, for example, Bathe and Wilson [10]). In practice, however, this formulation is normally overlooked in favor of the much simpler methodology of increasing mesh refinement. In this paper we discuss the derivation and implementation of a finite element with its own material property gradient for the solution of simple boundary value problems in FGM's. Although the ultimate usefulness of such an element is linked to the complexity of the stress field, we will examine several simple two-dimensional cases as a means of comparing among the alternative numerical modeling approaches and known exact analytical solutions.

## 2 Element Formulation

In the standard derivation of the stiffness matrix for an isoparametric finite element, one starts with an assumed set of interpolation functions such that (see, for example, Bathe and Wilson [10])

$$u(x)=\sum_{I=1}^{n} N_I(x)U_I. \tag{2}$$

[1]Current address: Aeronautics and Astronautics Department, University of Illinois at Champaign-Urbana, 310 Talbot Lab, MC 236, 104 S. Wright Street, Urbana, IL 61501.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Feb. 1, 2000, 1999; final revision, May 5, 2000. Associate Technical Editor: M. Ortiz. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

 Copyright © 2000 by ASME 

Here $u(x)$ is a matrix of the displacement components (each a function of position within the element), $N_I(x)$ is the matrix of interpolation or shape functions, and $U_I$ are the nodal displacement values for each of the ''$n$'' nodes of the element. To find the infinitesimal strain components, $\varepsilon(x)$ one takes the derivatives of the shape functions to get

$$\varepsilon(x)=\sum_{I=1}^{n} B_I(x)U_I \tag{3}$$

where the $B_I(x)$ matrix is populated with the appropriate derivatives of the $N_I(x)$ functions.

Assuming linear elastic material behavior, the stress components at a point $\sigma(x)$ are calculated from the strain at that point through the material property matrix $C(x)$ such that

$$\sigma(x)=C(x)\varepsilon(x). \tag{4}$$

Traditionally, the components of the $C(x)$ matrix are a set of constant material properties for each finite element. However, there is no fundamental reason that these elastic properties cannot be spatially variable functions within an element. In the case of an FGM, the components of $C(x)$ could be explicit functions describing the actual material property gradients. In any case, we define the element stiffness matrix $K^e$ as the linear function that maps the nodal displacements $U_I$ to the nodal forces $f_I$ for the element

$$f_I=K^eU_I. \tag{5}$$

By the principle of virtual work, the work done by the nodal forces must be equal to the work of deformation within the element. Setting these two quantities equal we can derive the following expression for the element stiffness matrix:

$$K^e=\int_{V_e} B^T(x)C(x)B(x)\,dV \tag{6}$$

where the integral is taken over the volume of the element $V_e$.

Up to this point, the only numerical approximation used in the formulation was the imposition of the interpolation functions in Eq. (2). On the element level, the accuracy of this numerical approximation is dependent on the compatibility of the assumed shape functions $N_I(x)$ to the exact displacement field. (Exact here refers to the analytical solution of the field equations under the assumptions of linear elastic response and infinitesimal strain.) In fact, if a set of shape functions is chosen which is perfectly compatible with the exact displacement solution, the finite element results will capture this solution with any level of mesh refinement. For example, a bilinear displacement field will be accurately predicted by one or more elements employing bilinear interpolation functions.

## 3 Material Property Gradients

The stiffness matrix for an element with graded material properties can be calculated using Eq. (6). The $C(x)$ matrix in this case will contain the material properties as explicit functions of the spatial variables. In general the material property fields can be continuous or discontinuous, isotropic, or anisotropic, and can be written in closed form or as a numerical approximation. Any functions which allow for the evaluation of the integral in Eq. (6) can be used. Since in this paper we intend to compare numerical results to analytical solutions from Erdogan and Wu [4] we will impose an isotropic property field with a constant Poisson's ratio and an exponential variation in elastic modulus as specified in Eq. (1). To further simplify the problem, we assume a rectilinear material property field and align the $x$-axis in the direction of the gradient by letting $\gamma=0$. Under these assumptions $E_0$ is the elastic modulus at $x=0$ and $1/\beta$ is the length scale of the nonhomogeneity. Although a single set of parameters will be used in each boundary value problem, the stiffness matrix of an individual element will depend on its location in the property field since $C(x)$ in equation (6) is a function of position.

For comparison, we will also solve the finite element problem using a step-wise constant material property field. This is the method used in Anlas et al. [8] as described in the Introduction. In this case, the $C(x)$ matrix for a specific element is assumed constant and assigned the property values at the centroid of the element. This approach allows the direct implementation of commercially available finite element packages, but imposes an additional discretization error in the modeling of the material property field.

## 4 Results

Consider a finite uncracked plate ($0\leqslant x\leqslant w, -h\leqslant y\leqslant h$) made of a continuously nonhomogeneous material with an exponentially varying elastic modulus in the $x$-direction according to Eq. (1) with $\gamma=0$ (see Fig. 1). The plate is subjected to a uniform traction or displacement perpendicular (Fig. 1($a$)) or parallel (Fig. 1($b$)) to the direction of the property gradient. The solutions will all be symmetric with respect to the midline, which in this case is the $x$-axis. Erdogan and Wu [4] analytically solved the problems with loading in the $y$-direction (Fig. 1($a$)) to be used as far-field solutions for the edge-cracked plate problem. In the special case where Poisson's ratio is zero, the exact solutions to the problems with loading in the $x$-direction (Fig. 1($b$)) are quite simple.

First, consider the constant displacement boundary condition on ($y=h$), as shown in Fig. 1($a$) where the arrows represent a uniform displacement. The normal stress in the $y$-direction $\sigma_{yy}(x,y)$ is quite easy to find and can be written

$$\sigma_{yy}(x,y)=E(x)\,\frac{u_0}{h} \tag{7}$$

where $u_0$ is the imposed constant displacement and $h$ is the half-length of the plate in the $y$-direction. All other stress components are zero and this solution is exact everywhere in the plate.

For the case of uniform applied traction on ($y=h$), Fig. 1($a$) where the arrows represent a uniform traction, the exact solution is not so simple. Under this loading, a constant stress field cannot be sustained in a generally nonhomogeneous elastic material since it would violate compatibility. However, it can be shown that a uniform applied traction in the $y$-direction gives the following exact solution along the line of symmetry ($x$-axis)

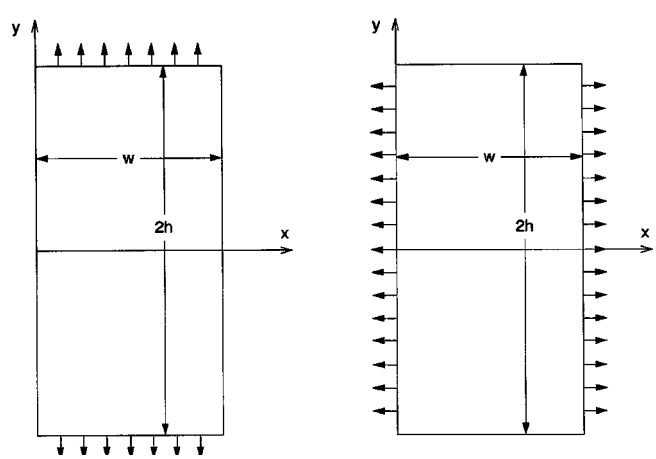



**Fig. 1 Finite nonhomogeneous plate subjected to a uniform displacement or traction; (*a*) perpendicular to the material gradient direction, (*b*) parallel to the material gradient direction**

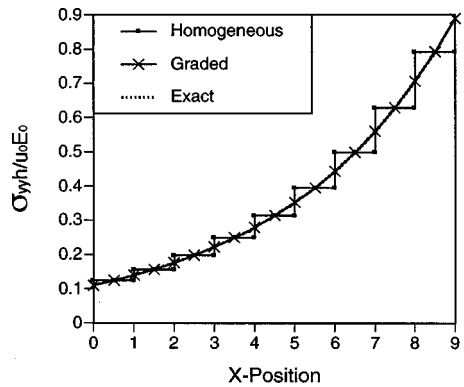



**Fig. 2 Normalized stress component versus position in nonhomogeneous plate subjected to uniform displacement in the *y*-direction**

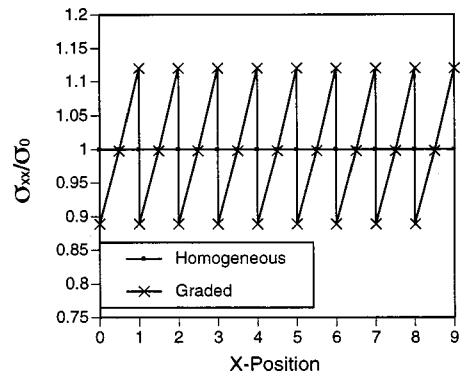



**Fig. 4 Normalized stress component versus position in nonhomogeneous plate subjected to uniform traction in the *x*-direction**

$$\sigma_{yy}(x,y=0) = \frac{\beta N e^{\beta w}}{2} \cdot \frac{2\beta x+\beta^2 xw - e^{\beta x}(2\beta x-\beta^2 xw+4+\beta^2 w^2)+4+\beta w}{e^{\beta w}\beta^2 w^2 - e^{2\beta w}+2e^{\beta w}-1} \tag{8}$$

where $N$ is the total applied load. Along this line all other stress components are zero. For the case of a very long strip, this solution can be considered an approximate solution far from the loaded edges.

For the cases where a uniform traction or displacement is applied parallel to the gradient direction, shown in Fig. 1(*b*), an exact solution can be found easily if the Poisson's ratio is set to zero. In this case, with either type of loading, the solution is essentially one-dimensional with

$$\sigma_{xx}(x,y)=\sigma_0 \tag{9}$$

where $\sigma_0$ is the applied uniform traction (or in the case of a uniform displacement, the traction resulting from the applied displacement) and all other stress components are equal to zero.

In the following graphs we compare these analytical solutions (exact at least along the line $y=0$) to the finite element results generated from a 100-node two-dimensional mesh (200 degrees-of-freedom). In both the exact solution and the finite element solutions the material modulus varies according to Eq. (1) with $E_0=1$, $\gamma=0$ and $\beta w=\ln 8$ so that $E(x,y)|_{(x=w)}=8$. In the solutions for loading in the *y*-direction, the Poisson's ratio is constant at 0.3 and in the *x*-direction loadings, Poisson's ratio is zero everywhere. In all the numerical solutions, four-node isoparametric plane-stress elements are used. The mesh is made up of 9 by 9 perfectly square uniform elements and represents the upper half $(0\leqslant y\leqslant h)$ of the symmetric domain. For the cases with loading in the *y*-direction, the model is subjected to a uniform traction or displacement along the upper edge $(y=h)$ and zero vertical displacement along the lower edge $(y=0)$ to model symmetry. For the cases with loading in the *x*-direction, the model is subjected to uniform displacement or traction along the right edge $(x=w)$ zero horizontal displacement along the left edge $(x=0)$ and zero vertical displacement along the bottom $(y=0)$. The results labeled ''homogeneous'' are generated using standard conventional finite elements each with a constant material property matrix. The term homogeneous refers to the properties within an element, this is still a model for a nonhomogeneous plate. Therefore, the property gradient is this case is a piecewise constant approximation to Eq. (1), discretized on the same size scale as the elements. The results labeled ''graded'' are generated using elements with their own material property fields. The $K^e$ matrices for these elements are calculated through Eq. (7) with variable elastic modulus as in Eq. (1). In this case the property gradient is incorporated into the element stiffness matrix and is continuous between elements.

Figure 2 shows the normalized stress component $\sigma_{yy}h/u_0E_0$ as a function of $x$ in a nonhomogeneous plate subjected to a uniform displacement in the *y*-direction. For this example, the stress is uniform in the *y*-direction, so this graph applies for all *y*. Figure 3 shows the normalized stress component $\sigma_{yy}w/N$ in the same plate subjected to a uniform traction in the *y*-direction. Recall that the exact solution is valid only along the line of symmetry $y=0$. Therefore, the finite element results are reported along this line as

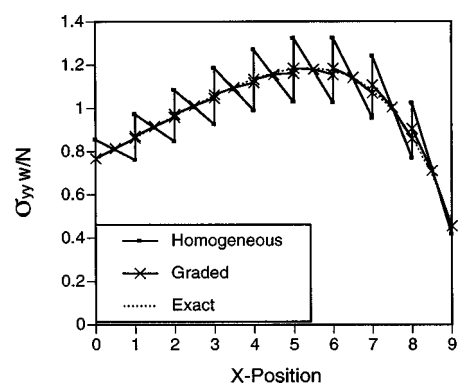



**Fig. 3 Normalized stress component versus position in nonhomogeneous plate subjected to uniform traction in the *y*-direction**

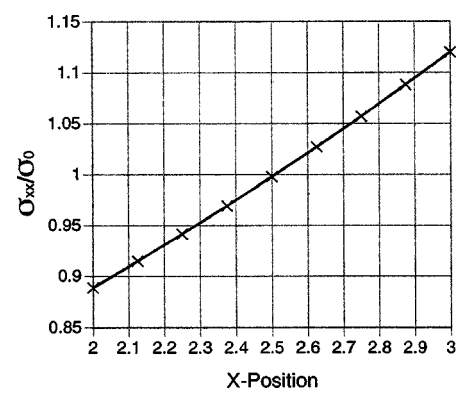



**Fig. 5 Rescaling of the data from Fig. 4 for a single graded element**

well. Figure 4 shows the normalized stress component $\sigma_{xx}/\sigma_0$ for the uniform traction loading in the $x$-direction and is valid everywhere in the plate. A closeup of the same data for a single element is shown in Fig. 5. The results are nearly identical for the case of an applied uniform displacement and therefore these results are omitted. In all the finite element solutions, the stress at a specific point is found by substituting the calculated nodal displacement values back into Eqs. (3) and (4). For each element, the material property matrix $C(x)$ is the same as the one used in the calculation of the element stiffness matrix (Eq. (7)).

## 5 Discussion

Focusing first on Fig. 2, notice that the graded solution and the exact solution are identical. This should not be surprising since the exact displacement field is linear and can therefore be captured by the linear isoparametric elements. In fact a single graded element could be used to predict this simple solution exactly. The homogeneous solution in this figure is piecewise constant. This is expected since the elements in the conventional formulation each have a single, homogeneous material property. These elements have the same linear shape functions as the graded elements, however, they possess a piecewise-constant material property approximation. Therefore, they predict the proper stress values only at their centroids where their properties exactly match the material gradient function.

Figure 3 shows the normalized stress component $\sigma_{yy}w/N$ in the same plate, this time subjected to a uniform vertical traction at $y=h$. Recall that the analytical solution given in Eq. (8) is exact only along the line of symmetry ($y=h$) so we will compare only the stresses calculated along that line. For this problem both finite element solutions are piecewise continuous. This is because the linear interpolation functions used for both element types cannot capture the actual displacement field (which in this case is exponential). Clearly, however, the graded elements provide a better approximation to the exact solution in every element. In the graded elements, the linear displacement results in an exponential stress field while in the homogeneous elements, the result is a linear stress variation. In particular, notice the elements on the left side of the domain (i.e., small $x$). The homogeneous elements give $\sigma_{yy}$ as a decreasing function of $x$ in each individual element while the exact solution shows an increasing trend. In the loading cases examined in Fig. 2 and 3, the stress components $\sigma_{xx}$ and $\sigma_{xy}$ are zero in the exact solution. Both finite element solutions give fairly good approximations to this (for example $|\sigma_{xx}/\sigma_{yy}|<0.1$ everywhere except the endpoints and is $<0.05$ in most elements).

The differences shown in Fig. 3 are the result of the piecewise-constant approximation to the actual material property gradient used in the conventional elements. The finite element solution provides a numerical ''best fit'' to the exact solution within the constraints of the formulation. In the exact solution, the normal strain in the $y$-direction is higher on the left side of the elements and both numerical formulations predict this. However, in the case of the conventional homogeneous elements, this higher strain automatically translates into a higher stress due to the fact that the material properties within the element are constant. The graded elements, on the other hand, can better accommodate this strain difference because of their nonhomogeneous properties.

Figure 4 shows the normalized stress component $\sigma_{xx}/\sigma_0$ for the case where the loading is in the $x$-direction. Here we see that the conventional elements capture the exact solution, $\sigma_{xx}/\sigma_0=1$, while the graded elements give a piecewise continuous solution. This is the reverse of the effect seen in the previous load cases. A closer examination of the results as seen in Fig. 5 reveals that these are actually piecewise exponential functions. In this load case, the stress in the $x$-direction is a constant throughout the body and therefore, according to Eq. (4) the strain must vary exponentially in this direction. In both formulations, the linear interpolation functions cannot capture this variation in strain within an element, so they instead predict a constant normal strain in each element. In the conventional homogeneous element, this results in a constant stress, while in the graded element, this gives an exponential variation in stress.

## 6 Conclusions

In this study graded finite elements were formulated using linear interpolation functions and an exponential material property gradient and compared to conventional homogeneous elements based on the same shape functions. This particular graded element is ideal for solving the uniform displacement-loading problem of Fig. 2, and therefore it is not surprising that it outperforms the conventional element for this problem. In fact, this graded element will capture this exact solution at any level of mesh refinement. In addition, this graded element also outperforms the conventional finite element when the load is an applied traction perpendicular the gradient direction. In both cases, the loss in accuracy for the conventional models is due to the material property discretization inherent in a homogeneous element. What was unexpected, however, is the fact that the same graded element gives a much poorer approximation to a uniform traction parallel to the gradient direction.

We have shown here that by using graded elements without increasing the degrees of freedom one can get more accurate local values for some stress components, but risks losing accuracy in others. Of course, it is possible to formulate other graded elements with higher order interpolation function and/or different material property gradients for specific applications. In each case, the analyst must weigh the additional cost of formulation against CPU time. For example, in an earlier study, Anlas et al., [8] showed that in many cases conventional elements can be used to provide reasonably accurate solutions for overall energy calculations at moderately low mesh densities. However, in some applications, such as failure analysis or crack path predictions, the local stress values may be of critical importance. The results of this study show that one must be cautious when using this type of graded element in calculating local stress values within an element.

## Acknowledgment

The authors wish to acknowledge support for this research through the National Science Foundation Grant CMS-9712831. We would like to thank Philippe Geubelle for his helpful insights.

## References

[1] Erdogan, F., 1995, ''Fracture Mechanics of Functionally Graded Materials,'' Composites Eng., **5**, No. 7, pp. 753–770.
[2] Delale, F., and Erdogan, F., 1993, The Crack Problem for a Nonhomogeneous Plane,'' J. Appl. Mech., **50**, pp. 609–614.
[3] Konda, N., and Erdogan, F., 1994, ''The Mixed Mode Crack Problem in a Nonhomogeneous Elastic Medium,'' Eng. Fract. Mech., **47**, No. 4, pp. 533–545.
[4] Erdogan, F., and Wu, B. H., 1997, ''The Surface Crack Problem for a Plate With Functionally Graded Properties,'' ASME J. Appl. Mech., **64**, pp. 449–456.
[5] Gu, P., and Asaro, R. J., 1997, ''Cracks in Functionally Graded Materials,'' Int. J. Solids Struct., **34**, pp. 1–7.
[6] Gu, P., and Asaro, R. J., 1997, ''Cracks Deflection in Functionally Graded Materials,'' Int. J. Solids Struct., **34**, pp. 3085–3098.
[7] Gu, P., Dao, M., and Asaro, R. J., 1999, ''A Simplified Method for Calculating the Crack Tip Field of Functionally Graded Materials Using the Domain Integral,'' ASME J. Appl. Mech., **66**, pp. 101–108.
[8] Anlas, G., Santare, M. H., and Lambros, J., 2000, ''Numerical Calculation of Stress Intensity Factors in Functionally Graded Materials,'' Int. J. Fract., **104**, No. 2, pp. 131–143.
[9] Li, H., Lambros, J., Cheeseman, B. A., and Santare, M. H., 2000, ''Experimental Investigation of the Quasi-Static Fracture of Functionally Graded Materials,'' Int. J. Solids Struct., **37**, No. 27, pp. 3715–3732.
[10] Bathe, K.-J., and Wilson, E. L., 1976, *Numerical Methods in Finite Element Analysis*, Prentice-Hall, Englewood Cliffs, NJ.
[11] Delale, F., and Erdogan, F., 1983, ''The Crack Problem for a Nonhomogeneous Plane,'' ASME J. Appl. Mech., **50**, pp. 609–614.

# Solution of the Moving Mass Problem Using Complex Eigenfunction Expansions

**K.-Y. Lee**

**A. A. Renshaw**
Mem. ASME

Department of Mechanical Engineering,
Columbia University,
Mail Code 4703,
New York, NY 10027

*A new solution technique is developed for solving the moving mass problem for nonconservative, linear, distributed parameter systems using complex eigenfunction expansions. Traditional Galerkin analysis of this problem using complex eigenfunctions fails in the limit of large numbers of trial functions because complex eigenfunctions are not linearly independent. This linear dependence problem is circumvented by applying a modal constraint on the velocity of the distributed parameter system (Renshaw, A. A., 1997, J. Appl. Mech.,* ***64****, pp. 238–240). This constraint is valid for all complete sets of eigenfunctions but must be applied with care for finite dimensional approximations of concentrated loads such as found in the moving mass problem. Numerical results indicate that the proposed method is competitive with Galerkin's method with real trial functions in terms of accuracy and rate of convergence.* [S0021-8936(00)00604-8]

## 1 Introduction

When analyzing the dynamics of a system that is related to or a perturbation of a well-known linear distributed parameter system, the most efficient trial functions to use for approximate solution techniques are frequently the eigenfunctions of the underlying distributed parameter system. For example, Wickert and Mote [1] show that one-term Rayleigh quotient approximations of the natural frequencies of an axially translating cable are much more accurate when the complex eigenfunctions of the axially moving string are used as trial functions rather than those of the stationary string. When the actual system under study is linear and time invariant, eigenfunction expansions are particularly efficient since the equations of motion decouple and exact closed-form solutions are obtained whenever the eigenfunctions are complete ([2–4]).

For complex eigenfunctions of non-self-adjoint systems, eigenfunction expansion techniques can be problematic because the eigenfunctions are not linearly independent. When using Galerkin's method, this lack of linear independence renders the mass matrix singular as the number of terms used in the expansion increases, which eventually causes the method to fail. For self-adjoint systems, it is straightforward to extract a complete, linearly independent set of eigenfunctions since complex conjugate eigenvalues possess the same real eigenfunction. For complex eigenfunctions such an extraction could be performed using Gram-Schmidt orthogonalization, but this would destroy the desired eigenfunction properties.

In the present paper, we develop a general technique for using complex eigenfunction expansions to approximate system dynamics and apply this method to the solution of the moving mass problem. In the moving mass problem, a concentrated mass moves across and is coupled to a linear distributed system such as a string or cable. The distributed system may be stationary or moving. This problem originally arose in the context of railroad bridge failures and has since been used to study cable tramways, ski lifts, and computer disk drives. A large number of techniques have been proposed for solving this problem including finite element methods ([5]), integro-differential equation techniques ([1,6]), and eigenfunction expansions ([7–11]). When an eigenfunction expansion technique has been used, the eigenfunctions used were almost always real and self-adjoint. In Pesterev and Bergman [9], a method was determined for the special case of non-self-adjoint systems with real eigenfunctions. When the eigenfunctions are complex, existing solution techniques cannot be used. This occurs, for instance, when the distributed parameter system either moves or possesses viscous damping.

In the present paper, we derive a method for using complex eigenfunction expansions to solve the moving mass problem. The linear dependence problem for complex eigenfunctions is circumvented by applying a modal constraint on the velocity of the distributed parameter system ([3]). This constraint is valid for all complete sets of eigenfunctions but must be applied with care in the case of finite dimensional approximations of concentrated loads such as found in the moving mass problem. Numerical results indicate that the proposed method is competitive with Galerkin's method using real trial functions in terms of accuracy and rate of convergence.

## 2 Mathematical Preliminaries

Consider the nonconservative, linear, distributed parameter system whose equation of motion is

$$Mw_{,tt}+Cw_{,t}+Kw=f \tag{1}$$

where $M$, $C$, and $K$ are linear, real, differential operators in a spatial domain, $M>0$, $x$ is the spatial position, $t$ is time, $w(x,t)$ is the system displacement, and $f(x,t)$ are external forces. We assume the existence of a sufficiently differentiable Hilbert space with complex inner product $\langle,\rangle$ on which $M$, $C$, $K$, and their adjoint operators, $M^*$, $C^*$, and $K^*$, can be defined for appropriate boundary conditions. In the special case of positive definite gyroscopic systems, $M=M^*$, $C=-C^*$, $K=K^*$ and $K>0$. These are not assumed in the following unless explicitly stated.

The eigenvalue problem associated with (1) is

$$\lambda_n^2 Mu_n+\lambda_n Cu_n+Ku_n=0 \tag{2}$$

and the eigensolutions $\{\lambda_n,u_n(x)\}$ are generally complex. Since $M$, $C$, and $K$ are real, if $\{\lambda_n,u_n\}$ is an eigensolution, so is its complex conjugate, which we denote as $\{\lambda_{-n},u_{-n}\}=\{\overline{\lambda_n},\overline{u_n}\}$ where the overbar denotes complex conjugation. (In the sequel, when we say complex eigenfunctions, we mean eigenfunctions that cannot be made entirely real.) The complete set of eigensolutions is ordered such that $n=\pm 1,\pm 2,\pm 3\ldots$. If $\{\lambda_n,u_n\}$ is an

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Sept. 15, 1999; final revision, Mar. 12, 2000. Associate Technical Editor: N. C. Perkins. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

 **Copyright © 2000 by ASME** 

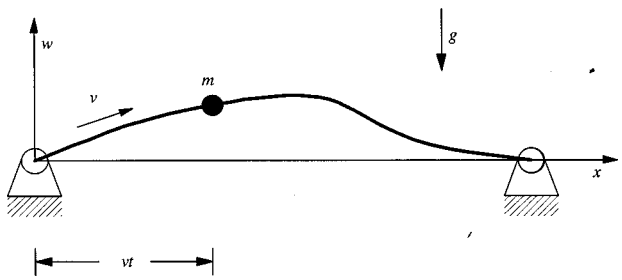



**Fig. 1 Schematic of a damped, axially moving string subject to a concentrated moving mass**

eigensolution of (2), then $\{\overline{\lambda_n}, u_n^*\}$ is an eigensolution of the adjoint eigenvalue problem. We assume that there are no vanishing eigenvalues and normalize the eigenfunctions such that

$$\lambda_n \lambda_m \langle M u_n, u_m^* \rangle - \langle K u_n, u_m^* \rangle = \delta_{nm} \tag{3}$$

where $\delta_{nm}$ is the Kronecker delta. The orthogonality

$$\lambda_n \lambda_m \langle C u_n, u_m^* \rangle + (\lambda_n + \lambda_m) \langle K u_n, u_m^* \rangle = -\lambda_n \delta_{nm} \tag{4}$$

also holds. For positive definite gyroscopic systems, the eigenvalues are imaginary and the eigenfunctions and adjoint eigenfunctions are the same up to a multiplicative constant. We normalize the eigenfunctions such that $u_n = -u_n^*$.

We assume that the eigenfunctions are complete in the sense that an eigenfunction expansion can simultaneously represent an arbitrary displacement and velocity at any instant of time. This is true for most problems of practical interest ([2]). For complete eigenfunctions, the general solution of (1) can be expressed as an eigenfunction expansion

$$w = \sum_{n=\pm 1, \pm 2, \ldots}^{\pm \infty} d_n(t) u_n(x). \tag{5}$$

Since the eigenfunctions are not linearly independent (in the finite dimensional case, for instance, there are twice as many eigenfunctions as spatial dimensions), Eqs. (1) and (5) do not uniquely determine $d_n(t)$. We therefore constrain the $d_n(t)$ such that

$$w_{,t} = \sum_{n=\pm 1}^{\pm \infty} d_{n,t} u_n = \sum_{n=\pm 1}^{\pm \infty} \lambda_n d_n u_n \tag{6}$$

holds. Upon substituting (5) into (1), taking the inner product with $\overline{\lambda_n} u_n^*$, differentiating (6) with respect to time and utilizing (3) and (4), the equation of motion decouples into the readily solved equation

$$d_{n,t} - \lambda_n d_n = \lambda_n \langle f, u_n^* \rangle \tag{7}$$

which uniquely determines $d_n$.

As a specific example of a system of form (1), consider the damped, axially moving string shown in Fig. 1 without the concentrated mass. A string of unit tension moves axially between eyelets at $x=0$ and $x=1$ with uniform axial velocity $v$. The string has viscous damping of magnitude $2\pi c$. The transverse displacement of the string, $w(x,t)$, is measured vertically upwards and vanishes at $x=0$ and $x=1$. Gravity acts vertically downwards. For this system

$$M = I, \quad C = 2\pi c + 2v \frac{\partial}{\partial x}, \quad K = (v^2 - 1) \frac{\partial^2}{\partial x^2}, \quad f = -g, \tag{8}$$

where $I$ is the identity operator ([1]). The system is positive definite and gyroscopic when $c=0$ and $v<1$; the system is self-adjoint when $c=v=0$.

The linear dependence of the eigensolutions can be demonstrated by calculating the determinant of the matrix

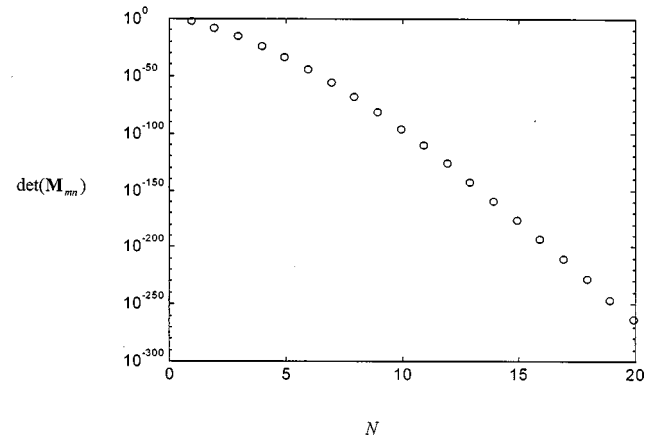



**Fig. 2 The determinant of $M_{mn} = \langle M u_n, u_m^* \rangle$ as a function of $N$, the number of trial functions used**

$$\mathbf{M}_{mn} = \langle M u_n, u_m^* \rangle \tag{9}$$

for $n$ and $m = \pm 1, \ldots, \pm N$ using the eigenfunctions of undamped, axially moving string,

$$\lambda_n = i n \pi (1 - v^2),$$

$$u_n = -u_n^* = \sin(n\pi x) \exp(i n \pi v x) / n \pi \sqrt{1 - v^2}. \tag{10}$$

Figure 2 shows $\det(\mathbf{M}_{mn})$ as a function of $N$ for $v = 0.25$. As $N$ increases, the determinant steadily becomes smaller and smaller due to the linear dependence of the eigenfunctions. At some finite $N$, the matrix becomes numerically singular and noninvertible. If the adjoint eigenfunction $u_m^*$ in (9) is replaced by eigenfunction $u_m$, as is done in Galerkin's method, the singularity problem remains, and the resulting equation of motion derived by Galerkin's method becomes numerically insoluble.

## 3 The Behavior of the Constraint

The crucial step in the eigenfunction analysis described in Section 2 that distinguishes it from self-adjoint analyses is the constraint (6). Since the $d_n(t)$ satisfy (7), the constraint (6) is equivalent to and can be evaluated directly by evaluating

$$\sum_{n=\pm 1}^{\pm \infty} (d_{n,t} - \lambda_n d_n) u_n = \sum_{n=\pm 1}^{\pm \infty} \lambda_n \langle f, u_n^* \rangle u_n = 0 \tag{11}$$

at any instant of time. From (11), it can be seen that for self-adjoint systems with purely imaginary eigenvalues and real eigenfunctions, the terms in the finite summation

$$S_N(x) = \sum_{n=\pm 1}^{\pm N} (d_{n,t} - \lambda_n d_n) u_n = \sum_{n=\pm 1}^{\pm N} \lambda_n \langle f, u_n^* \rangle u_n \tag{12}$$

cancel out in complex conjugate pairs. For complex eigenvalues and eigenfunctions, $S_\infty \to 0$ in the $L_2$ norm, but not in the pointwise sense. This is important to the moving mass problem since the solution of this problem depends crucially on the motion of a particular point in the span (e.g., the location of the concentrated mass).

Figure 3 shows $S_N$ for the uniform load $f=1$ for various $N$ while Fig. 4 shows $S_N$ for a concentrated load $f = \delta(x - 0.3)$ for various $N$, using the undamped, axially moving string eigensolutions (10). In both cases, $S_N \to 0$ as $N \to \infty$. However, while the uniform load appears to converge uniformly, the concentrated load is distinguished by large rapid oscillations at the point of application of the concentrated load. Because of these oscillations, the spatial derivative, $\partial S_N / \partial_x = 0$, is expected to be poorly satis-

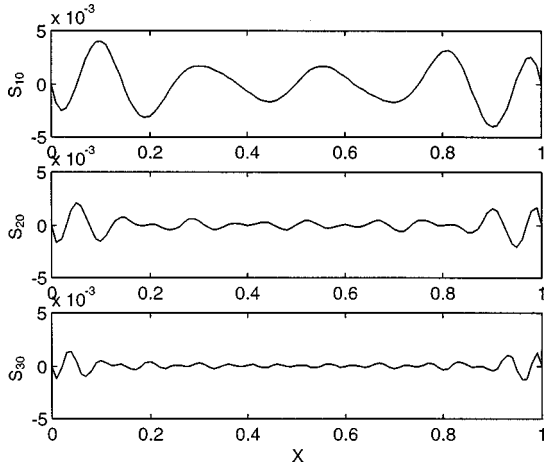



**Fig. 3 Plots of $S_N(x)$ for the uniform load $f=1$ for $N=10$, 20, and 30 using the eigenfunctions of the undamped string**

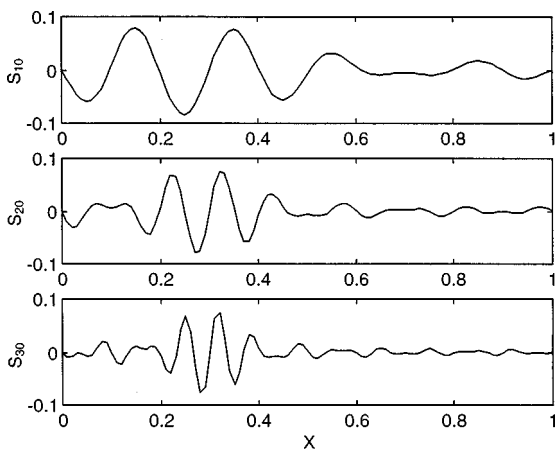



**Fig. 4 Plots of $S_N(x)$ for the uniform load $f=\delta(x-0.3)$ for $N=10$, 20, and 30 using the eigenfunctions of the undamped string**

fied for finite $N$. This bears directly on the moving mass problem since the constraint must be differentiated in order to reduce the order of the differential equation.

Despite these oscillations, the value of $S_N$ at the point of application of the concentrated load vanishes. This is always true for positive definite gyroscopic systems, since the terms of $S_N$ cancel out in pairs, and can be demonstrated for the damped axially moving string (8) (see the Appendix). It remains an open question whether this is true for general non-self-adjoint systems. *The fact that $S_N$ vanishes at the point of application of a concentrated load for positive definite gyroscopic systems and for the damped, axially moving string is essential to the following solution of the moving mass problem.*

## 4 The Moving Mass Problem

We now turn our attention to the moving mass problem shown in Fig. 1 and governed by the dimensionless equation

$$w_{,tt}+2\pi c w_{,t}+2v w_{,tx}+(v^2-1)w_{,xx} =[-mg-m(w_{,tt}+2v w_{,tx}+v^2 w_{,xx})|_{x=vt}]\delta(x-vt). \tag{13}$$

Equation (13) describes the coupled dynamics of a damped, axially moving string translating at speed $v$ and a concentrated mass of mass $m$ that translates with the string, entering the span at $t=0$ and exiting the span at $t=1/v$. For convenience, the displacement of the string due to its own weight has been omitted, since that solution can be found directly using the method given in Section 2 and superimposed on the moving mass solution. We consider two alternative solution methods.

**Galerkin's Method Using Real Trial Functions.** In Galerkin's method, a set of real, linearly independent, complete trial functions, $p_n(x)$, are selected and the response, $w$, is approximated by

$$w=\sum_{n=1}^{N} q_n(t)p_n(x) \tag{14}$$

for some finite $N$. Expansion (14) is substituted into (13) and the inner product with each $p_m$ is taken. This produces to a set of $N$ second order, ordinary differential equations in time of the form

$$\mathbf{A}\mathbf{q}_{,tt}+\mathbf{B}\mathbf{q}_{,t}+\mathbf{D}\mathbf{q}=\mathbf{f} \tag{15}$$

where

$$\mathbf{q}=[q_1 \;\; q_2 \;\; \ldots \;\; q_N]^T \tag{16}$$

$$\mathbf{A}_{mn}=\langle p_n,p_m\rangle+m p_{no}p_{mo} \tag{17}$$

$$\mathbf{B}_{mn}=2c\pi\langle p_n,p_m\rangle+2v\langle(p_{n,x}),p_m\rangle+2v m p_{no,x}p_{mo} \tag{18}$$

$$\mathbf{D}_{mn}=(v^2-1)\langle(p_{n,xx}),p_m\rangle+v^2 m p_{no,xx}p_{mo} \tag{19}$$

$$\mathbf{f}_m=-mgp_{mo}. \tag{20}$$

Subscript $o$ to indicates evaluation at $x=vt$. Equation (15) is numerically integrated in time to find the coefficients $q_n(t)$. For (13), we choose $p_n=\sin(n\pi x)$, the eigenfunctions of the stationary string.

If complex eigenfunctions are used as trial functions with Galerkin's method, the linear dependence of the eigenfunctions renders this approach unusable for $N$ greater than about 10 or so.

**Eigenfunction Expansion Method.** In the eigenfunction expansion method developed here, the response is written as

$$w=\sum_{n=\pm 1}^{\pm N} d_n(t)u_n(x). \tag{21}$$

Substitution of (21) into (13), taking the inner product with $\overline{\lambda}_m u_m^*$, differentiating the constraint (6) in time and applying it on the left-hand side of (13), gives the complex valued second-order ordinary differential equations

$$d_{m,t}-\lambda_m d_m=-mg\lambda_m\overline{u}_{mo}^*-m\lambda_m\sum_{n=\pm 1}^{\pm N}\{d_{n,tt}u_{no}\overline{u}_{mo}^* +2v d_{n,t}u_{no,x}\overline{u}_{mo}^*+v^2 d_n u_{no,xx}\overline{u}_{mo}^*\}. \tag{22}$$

In this form, these equations are not resolvable since the leading order coefficient matrix $m\lambda_m u_{no}\overline{u}_{mo}^*$ is singular for small $t$. Since the constraint (6) is valid for all $x$, including $x=vt$, one could simply differentiate it in time, evaluate at $x=vt$ and substitute into (22) to reduce the second-order ordinary differential equation to first order. However, numerical experiments indicate that this technique does not work because the derivative of (6) applied at $x=vt$ is not valid for finite $N$.

Instead, we take advantage of the fact the constraint (6) *holds identically at the point of application* of the concentrated load and differentiate (6) along the trajectory of the concentration load, i.e.,

$$\left(\frac{\partial}{\partial t}+v\frac{\partial}{\partial x}\right)\sum_{n=\pm 1}^{\pm N}(d_{n,t}-\lambda_n d_n)u_n$$
$$=\sum_{n=\pm 1}^{\pm N}(d_{n,tt}u_n-\lambda_n d_{n,t}u_n+v d_{n,t}u_{n,x}-v\lambda_n d_n u_{n,x})=0. \tag{23}$$

Since the constraint vanishes along this trajectory for finite $N$, (23) holds identically at $x=vt$. Consequently, substitution of (23) evaluated at $x=vt$ into (22) gives the $2N$ first order, complex ordinary differential equations

$$d_{m,t}-\lambda_m d_m=-mg\lambda_m\overline{u}^*_{mo}-m\lambda_m\sum_{n=\pm 1}^{\pm N}\{(\lambda_n d_{n,t}u_{no}-v d_{n,t}u_{no,x}+v\lambda_n d_n u_{no,x}+2v d_{n,t}u_{no,x}+v^2 d_n u_{no,xx})\overline{u}^*_{mo}\}. \tag{24}$$

Numerical integration of Eq. (24) gives the eigenfunction expansion solution considered here.

## 5 Convergence Results

Figures 5–8 compare the convergence results for Galerkin's method with real trial functions and the complex eigenfunction expansion method for a variety of cases. In each case, the displacement of the concentrated mass at the instance of time when it is halfway across the span, $t=1/2v$, is shown as a function of $N$, the number of terms taken.

Figures 5 and 6 show solutions for the undamped string. In Fig. 5, $c=0$, $v=0.25$, $m=g=0.375$. In Fig. 6, $c=0$, $v=0.5$, $m=g=0.125$. Figures 7 and 8 show solutions for the damped string. In Fig. 7, $c=0.3$, $v=0.25$, $m=g=0.375$. In Fig. 8, $c=0.3$, $v=0.5$, $m=g=0.125$. Also shown in Figs. 5 and 6 and presented by a dash line is the solution calculated using the integro-differential technique of Wickert and Mote [12]. This technique is only applicable to the undamped string and cannot be used for Figs. 7 and 8. It does, however, give an alternative estimate of the accuracy of the Galerkin and eigenfunction expansion results.

The results of each method appear to converge to the same answer. The results also indicate that the proposed method is competitive with Galerkin's method with real trial functions in terms of speed of convergence and accuracy.

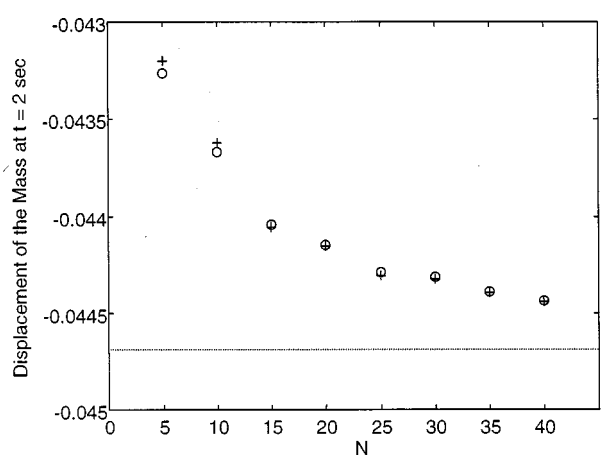



**Fig. 5 The displacement of the concentrated mass at the midpoint of the span as a function of *N* for Galerkin's method using real trial functions (plus) and the proposed eigenfunction expansion technique (circle). The dashed line is the result obtained using the method of Wickert and Mote [12]. $c=0$, $v=0.25$, $t=2$, and $m=g=0.375$.**

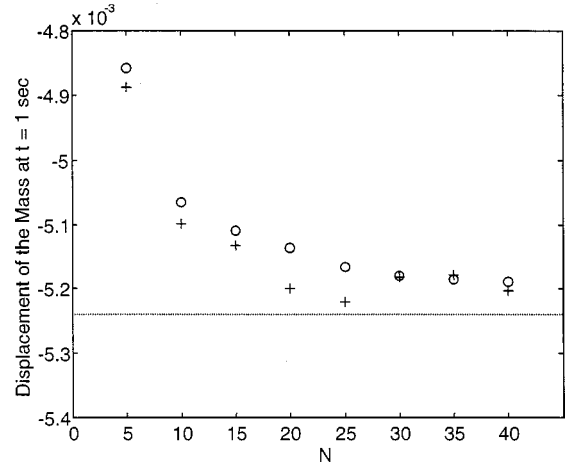



**Fig. 6 Same as Fig. 5 except $t=1$, $v=0.5$, $m=g=0.125$. Galerkin's method using real trial functions (plus) and the proposed eigenfunction expansion technique (circle). The dashed line is the result obtained using the method of Wickert and Mote [12].**

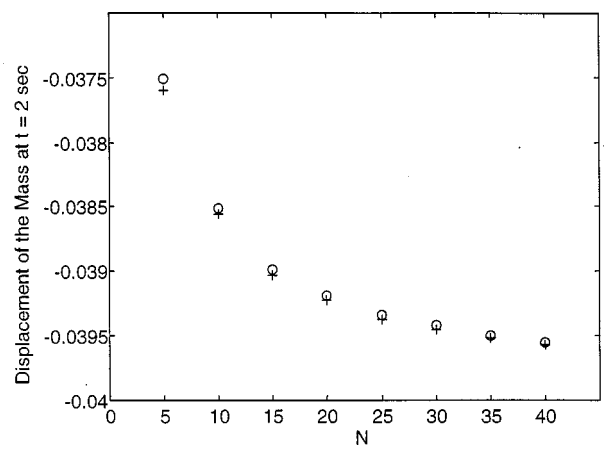



**Fig. 7 The displacement of the concentrated mass at the midpoint of the span as a function of *N* for Galerkin's method using real trial functions (plus) and the proposed eigenfunction expansion technique (circle). $c=0.3$, $v=0.25$, and $m=g=0.375$.**

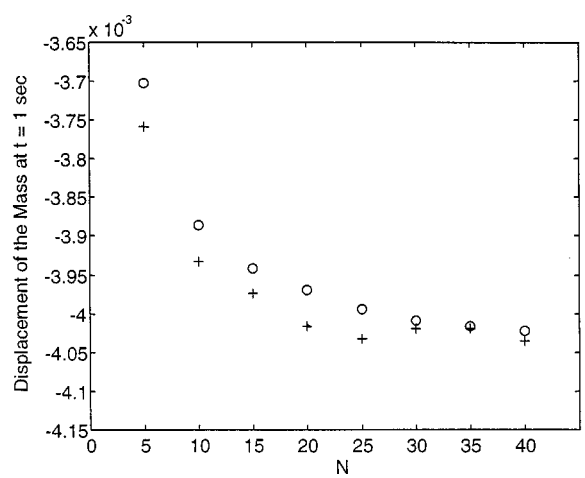



**Fig. 8 Same as Fig. 7 except $t=1$, $v=0.5$, $m=g=0.125$. Galerkin's method using real trial functions (plus) and the proposed eigenfunction expansion technique (circle).**

## 6 Discussion

Unlike alternative solution techniques (e.g., [12]), the eigenfunction solution method can be applied to any moving mass problem, including the cases when the moving mass is attached to the transporting medium by a spring and damper and when modeling of the transporting medium includes bending stiffness. This may be relevant to the analysis of disk drive designs since existing results have only studied the interaction of stationary recording heads with the rotating memory disks or moving recording heads with stationary memory disks [10,13]. The method can also be extending to arbitrary motion of the moving mass as long as the derivative of the trajectory of the mass is piecewise continuous in the $x-t$ coordinate plane.

The unresolvable Eq. (22) can be derived in a number of well-known methods. For instance, if one recasts the second order equation of motion (1) into first order, state space form ([3,4]), Galerkin analysis will result in Eq. (22). Pesterev and Bergman [9] derive (22) using an integral equation technique. However, the constraint (6) is only implicitly assumed in both of these derivations. As a result, previous researchers have not used (6) to reduce the unresolvable second-order ordinary differential equation (22) to the first-order (24). This step, together with the insight to differentiate $S_N$ along the trajectory of the concentrated mass in the $x-t$ coordinate space, are the primary contributions of this paper.

## 7 Conclusions

A new solution technique is developed for solving the moving mass problem for nonconservative, linear, distributed parameter systems using complex eigenfunction expansions. Traditional Galerkin analysis of this problem using complex eigenfunctions fails in the limit of large numbers of trial functions because complex eigenfunctions are not linearly independent. This linear dependence problem is circumvented in the method proposed here by applying a modal constraint on the velocity of the distributed parameter system [3]. This constraint is valid for all complete sets of eigenfunctions but must be applied with care for finite dimensional approximations of concentrated loads such as found in the moving mass problem. Numerical results indicate that the proposed method is competitive with Galerkin's method using real trial functions in terms of accuracy and rate of convergence.

## Appendix

For the damped, axially moving string with a point force, it can be shown $S_N=0$ exactly at the point force. The operators of the system are given in (8) and the eigensolutions and the adjoint eigensolutions are of the form

$$\begin{aligned} \lambda_n &= a+ib_n \\ \overline{\lambda}_n &= a-ib_n \\ u_n &= H_n \exp(v\lambda_n x/1-v^2)\sin n\pi x \\ u_n^* &= H_n^* \exp(-v\overline{\lambda}_n x/1-v^2)\sin n\pi x \end{aligned} \tag{25}$$

where $n=\pm 1,\pm 2\ldots$, $H_n$ and $H_n^*$ are normalization constants to be determined, and

$$a=-(1-v^2)c\pi, \quad b_n=\mathrm{sgn}(n)(1-v^2)\pi\sqrt{n^2-c^2}. \tag{26}$$

Substituting (25) into (3) gives

$$H_n\overline{H}_n^*=\frac{2}{\Delta_1+i\Delta_2}=\frac{2(\Delta_1-i\Delta_2)}{\Delta_1^2+\Delta_2^2} \quad \text{for } n>0, \tag{27}$$

$$H_n\overline{H}_n^*=\frac{2}{\Delta_1-i\Delta_2}=\frac{2(\Delta_1+i\Delta_2)}{\Delta_1^2+\Delta_2^2} \quad \text{for } n<0. \tag{28}$$

where $\Delta_1=(a^2-b_n^2-(1-v^2)^2n^2\pi^2)/(1-v^2)$ and $\Delta_2=2ab_n/(1-v^2)$. Equations (25), (27), and (28) give

$$\lambda_n u_n(x_0)\overline{u}_n^*(x_0)=\frac{2}{\Delta_1^2+\Delta_2^2}[(\Delta_1 a+\Delta_2 b_n) + i(-\Delta_2 a+\Delta_1 b_n)]\sin^2 n\pi x \quad \text{for } n>0, \tag{29}$$

$$\lambda_n u_n(x_0)\overline{u}_n^*(x_0)=\frac{2}{\Delta_1^2+\Delta_2^2}[(\Delta_1 a+\Delta_2 b_n) + i(\Delta_2 a-\Delta_1 b_n)]\sin^2 n\pi x \quad \text{for } n<0, \tag{30}$$

Therefore (12) becomes

$$\begin{aligned} S_N(x_0) &= \sum_{n=\pm 1}^{\pm N} \lambda_n\langle\delta(x-x_0),u_n^*(x)\rangle u_n(x_0) \\ &= \sum_{n=\pm 1}^{\pm N} \lambda_n\overline{u}_n^*(x_0)u_n(x_0) \\ &= \sum_{n=1}^{N} \frac{4}{\Delta_1^2+\Delta_2^2}(\Delta_1 a+\Delta_2 b_n)\sin^2 n\pi x=0 \end{aligned} \tag{31}$$

since $(\Delta_1 a+\Delta_2 b_n)=0$ when expanded.

## References


[1] Wickert, J. A., and Mote, Jr., C. D., 1991, ''Response and Discretization Methods for Axially Moving Materials,'' Appl. Mech. Rev., **44**, pp. S279–S284.

[2] Hryniv, R. O., Lancaster, P., and Renshaw, A. A., 1999, ''A Stability Criterion for Parameter Dependent Gyroscopic Systems,'' ASME J. Appl. Mech., **66**, pp. 660–664.

[3] Renshaw, A. A., 1997, ''Modal Decoupling of Systems Described by Three Linear Operators,'' ASME J. Appl. Mech., **64**, pp. 238–240.

[4] Wickert, J. A., and Mote, Jr., C. D., 1990, ''Classical Vibration Analysis of Axially Moving Continua,'' ASME J. Appl. Mech., **57**, pp. 738–744.

[5] Olsson, M., 1991, ''On the Fundamental Moving Load Problem,'' J. Sound Vib., **145**, pp. 299–307.

[6] Sadiku, S., and Leipholz, H. H. E., 1987, ''On the Dynamics of Elastic Systems With Moving Concentrated Masses,'' Ing.-Arch., **57**, pp. 223–242.

[7] Lee, U., 1996, ''Revisiting the Moving Mass Problem: Onset of Separation Between the Mass and Beam,'' ASME J. Vibr. Acoust., **118**, pp. 516–521.

[8] Nelson, H. D., and Conover, R. A., 1971, ''Dynamic Stability of a Beam Carrying Moving Masses,'' ASME J. Appl. Mech., **93**, pp. 1003–1006.

[9] Pesterev, A. V., and Bergman, L. A., 1998, ''Response of a Nonconservative Continuous System to a Moving Concentrated Load,'' ASME J. Appl. Mech., **65**, pp. 436–444.

[10] Shen, I. Y., 1993, ''Response of a Stationary, Damped, Circular Plate Under a Rotating Slider Bearing System,'' ASME J. Vibr. Acoust., **115**, pp. 65–69.

[11] Stanisic, M. M., 1985, ''On a New Theory of the Dynamic Behavior of the Structures Carrying Moving Masses,'' Ing.-Arch., **55**, pp. 176–185.

[12] Wickert, J. A., and Mote, Jr., C. D., 1991, ''Traveling Load Response of an Axially Moving String,'' J. Sound Vib., **149**, pp. 267–284.

[13] Iwan, W. D., and Moeller, T. L., 1976, ''The Stability of a Spinning Elastic Disk With a Transverse Load System,'' ASME J. Appl. Mech., **43**, pp. 485–490.

Journal of Applied Mechanics

# Brief Notes

A Brief Note is a short paper that presents a specific solution of technical interest in mechanics but which does not necessarily contain new general methods or results. A Brief Note should not exceed 1500 words *or equivalent* (a typical one-column figure or table is equivalent to 250 words; a one line equation to 30 words). Brief Notes will be subject to the usual review procedures prior to publication. After approval such Notes will be published as soon as possible. The Notes should be submitted to the Technical Editor of the JOURNAL OF APPLIED MECHANICS. Discussions on the Brief Notes should be addressed to the Editorial Department, ASME, United Engineering Center, Three Park Avenue, New York, NY 10016-5990, or to the Technical Editor of the JOURNAL OF APPLIED MECHANICS. Discussions on Brief Notes appearing in this issue will be accepted until two months after publication. Readers who need more time to prepare a Discussion should request an extension of the deadline from the Editorial Department.

## On the Relation Between the *L*-Integral and the Bueckner Work-Conjugate Integral

**J. P. Shi, X. H. Liu, and J. M. Li**
Department of Engineering Mechanics, Xi'an University of Technology, 710048 Xi'an, China

*A simple but inherent relation between the L-integral and the Bueckner work conjugate integral is deduced for crack problem in isotropic, anisotropic, and dissimilar materials, respectively. It is proved the L-integral, from the mathematical point of view as well as in principle, is arising from the Betti's reciprocal theorem.*
[S0021-8936(00)00103-3]

### 1 Introduction

Knowles and Sternberg [1] have shown that the $L$-integral is given by

$$L=\oint_{\Gamma} e_{3ij}(Wx_jn_i-T_iu_j-T_ku_{k,i}x_j)ds \tag{1}$$

where $\Gamma$ is a closed contour in the $x_1=x$, $x_2=y$ plane surrounding a whole crack; $W$ is the strain energy density, and $T_i$ is the traction acting on the outer side of the $\Gamma$. The characteristics of the $L$-integral and the $J_k$-integral are different. It can be proven that the $L$-integral is a path-independent integral. We can also verify that the $L$-integral is independent of the selection of the coordinate system.

The Bueckner work-conjugate integral ([2]) was derived from the well-known Betti's reciprocal theorem, which could be formulated as follows:

$$I_{\Gamma}=\int_{\Gamma}(u_i^{(I)}\sigma_{ij}^{(II)}-u_i^{(II)}\sigma_{ij}^{(I)})n_jds \quad (i,j=x,y) \tag{2}$$

where the superscripts ($I$) and ($II$) refer to two possible displacement-stress fields which satisfy the traction-free conditions on the crack faces. The property of the path-independent integral is proved by Bueckner using Betti's work reciprocal theorem, that is $I_{\Gamma}=I_C$.

### 2 Proof

For homogeneous isotropic materials, assume that the first possible displacement-stress field is induced by the following complex potentials $\varphi(z)$ and $\psi(z)$:

$$\begin{aligned}\varphi'(z)&=\frac{1}{\sqrt{z^2-a^2}}\sum_{k=1}^{+\infty}E_kz^k+\sum_{k=1}^{+\infty}F_kz^{k-1}\\ \psi'(z)&=\frac{1}{\sqrt{z^2-a^2}}\sum_{k=1}^{+\infty}E_kz^k-\sum_{k=1}^{+\infty}F_kz^{k-1}\end{aligned} \tag{3}$$

where $E_k$ and $F_k$ are complex coefficients which can be defined by remote conditions. Introduce a supplemental displacement-stress field defined by the following complex potentials $\varphi^{(II)}(z)$ and $\psi^{(II)}(z)$:

$$\begin{aligned}\varphi^{(II)}(z)&=-iz\varphi'(z)\\ \psi^{(II)}(z)&=-iz\psi'(z)+2i\bar{z}\varphi'(z).\end{aligned} \tag{4}$$

The corresponding displacement and stress components are derived as follows:

$$\begin{aligned}&u_i^{(II)}=yu_{i,x}-xu_{i,y}\\ &\sigma_{ij}^{(II)}=y\sigma_{ij,x}-x\sigma_{ij,y}+\frac{1}{2}\int\sigma_{ij,x}dy-\frac{1}{2}\int\sigma_{ij,y}dx \quad (i,j=1,2)\end{aligned} \tag{5}$$

where $u_i$ and $\sigma_{ij}$, as the ($I$) field, are the displacement and stress components induced by Eq. (3). It can be examined that the stress $\sigma_{ij}^{(II)}$ satisfy the traction-free conditions on the crack faces. Substituting Eq. (5) and $u_i$ and $\sigma_{ij}$ into Eq. (2), we obtain

$$\begin{aligned}I_{\Gamma}=\int\Big[&u_i\sigma_{ij,x}y-u_i\sigma_{ij,y}x-u_{i,x}\sigma_{ij}y+u_{i,y}\sigma_{ij}x\\ &+\frac{1}{2}u_i\int\sigma_{ij,x}dy-\frac{1}{2}u_i\int\sigma_{ij,y}dx\Big]n_jds\end{aligned} \tag{6}$$

where $dx=-n_2ds$, $dy=n_1ds$.

Now, the $I_{\Gamma}-2L$ is examined. Utilizing the equilibrium conditions in plane problems and noting the integral terms $T_iu_i$, $\sigma_{ij}u_i$ and $u_i\int\sigma_{ij,x}dy$ have no contribution for $I_{\Gamma}-2L$ when $\rho\to 0$ (at the near of crack tip). Thus, $I_{\Gamma}-2L$ is equal to zero. We obtain

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, June 6, 1999; final revision, January 21, 2000. Associate Technical Editor: J. T. Ju.

 **Copyright © 2000 by ASME** 

$$I_\Gamma = 2L. \tag{7}$$

Equation (7) shows that between the $L$-integral and the Bueckner work conjugate integral there is a simple but inherent relationship. We need not to know the obvious function expressions of complex potentials for the crack beforehand, but the traction-free conditions must be satisfied.

If there are two displacement-stress fields, namely, $u_i$, $\sigma_{ij}$ and $u_i^{(II)}$, $\sigma_{ij}^{(II)}$, using the Betti's reciprocal theorem to the region bounded by crack borders, one can divide the contour $C$ into $C_L$, $C_R$ and $C_+, C_-$, where $C_L, C_R$ are circles around the left and the right crack tips and $C_+, C_-$ are the straight line along the upper and lower crack faces, respectively. Because the stresses are free on the crack faces, then $I_\Gamma = I_{CL} + I_{CR}$. In this case, if integrals $I_{CL}$ and $I_{CR}$ can be evaluated for some displacement-stress fields, then the path-independent integral $I_\Gamma$ can also be defined and evaluated. If the displacement-stress fields are defined by Eqs. (3) and (4), we can deduce the following relationship between the $L$-integral and the stress intensity factors:

$$L = -\frac{3(\kappa-1)a}{4\mu}(K_{1L}K_{2L} + K_{1R}K_{2R}) \tag{8}$$

where $K_L = K_{1L} + iK_{2L}$, $K_R = K_{1R} + iK_{2R}$ are stress intensity factors at the left and the right crack tips, respectively. $\kappa$ and $\mu$ are elastic constants.

## 3 Discussion

The complex potentials of the center crack, $\varphi_1(z), \omega_1(z), \varphi_2(z), \omega_2(z)$, were obtained by Chen and Shi [3] by using the same method obtained the eigenfunction expansion form by Rice [4] in interfacial cracks for dissimilar material. The stress and displacement fields that are obtained from these complex potentials satisfy the traction-free conditions on the crack faces and the continuous condition along the entire interface.

A supplemental displacement-stress field defined by the complex potentials $\varphi_1^{(II)}(z), \omega_1^{(II)}(z), \varphi_2^{(II)}(z), \omega_2^{(II)}$ is introduced. The relations between $\varphi_1^{(II)}(z), \omega_1^{(II)}(z), \varphi_2^{(II)}(z), \omega_2^{(II)}$ and $\varphi_1(z), \omega_1(z), \varphi_2(z), \omega_2(z)$ are analogous to Eq. (4).

In a similar manner, the displacement and stress of the ($II$) field are presented in Eq. (5). They satisfy the traction-free conditions on the crack faces also. The corresponding displacement and stress components will be substituted into Eq. (6). Note that the curve $\Gamma$ can be divided into two sections: curve $\Gamma_1$ of the upper plane and the curve $\Gamma_2$ of the below plane. The deductions of Eqs. (6) to (7) relate to the equilibrium equations in plane problems and traction-free conditions only, but don't involve the material parameter. The process is the same as the above homogeneous isotropic material. Finally, we still obtain Eq. (7) in the interface crack, that is $I_\Gamma = 2L$.

Between $L$-integral and stress intensity factors there is the following relation:

$$L = -\left(\frac{\kappa_1 - 1}{\mu_1} + \frac{\kappa_2 - 1}{\mu_2}\right)\frac{3(K_{1L}K_{2L} + K_{1R}K_{2R})}{8\cosh^2(\pi\varepsilon)}a \tag{9}$$

where $\varepsilon$ is ''oscillation index'' and $K_L = K_{1L} + iK_{2L}$, $K_R = K_{1R} + iK_{2R}$ are complex stress intensity factors at the left and the right crack-tips, respectively. They cannot be separated into the pure I model and II model; $\kappa_1, \mu_1$ and $\kappa_2, \mu_2$ stand for the material parameters of the upper and lower plane.

For anisotropy material, the Lekhenitski complex potential theory needs to be used ([5]). According to the need of the Bueckner work conjugate integral, the subsidiary stress-displacement fields, which represents ($II$) field, are

$$\varphi^{(II)}(z_1) = -iz_1\varphi'(z_1)$$
$$\psi^{(II)}(z_2) = -iz_2\psi'(z_2) + 2i\overline{z_2}\varphi'(z_1). \tag{10}$$

The stress fields caused by Eq. (10) satisfy the traction-free conditions. The relation between the ($II$) field and a physical stress field are analogous to Eq. (5).

By proceeding in the same manner as the isotropic case from Eq. (6) to Eq. (7), we draw a conclusion $I_\Gamma = 2L$.

It can be seen that a simple but inherent relation between the $L$-integral and the Bueckner work conjugate integral is right all along, although the characteristic of material is more complex than isotropic and the complex potentials in these two cases are more different with in isotropic.

## 4 Conclusions

Using the Bueckner work conjugate integral through introducing a special subsidiary stress-displacement field, one can render the $L$-integral. The relation between $L$-integral and the Bueckner work conjugate integral seems independent of the stress oscillatory singularities on the interface crack tips and the eigenroot in the anisotropy. It is found that the $L$-integral, from the mathematical point of view as well as in principle, is arising from the Betti's reciprocal theorem. This means that the Bueckner work conjugate integral is a more general path-independent integral than others are. Using the Bueckner integral through choosing a different subsidiary stress-displacement field could render any other path-independent integrals.

## References


[1] Knowles, J. K., and Sternberg, E., 1972, ''On a Class of Conservation Laws in Linearized and Finite Elastostatics,'' Arch. Ration. Mech. Anal., **44**, pp. 187–211.

[2] Bueckner, H. F., 1973, *Mechanics of Fracture: Methods of Analysis and Solution of Crack Problems*, G. C. Sih, ed., Noordhoff, Leyden, pp. 239–314.

[3] Chen, Y. H., and Shi, J. P., 1998, ''On the Relation Between the *M*-integral and the Bueckner Work-Conjugate Integral,'' Acta Mech. Sin., **30**, pp. 495–502 (in Chinese).

[4] Rice, J. R., 1988, ''Elastic Fracture Mechanics Concepts for Interface Cracks,'' J. Appl. Mech., **55**, pp. 98–103.

[5] Sih, G. C., and Chen, E. P., 1981, ''Cracks in Composite Materials,'' Mech. Fracture, **6**, pp. 1–99.


# A Note on the Driving Traction Acting on a Propagating Interface: Adiabatic and Non-Adiabatic Processes of a Continuum


**R. Abeyaratne**
Mem. ASME, Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139

**J. K. Knowles**
Fellow ASME, Division of Engineering and Applied Sciences, California Institute of Technology, Pasadena, CA 91125



*An expression for the driving traction on an interface is derived for an arbitrary continuum undergoing an arbitrary thermomechanical process which may or may not be adiabatic.*



[S0021-8936(00)00403-7]

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, July 30, 1997; final revision, Aug. 5, 1997. Associate Technical Editor: L. T. Wheeler.

## 1 Summary

In this note we derive an expression for the ''driving traction,'' or Eshelby force ([1]), acting on a propagating interface in a continuum. The interfaces that we have in mind might represent, for example, a shock wave or a boundary between two phases of a material, and the thermomechanical processes which the continuum is permitted to undergo may or may not be adiabatic. From the perspective of irreversible thermodynamics, the driving traction corresponds to a ''thermodynamic affinity''; see, for example, [2–4]. It plays a central role in modeling the kinetics of phase transformations by characterizing the rate of propagation of phase boundaries (e.g., see [5–8]).

The derivation sketched below makes no assumptions about the constitutive law for the continuum under consideration. When specialized to a thermoelastic material, the expression for the driving traction obtained here has certain similarities with the Legendre transform of the Helmholtz free-energy $\hat{\psi}(\mathbf{F},\theta)$ with respect to *both* the deformation gradient tensor $\mathbf{F}$ and the temperature $\theta$, as well as with the Legendre transform of the internal energy $\hat{\varepsilon}(\mathbf{F},\eta)$ with respect to $\mathbf{F}$ and the specific entropy $\eta$.

The result derived here generalizes an earlier one which had been established for non-adiabatic processes ([9,10,5]). This former characterization of driving traction was not valid in adiabatic processes, and therefore did not, in particular, apply to shock waves in classical gas dynamics or to impact-induced rapidly moving phase boundaries in solids. A one-dimensional version of the present result was obtained in [11].

## 2 Momentum and Energy

Consider a body which occupies a region $R$ in a reference configuration. Let $\mathbf{x}\in R$ denote the position of a particle in this configuration and let $t$ denote time. Consider a thermomechanical process of this body on some time interval $[t_1,t_2]$ which is characterized by the motion $\mathbf{y}(\mathbf{x},t)$, body force per unit mass $\mathbf{b}(\mathbf{x},t)$, Piola-Kirchhoff stress $\boldsymbol{\sigma}(\mathbf{x},t)$, heat flux $\mathbf{q}(\mathbf{x},t)$, heat supply $r(\mathbf{x},t)$ and internal energy per unit mass $\varepsilon(\mathbf{x},t)$. Suppose that during this process $\mathbf{y}$ is continuous with piecewise continuous first and second derivatives on $R\times[t_1,t_2]$; $\mathbf{b}(\cdot,t)$ and $r(\cdot,t)$ are continuous on $R$; $\boldsymbol{\sigma}(\cdot,t)$ and $\mathbf{q}(\cdot,t)$ are piecewise continuous with piecewise continuous gradient on $R$; and $\varepsilon$ is piecewise continuous with piecewise continuous first derivatives on $R\times[t_1,t_2]$. During this process, the usual balance laws of linear and angular momentum and the first law of thermodynamics require that for any subregion $D$,

$$\int_{\partial D}\boldsymbol{\sigma}\mathbf{n}\,dA+\int_D\rho\mathbf{b}\,dV=\frac{d}{dt}\int_D\rho\mathbf{v}\,dV, \tag{1}$$

$$\int_{\partial D}\mathbf{y}\times\boldsymbol{\sigma}\mathbf{n}\,dA+\int_D\mathbf{y}\times\rho\mathbf{b}\,dV=\frac{d}{dt}\int_D\mathbf{y}\times\rho\mathbf{v}\,dV, \tag{2}$$

$$\int_{\partial D}(\boldsymbol{\sigma}\mathbf{n}\cdot\mathbf{v}+\mathbf{q}\cdot\mathbf{n})dA+\int_D(\rho\mathbf{b}\cdot\mathbf{v}+\rho r)dV =\frac{d}{dt}\int_D\left(\frac{1}{2}\rho\mathbf{v}\cdot\mathbf{v}+\rho\varepsilon\right)dV. \tag{3}$$

Here $\mathbf{v}=\dot{\mathbf{y}}$ denotes particle velocity, $\rho(\mathbf{x})$ is the mass density in the reference configuration which is assumed to be continuous on $R$, and $\mathbf{n}$ is a unit outward normal vector on $\partial D$.

At a point in $R$ at which the fields are smooth the balance laws (1)–(3) yield the usual field equations

$$\mathrm{Div}\,\boldsymbol{\sigma}+\rho\mathbf{b}=\rho\dot{\mathbf{v}}, \tag{4}$$

$$\boldsymbol{\sigma}\mathbf{F}^T=\mathbf{F}\boldsymbol{\sigma}^{\mathrm{T}}, \tag{5}$$

$$\boldsymbol{\sigma}\cdot\dot{\mathbf{F}}+\mathrm{Div}\,\mathbf{q}+\rho r=\rho\dot{\varepsilon}, \tag{6}$$

where $\mathbf{F}=\mathrm{Grad}\,\mathbf{y}$ is the deformation gradient tensor.

Next, suppose that there is a surface $S_t$ in $R$ such that the fields $\mathbf{F}$, $\mathbf{v}$, $\mathbf{q}$, $\boldsymbol{\sigma}$ and $\varepsilon$ suffer jump discontinuities across $S_t$ while being continuous on either side of it. Such a surface may represent, for example, the Lagrangian image of a shock wave or an interface separating two material phases. Let $V_n\geqslant 0$ denote the normal velocity of propagation of this interface. We refer to the side into which $V_n$ points as the positive side of $S_t$. For any field quantity $g(\mathbf{x},t)$ let $\overset{+}{g}$ and $\overline{g}$ denote the limiting values of $g$ as a point on $S_t$ is approached from its positive and negative side, respectively. Then, we let $[[g]]$ and $\langle g\rangle$ denote the jump and the average values of $g$ on $S_t$:

$$[[g]]=\overset{+}{g}-\overline{g},\quad \langle g\rangle=\frac{1}{2}(\overset{+}{g}+\overline{g}). \tag{7}$$

At a point on $S_t$, the balance laws (1)–(3) yield the usual jump conditions

$$[[\boldsymbol{\sigma}\mathbf{n}]]+[[\rho\mathbf{v}]]V_n=0, \tag{8}$$

$$[[\boldsymbol{\sigma}\mathbf{n}\cdot\mathbf{v}]]+\left[\left[\rho\varepsilon+\frac{1}{2}\rho\mathbf{v}\cdot\mathbf{v}\right]\right]V_n+[[\mathbf{q}\cdot\mathbf{n}]]=0. \tag{9}$$

The energy jump condition (9) can be written in the following alternative form by making use of (8) and $[[\mathbf{v}]]+V_n[[\mathbf{F}\mathbf{n}]]=\mathbf{0}$ which follows from the continuity of the deformation (for algebraic details see, for example, [5]):

$$([[\rho\varepsilon]]-\langle\boldsymbol{\sigma}\rangle\cdot[[\mathbf{F}]])V_n=-[[\mathbf{q}\cdot\mathbf{n}]]. \tag{10}$$

## 3 Rate of Entropy Production

In order to address the second law of thermodynamics one must consider two additional fields, viz. the temperature $\theta(\mathbf{x},t)$ and the entropy per unit mass $\eta(\mathbf{x},t)$. Suppose that $\theta(\cdot,t)$ is piecewise continuous with a piecewise continuous gradient on $R$, and that $\eta$ is piecewise continuous with piecewise continuous first derivatives on $R\times[t_1,t_2]$; $\theta$ and $\eta$ are permitted to suffer jump discontinuities across $S_t$. The rate of entropy production associated with a subregion $D$ is defined by

$$\Gamma=\frac{d}{dt}\int_D\rho\eta\,dV-\int_{\partial D}\frac{\mathbf{q}\cdot\mathbf{n}}{\theta}dA-\int_D\frac{\rho r}{\theta}dV, \tag{11}$$

and the second law of thermodynamics requires that $\Gamma\geqslant 0$ for all regions $D$ and all processes. When the region $D$ intersects the interface $S_t$ one can rewrite (11) in the form

$$\Gamma=\int_D\left\{\rho\dot{\eta}-\mathrm{Div}\left(\frac{\mathbf{q}}{\theta}\right)-\frac{\rho r}{\theta}\right\}dV -\int_{S_t\cap D}\left\{[[\rho\eta]]V_n+\left[\left[\frac{\mathbf{q}\cdot\mathbf{n}}{\theta}\right]\right]\right\}dA \tag{12}$$

by carrying out a standard calculation; e.g. see page 116 of [12]. The first term in (12) represents the entropy production rate in the bulk of the body and the second term is associated with the moving interface. Let $\Gamma_s$ denote the rate of entropy production due to the propagating surface:

$$\Gamma_s=-\int_{S_t}\left\{[[\rho\eta]]V_n+\left[\left[\frac{\mathbf{q}\cdot\mathbf{n}}{\theta}\right]\right]\right\}dA. \tag{13}$$

One finds by using (13) and (10), that $\Gamma_s$ can be alternatively expressed as

$$\Gamma_s=\int_{S_t}\left\{\left[\left\langle\frac{1}{\theta}\right\rangle([[\rho\varepsilon]]-\langle\boldsymbol{\sigma}\rangle\cdot[[\mathbf{F}]])-[[\rho\eta]]\right]V_n +\langle\mathbf{q}\cdot\mathbf{n}\rangle\left[\left[\frac{1}{\theta}\right]\right]\right\}dA. \tag{14}$$

In an adiabatic process there is no heat transfer: $\mathbf{q}=\mathbf{0}$ and $r=0$. On the other hand if the process is not adiabatic, the typical heat conduction law, whatever it may be, involves Grad $\theta$ and therefore the partial differential equations resulting from using the constitutive relationships in the energy Eq. (6) involve (at least) the second spatial derivative of $\theta$; thus, one usually requires the temperature to be continuous in non-adiabatic processes: $[[\theta]]=0$ on $S_t$. Thus in both the adiabatic and non-adiabatic cases one has $[[\theta]]\overset{\pm}{\mathbf{q}}=\mathbf{0}$ on $S_t$ and therefore necessarily

$$\langle \mathbf{q}\cdot\mathbf{n}\rangle\left[\left[\frac{1}{\theta}\right]\right]=0 \quad \text{and} \quad \left(\left\langle\frac{1}{\theta}\right\rangle-\frac{1}{\langle\theta\rangle}\right)[[\mathbf{q}\cdot\mathbf{n}]]=0. \tag{15}$$

In view of this and (10), we can write $\Gamma_s$ as

$$\Gamma_s=\int_{S_t}\frac{[[\rho\varepsilon]]-\langle\boldsymbol{\sigma}\rangle\cdot[[\mathbf{F}]]-\langle\theta\rangle[[\rho\eta]]}{\langle\theta\rangle}V_n\,dA, \tag{16}$$

or in terms of the Helmholtz free-energy $\psi=\varepsilon-\eta\theta$ as

$$\Gamma_s=\int_{S_t}\frac{[[\rho\psi]]-\langle\boldsymbol{\sigma}\rangle\cdot[[\mathbf{F}]]+\langle\rho\eta\rangle[[\theta]]}{\langle\theta\rangle}V_n\,dA. \tag{17}$$

## 4 Driving Traction

The rate of entropy production due to the propagating interface can be written as

$$\Gamma_s=\int_{S_t}\frac{fV_n}{\langle\theta\rangle}dA \tag{18}$$

where

$$\begin{aligned} f&=[[\rho\psi]]-\langle\boldsymbol{\sigma}\rangle\cdot[[\mathbf{F}]]+\langle\rho\eta\rangle[[\theta]]\\ &=[[\rho\varepsilon]]-\langle\boldsymbol{\sigma}\rangle\cdot[[\mathbf{F}]]-\langle\theta\rangle[[\rho\eta]] \end{aligned} \tag{19}$$

is called the driving traction or Eshelby force. The second law of thermodynamics requires that $fV_n\geqslant 0$ on $S_t$ which specifies the direction in which the interface is permitted to move. This result is valid for any continuum undergoing an arbitrary thermomechanical process which may or may not be adiabatic. If the process is adiabatic, (19) and (10) yield $f=-\langle\theta\rangle[[\rho\eta]]$. If it is not adiabatic, (19) specializes to $f=[[\rho\psi]]-\langle\boldsymbol{\sigma}\rangle\cdot[[\mathbf{F}]]$.

In the special case of a thermoelastic material one has $\psi=\hat{\psi}(\mathbf{F},\theta)$ and the stress and entropy are given by the constitutive relationships $\boldsymbol{\sigma}=\rho\hat{\psi}_{\mathbf{F}}$, $\eta=-\hat{\psi}_\theta$. Equivalently one has $\varepsilon=\hat{\varepsilon}(\mathbf{F},\eta)$ with the stress and temperature given by $\boldsymbol{\sigma}=\rho\hat{\varepsilon}_{\mathbf{F}}$, $\theta=\hat{\varepsilon}_\eta$. Thus for a thermoelastic material (19) can be written as

$$\begin{aligned} f&=[[\rho\hat{\psi}]]-\langle\rho\hat{\psi}_{\mathbf{F}}\rangle\cdot[[\mathbf{F}]]-\langle\rho\hat{\psi}_\theta\rangle[[\theta]]\\ &=[[\rho\hat{\varepsilon}]]-\langle\rho\hat{\varepsilon}_{\mathbf{F}}\rangle\cdot[[\mathbf{F}]]-\langle\rho\hat{\varepsilon}_\eta\rangle[[\eta]] \end{aligned} \tag{20}$$

which is reminiscent of the Legendre transforms of $\rho\hat{\psi}(\mathbf{F},\theta)$ and $\rho\hat{\varepsilon}(\mathbf{F},\eta)$.


## Acknowledgment

The results reported here were obtained in the course of research supported by the National Science Foundation.


## References


[1] Eshelby, J. D., 1956, ''Continuum Theory of Lattice Defects,'' *Solid State Physics*, Vol. 3, F. Seitz and D. Turnbull, eds., Academic Press, San Diego, pp. 79–144.

[2] Callen, H. B., 1985, *Thermodynamics and an Introduction to Thermostatistics*, Second Ed., John Wiley and Sons, New York, Chapter 14.

[3] Kestin, J., 1968, *A Course on Thermodynamics*, Vol. II, McGraw-Hill, New York, Chapter 14.

[4] Truesdell, C., 1969, *Rational Thermodynamics* (Lecture 7), Springer-Verlag, New York.

[5] Abeyaratne, R., and Knowles, J. K., 1990, ''On the Driving Traction Acting on a Surface of Strain Discontinuity in a Continuum,'' J. Mech. Phys. Solids, **38**, pp. 345–360.

[6] Abeyaratne, R., and Knowles, J. K., 1991, ''Kinetic Relations and the Propagation of Phase Boundaries in Solids,'' Arch. Ration. Mech. Anal., **114**, pp. 119–154.

[7] Abeyaratne, R., Kim, S-J., and Knowles, J. K., 1994, ''A One-Dimensional Continuum Model for Shape-Memory Alloys,'' Int. J. Solids Struct., **31**, pp. 2229–2249.

[8] Rosakis, P., and Tsai, H., 1995, ''Dynamic Twinning Processes in Crystals,'' Int. J. Solids Struct., **32**, pp. 2711–2723.

[9] Heidug, W., and Lehner, F. K., 1985, ''Thermodynamics of Coherent Phase Transformations in Non-Hydrostatically Stressed Solids,'' Pure Appl. Geophys., **123**, pp. 91–98.

[10] Truskinovsky, L., 1985, ''Structure of an Isothermal Phase Discontinuity,'' Sov. Phys. Dokl., **30**, pp. 945–948.

[11] Abeyaratne, R., and Knowles, J. K., 1994, ''Dynamics of Propagating Phase Boundaries: Adiabatic Theory for Thermoelastic Solids,'' Physica D, **79**, pp. 269–288.

[12] Chadwick, P., 1976, *Continuum Mechanics*, John Wiley and Sons, New York.


# Characterizing Damping and Restitution in Compliant Impacts via Modified K-V and Higher-Order Linear Viscoelastic Models


**E. A. Butcher**
Assoc. Mem. ASME, Department
of Mechanical Engineering,
University of Alaska, Fairbanks, AK 99775-5905

**D. J. Segalman**
Fellow ASME, Sandia National Laboratories,[1]
P.O. Box 5800,
MS 0847, Albuquerque, NM 87185-0847


## 1 Introduction

Time-domain models for compliant impacts have been widely used to model collision dynamics as finite-time events. The most common way to account for energy dissipation in the compliant impact model has been via the standard Kelvin-Voigt (K-V) viscoelastic model

$$F(t)=kx+c\dot{x} \tag{1}$$

in which the resulting equation of motion assumes the familiar linear form

$$\ddot{x}+2\zeta\omega_n\dot{x}+\omega_n^2x=0 \tag{2}$$

from vibration theory where $\omega_n=\sqrt{k/m}$ and $\zeta=c/(2\sqrt{km})$. The initial conditions $x(0)=0$ and $\dot{x}(0)=v_0$ yield the solution

$$x(t)=\frac{v_0}{\omega_d}\exp(-\zeta\omega_n t)\sin\omega_d t \tag{3}$$

where $\omega_d=\omega_n\sqrt{1-\zeta^2}$. If the impact duration is assumed to be a half-period of vibration associated with the damped frequency, then the exact restitution coefficient is obtained easily in terms of the dimensionless damping ratio as

[1]Sandia National Laboratories is a multiprogram laboratory operated by Sandia Corporation, A Lockheed Martin Company for the U.S. Department of Energy under Contract DE-ACO4, 94AL85000.


Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, August 25, 1998; final revision, February 10, 2000. Associate Technical Editor: V. K. Kinra.




Copyright © 2000 by ASME

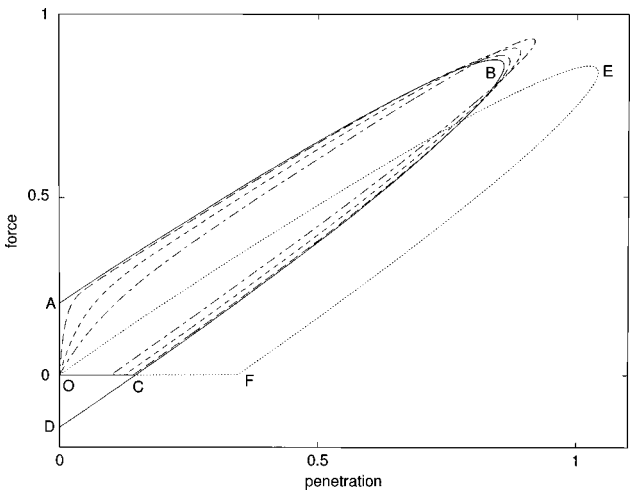



**Fig. 1 Hysteresis diagrams for the Kelvin-Voigt (solid), Maxwell (dotted), and standard linear impact models with $\omega_n=1$ and $\zeta=0.1$ where $\eta=0.0$ (solid), 0.05 (long-dashed), 0.2 (short-dashed), and 0.4 (short-long-dashed). The modified K-V and standard linear models omit the tension at the conclusion of the restitution phase of impact.**

$$e=-\frac{\dot{x}(t_r)}{\dot{x}(0)}=\exp\left(-\frac{\zeta\pi}{\sqrt{1-\zeta^2}}\right) \tag{4}$$

where $t_r=\pi/\omega_d$ is the release time ([1,2]). While the undamped collision is elastic, for critical damping or overdamping the collision is purely plastic. Another reason the half-period K-V model has been widely used is that Eq. (4) may easily be used to obtain the impact damping parameter $c$ as

$$c=2|\ln e|\sqrt{\frac{km}{(\ln e)^2+\pi^2}} \tag{5}$$

in terms of an experimentally obtained restitution coefficient ([3]).

The energy $E_L$ lost in the impact is

$$E_L=E_0(1-e^2)=E_0\left(1-\exp\left(-\frac{2\zeta\pi}{\sqrt{1-\zeta^2}}\right)\right) \tag{6}$$

where $E_0$ is the initial kinetic energy. This energy loss is represented in Fig. 1 by the area enclosed inside the hysteresis curve O-A-B-D-O. The peak elastic potential energy stored in the impact is

$$U=\frac{1}{2}kx_{\max}^2=E_0\exp\left(-\frac{2\zeta}{\sqrt{1-\zeta^2}}\tan^{-1}\frac{\sqrt{1-\zeta^2}}{\zeta}\right) \tag{7}$$

from which an equivalent linear damping ratio $\zeta_{eq}$ may be found via the loss factor $\eta=E_L/(2\pi U)$ as

$$\zeta_{eq}=\frac{\eta}{2}=\frac{1}{4\pi}\exp\left(\frac{2\zeta}{\sqrt{1-\zeta^2}}\tan^{-1}\frac{\sqrt{1-\zeta^2}}{\zeta}\right)\times\left(1-\exp\left(-\frac{2\zeta\pi}{\sqrt{1-\zeta^2}}\right)\right) \tag{8}$$

Thus the bilinear impact model may be replaced by an equivalent linear Kelvin-Voigt model with damping constant $c_{eq}=2\zeta_{eq}\sqrt{km}$ which dissipates $E_L$ energy per period of vibration. This technique is often advantageous in vibratory impact problems.

Hunt and Crossley [4] noticed, however, that linear viscous damping in the K-V model gives an unrealistic hysteresis diagram for the impact force-deflection curve. Specifically, they noted that the parallel linear dashpot results in discontinuous force profiles at initial contact and release as well as a nonphysical tensile force applied during the end of the restitution phase. In order to eliminate this force discontinuity, they suggested a nonlinear damping function for use with the Hertzian stiffness model which complies with the expected boundary conditions of vanishing force at contact and release. Estimates of the corresponding restitution coefficient in terms of the model's damping parameters were made by the above authors and Herbert and McWhannell [5], who also noted that the effects of eliminating the force discontinuities include a more realistic frequency content in the impulse generated.

While other authors (e.g., [6,7]) have proposed different nonlinear models which also satisfy the expected boundary conditions, there have been few efforts to eliminate the force discontinuities in the K-V model while remaining within the framework of linear viscoelasticity. This paper attempts to help fill this gap. In contrast to using the various nonlinear models, this approach enables the associated damping and restitution to be characterized analytically without the need to approximate. First, the standard K-V model is reconsidered here under a different assumption regarding the restitution phase: that the mass releases when the *force* vanishes—*before* the initial contact location is reached. This approach was recently used by Luo and Hanagud [8] in order to improve the modeling and simulation of vibration absorbers with motion-limiting stops. In order also to guarantee that the force vanishes upon impact, higher-order viscoelastic (Maxwell and standard linear) models are implemented in which the boundary conditions in the force-displacement hysteresis curve are all satisfied and the force history is entirely continuous. A similar model for the impact surface has been utilized in a previous paper ([9]) in an effort to circumvent the previously mentioned discontinuities in the dynamic model of robotic manipulator collisions. Unlike these studies which were concerned with simulation, however, this paper presents analytical values of the restitution coefficient and related quantities for the viscoelastic models in terms of the dimensionless viscoelastic parameters. Hysteresis diagrams and restitution coefficients for each model are plotted and compared.

## 2 Modified K-V Model

In order to eliminate the tension in the K-V model, a better representation of the dynamics allows the mass to release when the net force vanishes. In Fig. 1, this occurs at point C. The resulting area of the hysteresis curve O-A-B-C-O (the energy loss) is thus smaller than that obtained using a half-period of vibration. Setting the force in Eq. (1) to zero, the release time is found as

$$t_r=\frac{1}{\omega_d}\tan^{-1}\left(\frac{2\zeta\sqrt{1-\zeta^2}}{2\zeta^2-1}\right) \tag{9}$$

which yields the restitution coefficient

$$e=\exp\left(-\frac{\zeta}{\sqrt{1-\zeta^2}}\tan^{-1}\left(\frac{2\zeta\sqrt{1-\zeta^2}}{2\zeta^2-1}\right)\right);\quad \zeta<1 \tag{10}$$

As $\zeta\to1, e\to\exp(-2)\approx0.14$ so that, unlike the half-period version, the impact is not perfectly plastic when the damping is critical. Instead, the nonzero restitution coefficient

$$e=\left(\frac{\zeta-\sqrt{\zeta^2-1}}{\zeta+\sqrt{\zeta^2-1}}\right)^{\zeta/\sqrt{\zeta^2-1}};\quad \zeta>1 \tag{11}$$

which matches Eq. (10) for $\zeta=1$ is obtained for overdamping. The energy lost in an impact is $E_L=E_0(1-e^2)$. Since the peak potential energy is given by Eq. (7) for $\zeta<1$ and by $U=E_0e$ for $\zeta>1$, the equivalent linear damping ratio may be obtained as

$$\zeta_{\mathrm{eq}}=\begin{cases}\dfrac{1}{4\pi}\exp\left(\dfrac{2\zeta}{\sqrt{1-\zeta^2}}\tan^{-1}\dfrac{\sqrt{1-\zeta^2}}{\zeta}\right)\left(1-\exp\left(-\dfrac{2\zeta}{\sqrt{1-\zeta^2}}\tan^{-1}\left(\dfrac{2\zeta\sqrt{1-\zeta^2}}{2\zeta^2-1}\right)\right)\right); & \zeta<1\\ \dfrac{1}{2\pi}\sinh\left(\dfrac{-\zeta}{\sqrt{\zeta^2-1}}\ln\left(\dfrac{\zeta-\sqrt{\zeta^2-1}}{\zeta+\sqrt{\zeta^2-1}}\right)\right); & \zeta>1\end{cases} \tag{12}$$

from which an equivalent damping constant for use in vibratory impact is $c_{eq}=2\zeta_{eq}\sqrt{km}$. Although $\zeta_{eq}$ is less than in the half-period model, the difference remains less than 0.02 for $\zeta<1$. Thus, eliminating the force discontinuity at release results in a minimal decrease in equivalent damping.

## 3 Maxwell Model

The force discontinuity at impact cannot be eliminated in the K-V model due to the model's lack of an instantaneous elasticity. Instead, a higher-order viscoelastic model which portrays instantaneous elasticity may be utilized. The most basic of these is the Maxwell model in which the force-displacement relation is ([10])

$$F+\frac{c}{k}\dot{F}=c\dot{x}. \tag{13}$$

Equation (13) results in the third-order differential equation

$$\dddot{x}+2\zeta\omega_n\ddot{x}+\omega_n^2\dot{x}=0 \tag{14}$$

describing the impact dynamics where $\omega_n=\sqrt{k/m}$ and $\zeta=\sqrt{km}/(2c)$. The initial conditions $x(0)=0$, $\dot{x}(0)=v_0$, and $\ddot{x}(0)=0$ yield the solution

$$x(t)=\exp(-\zeta\omega_n t)\left(\frac{v_0}{\omega_d}(1-2\zeta^2)\sin\omega_d t-\frac{2\zeta v_0}{\omega_n}\cos\omega_d t\right)+\frac{2\zeta v_0}{\omega_n} \tag{15}$$

where $\omega_d=\omega_n\sqrt{1-\zeta^2}$.

As seen in Fig. 1 (hysteresis curve O-E-F-O), the discontinuity upon impact has been eliminated via the third initial condition. Furthermore, if release occurs when the force vanishes then all boundary conditions are satisfied. Since the release time is equivalent to a half-period ($t_r=\pi/\omega_d$), the restitution coefficient $e$ and energy lost $E_L$ are thus found to be equivalent to those obtained in the half-period K-V model. Thus as $\zeta\to 1, e\to 0$ and the collision becomes perfectly plastic. Since Eq. (4) also applies for the Maxwell model, the damping parameter is easily calculated in terms of the coefficient of restitution (as in the K-V model) as

$$c=\frac{\sqrt{km((\ln e)^2+\pi^2)}}{2|\ln e|}. \tag{16}$$

Because the spring and dashpot are in series, the peak elastic potential energy stored in the spring is found in terms of the maximum *force* as

$$U=\frac{1}{2}k\left(\frac{F_{\max}}{k}\right)^2=E_0\exp\left(-\frac{2\zeta}{\sqrt{1-\zeta^2}}\tan^{-1}\frac{\sqrt{1-\zeta^2}}{\zeta}\right). \tag{17}$$

The equivalent linear damping ratio, therefore, is also equivalent to that for the half-period K-V model so that the Maxwell impact model may be replaced by a linear K-V model with $c_{eq}=2\zeta_{eq}\sqrt{km}$ which dissipates $E_L$ energy per period of vibration. It should be observed that, although certain quantities of the half-period K-V and Maxwell impact models are conveniently equivalent, their inherent physics are completely different as represented by the corresponding hysteresis curves.

## 4 Standard Linear Model

Another instantaneously elastic higher-order viscoelastic model which can be utilized is the standard linear model which consists of a K-V element (with spring constant $k_2$) in series with another spring $k_1$. The force-displacement relation is ([10])

$$(k_1+k_2)F+c\dot{F}=k_1k_2x+k_1c\dot{x}. \tag{18}$$

Equation (18) results in the third-order equation

$$\frac{2\zeta\eta}{\omega_n}\dddot{x}+\ddot{x}+2\zeta\omega_n\dot{x}+\omega_n^2x=0 \tag{19}$$

for the impact dynamics where $\omega_n=\sqrt{k_1k_2/((k_1+k_2)m)}$, while $\zeta=k_1c/(2(k_1+k_2)m\omega_n)$ and $\eta=k_2/(k_1+k_2)$ are the dimensionless viscoelastic parameters. In the limit as $k_1\to\infty$, then $\omega_n\to\sqrt{k_2/m}, 2\zeta\omega_n\to c/m, \eta\to 0$, and the system becomes a K-V model with natural frequency $\omega_n$ and damping ratio $\zeta$. Hence for $\eta\ll 1$, Eq. (19) represents a perturbation of the standard K-V model (Eq. (2)). Although the exact solution and restitution coefficient are intractable in this model, an approximate closed-form solution may be obtained be means of a singular perturbation technique ([11]) in which the three roots $\alpha$ and $\beta\pm i\gamma$ are obtained to first order in $\eta$ as

$$\alpha=-\frac{\omega_n}{2\zeta\eta}+2\zeta\omega_n+\eta(1-4\zeta^2)\zeta\omega_n$$
$$\beta=-\zeta\omega_n+\eta(1-4\zeta^2)\zeta\omega_n$$
$$\gamma=\omega_n\sqrt{1-\zeta^2}+\eta(3-4\zeta^2)\frac{\zeta^2}{\sqrt{1-\zeta^2}}\omega_n. \tag{20}$$

The perturbation series converges providing $\zeta<1/(2\sqrt{\eta})$.

The force-displacement hysteresis curves in Fig. 1 correspond to different values of the dimensionless parameter $\eta$. This parameter affects the model's instantaneous elasticity and can be adjusted to sufficiently ''smooth out'' the K-V force discontinuity at the origin. By allowing the mass to release at vanishing force, each of the force boundary conditions remain satisfied. Hence, for small values of $\eta$, the damping and restitution for this model are perturbations of those for the modified K-V model. The restitution coefficient was found for $\zeta<1$ to first order in $\eta$ as

$$e=\exp\left(\left(-\frac{\zeta}{\sqrt{1-\zeta^2}}+\eta f_1(\zeta)\right)\left(\tan^{-1}\left(\frac{2\zeta\sqrt{1-\zeta^2}}{2\zeta^2-1}\right)+\eta f_2(\zeta)\right)\right) \tag{21}$$

where $f_1(\zeta)$ and $f_2(\zeta)$ were found to expand as $f_1(\zeta)=\zeta-\zeta^3/2+O(\zeta^5)$ and $f_2(\zeta)=2\zeta-3\zeta^3+O(\zeta^5)$. These approximations are accurate for small $\zeta$ and break down for $\zeta$ near unity. In order to verify the analytical expression in Eq. (21) using the expansions for $f_1(\zeta)$ and $f_2(\zeta)$, the restitution for different $\eta$ values was also obtained numerically from the final velocity at release. It was found that the two results are practically identical for small values of both $\eta$ and $\zeta$. Finally, unlike the previous models considered, an equivalent linear damping ratio is not easily obtained for this model in terms of $\eta$ and $\zeta$.

## 5 Discussion

The restitution coefficients for each of the viscoelastic impact models are plotted as a function of $\zeta$ in Fig. 2 in which the half-

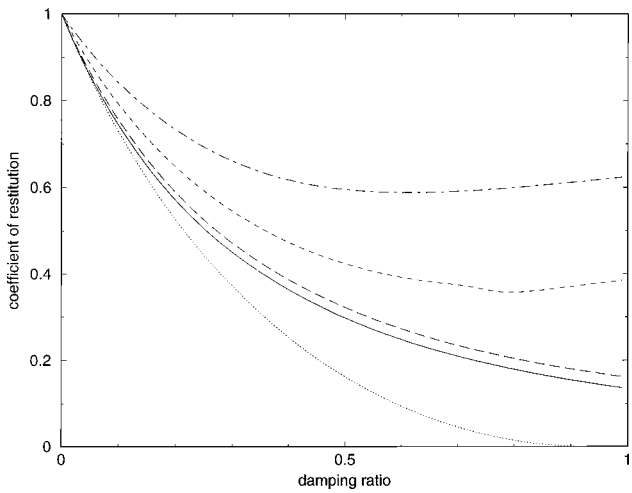



**Fig. 2 Restitution coefficients for the half-period K-V and Maxwell models (dotted), modified K-V (solid), and standard linear impact model (numerical results) where $\eta$=0.0 (solid), 0.05 (long-dashed), 0.2 (short-dashed), and 0.4 (short-long-dashed)**

period K-V and Maxwell results are identical. The numerical results for the standard linear model are shown for several values of $\eta$. The corresponding perturbation results (not shown) are essentially the same for small $\eta$ and $\zeta$. It is seen that the restitution coefficient of the half-period K-V and Maxwell models indeed vanishes as $\zeta \rightarrow 1$. Hence, a finite damping constant $c_p$ may be associated with purely plastic impacts in these two models where $c_p = 2\sqrt{km}$ for K-V and $c_p = \sqrt{km}/2$ for Maxwell. In the modified K-V model, however, the restitution vanishes only as $\zeta \rightarrow \infty$ so that the impact can never be purely plastic. It is also seen that the modified K-V model has a restitution coefficient which is very close to that of the half-period K-V and Maxwell models for small damping ratios since the release times in these models for low damping are nearly the same. The advantages of analytically obtaining both the impact damping parameter (in terms of an experimentally obtained restitution coefficient) and the equivalent linear damping constant for use in vibroimpact, together with the even more significant fact that all of the force boundary conditions are satisfied, leads to the conclusion that the Maxwell model is an attractive choice for practical implementation in the modeling of dissipative compliant impacts. However, the standard linear model may also be helpful in ''smoothing out'' the K-V impact discontinuity. Especially if the impact duration is relatively long or there are additional static forces on the impact surface, the finite static deformation of this model is preferable to the fluid-like behavior of the Maxwell model. Furthermore, if the damping is small and the instantaneous stiffness is large, then the impact dynamics and restitution may be found as a perturbation of those for the modified K-V model as was done here. Finally, further work is needed to extend these results to general planar and three-dimensional collision theories and to include the use of kinetic and energetic restitution coefficients.

## References


[1] Yigit, A. S., Ulsoy, A. G., and Scott, R. A., 1990, ''Dynamics of a Radially Rotating Beam With Impact, Part 1: Theoretical and Computational Model,'' ASME J. Vibr. Acoust., **112**, pp. 65–70.

[2] Brach, R. M., 1991, *Mechanical Impact Dynamics: Rigid Body Collisions*, John Wiley and Sons, New York.

[3] Brogliato, B., 1996, *Nonsmooth Impact Mechanics: Models, Dynamics, and Control*, Springer-Verlag, New York.

[4] Hunt, K. H., and Crossley, F. R. E., 1975, ''Coefficient of Restitution Interpreted as Damping in Vibroimpact,'' ASME J. Appl. Mech., **42**, pp. 440–445.

[5] Herbert, R. G., and McWhannell, D. C., 1977, ''Shape and Frequency Composition of Pulses From an Impact Pair,'' J. Eng. Ind., **99**, pp. 513–518.

[6] Lee, T. W., and Wang, A. C., 1983, ''On the Dynamics of Intermittent-Motion Mechanisms,'' ASME J. Mech. Des., **105**, pp. 534–540.

[7] Khulief, Y. A., and Shabana, A. A., 1987, ''A Continuous Force Model for the Impact Analysis of Flexible Multibody Systems,'' Mech. Mach. Theory, **22**, No. 3, pp. 213–224.

[8] Luo, H., and Hanagud, S., 1998, ''On the Dynamics of Vibration Absorbers With Motion-Limiting Stops,'' ASME J. Appl. Mech., **65**, pp. 223–233.

[9] Mills, J. K., and Nguyen, C. V., 1992, ''Robotic Manipulator Collisions: Modeling and Simulation,'' ASME J. Dyn. Syst., Meas., Control, **114**, pp. 650–659.

[10] Haddad, Y. M., 1995, *Viscoelasticity of Engineering Materials*, Chapman and Hall, New York.

[11] Nayfeh, A. H., 1981, *Introduction to Perturbation Techniques*, John Wiley and Sons, New York.


# Finite-Amplitude Elastic Instability of Plane-Poiseuille Flow of Viscoelastic Fluids


**R. E. Khayat**
e-mail: rkhayat@eng.uwo.ca

**N. Ashrafi**

Department of Mechanical and Materials Engineering,
University of Western Ontario, London,
Ontario N6A 5B9, Canada



*The purely elastic stability and bifurcation of the one-dimensional plane Poiseuille flow is determined for a large class of Oldroyd fluids with added viscosity, which typically represent polymer solutions composed of a Newtonian solvent and a polymeric solute. The problem is reduced to a nonlinear dynamical system using the Galerkin projection method. It is shown that elastic normal stress effects can be solely responsible for the destabilization of the base (Poiseuille) flow. It is found that the stability and bifurcation picture is dramatically influenced by the solvent-to-solute viscosity ratio, $\varepsilon$. As the flow deviates from the Newtonian limit and $\varepsilon$ decreases below a critical value, the base flow loses its stability. Two static bifurcations emerge at two critical Weissenberg numbers, forming a closed diagram that widens as the level of elasticity increases.* [S0021-8936(00)00703-0]


## 1 Introduction

While the problem of stability of plane-Poiseuille flow (PPF) has been extensively investigated for Newtonian fluids, relatively little attention has been devoted to the flow of viscoelastic fluids. The presence of viscoelasticity is expected to dramatically alter the stability and bifurcation picture in PPF, and yet no study has so far predicted the nonlinear bifurcation from the base flow. The presence of additional nonlinearities that are usually part of any realistic constitutive model ([1]) are expected to lead to the departure from the Newtonian picture. Similarly to the case of Taylor-Couette flow, there is experimental evidence that the base flow in a channel may lose its stability as a result of fluid elasticity inside the tube ([2]). This mechanism is now known as constitutive instability, as opposed to stick-slip induced instability. This mechanism of loss of stability should not be confounded with the short-wave instability due to a change in type of the field equations,


Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, April 28, 1999; final revision, July 27, 1999. Associate Technical Editor: A. K. Mal.




**Copyright © 2000 by ASME**

which is known as Hadamard instability ([3]). The emergence of surface instability at the exit of an extrusion die (sharkskin and melt fracture) keeps hinting at the possibility of a link with a hydrodynamic instability inside the channel, away and upstream from the exit ([4,5]). However, linear stability analyses of channel (Couette and Poiseuille) flows, using elementary constitutive models, such as Maxwell and Oldroyd-B fluids, failed to assert that the base flow may be linearly unstable when the level of elasticity (Weissenberg number) exceeds a critical level ([4,5]). More recent studies based on more generalized constitutive models of the Oldroyd class showed that the base flow in a channel can become unstable to small perturbations for some range of Weissenberg numbers ([6–8]). These generalized constitutive models display a nonmonotonic shear-stress/shear-rate curve. The range of instability coincides with the negative slope of the stress curve. However, only linear stability analyses were carried out.

The present study focuses on the nonlinear constitutive instability of the PPF of high-molecular-weight fluids. These fluids are typically composed of a Newtonian solvent and a polymeric solute. The Johnson-Segalman (JS) constitutive model is used, which is a highly nonlinear equation, and is one of the very few constitutive models that exhibit a nonmonotonic stress/shear-rate curve. It is thus expected that, while the presence of inertia and shear thinning alone can destabilize the flow, fluid elasticity or normal stresses will give rise to additional nonlinearities and coupling among the flow variables, making an already complex problem (due mainly to inertia) even more difficult to solve. Similarly to any flow in the transition regime, the PPF of viscoelastic fluids involves a continuous range of excited spatio-temporal scales. In order to assess the influence of the arbitrarily many smaller length and time scales on the flow, one would have to resort to the resolution of the flow at the small-scale level. This issue remains unresolved since, despite the great advances in storage and speed of modern computers, it will not be possible to resolve all of the continuous ranges of scales in the transition regime.

It is by now well established that dynamical systems can be a viable alternative to conventional numerical methods as one probes the nonlinear range of flow behavior ([9]). Dynamical systems are obtained using the Galerkin approximation. The velocity and stress components assume truncated Fourier or other orthogonal representations in space, depending on the boundary conditions. The expansion coefficients are functions of time alone, thus leading to a nonlinear system upon projection of the equations onto the various modes. The relative simplicity of dynamical systems, and the rich sequence of nonlinear flow phenomena exhibited by their solution, have been the major contributing factors to their widespread use as models for examining the onset of nonlinear behavior. The dynamical system approach has typically been used to handle simple flow configurations, and most particularly Newtonian flows. Recently, this approach has been attempted for non-Newtonian flows in thermal convection ([10–12]) and rotating flow ([13–17]). For Taylor-Couette flow, comparison was carried out with the experiments of Muller et al. [18], leading to excellent agreement ([15]). A modal expansion similar to that in [14,15] is used to solve the current problem.

## 2 Problem Formulation and Solution Procedure

Consider the plane channel (Poiseuille) flow of an incompressible viscoelastic fluid of density $\rho$, relaxation time $\lambda$, and viscosity $\eta$. In this study, only fluids that can be reasonably represented by a single relaxation time and constant viscosity are considered. The fluid considered here, is a polymer solution composed of a Newtonian solvent and a polymer solute of viscosities $\eta_s$ and $\eta_p$, respectively. Therefore $\eta=\eta_s+\eta_p$. The velocity, time, space coordinates, pressure, and stress are nondimensionalized by $d/\lambda$, $\lambda$, $d$, $\eta_p U/d$ and $\eta_p/\lambda$, respectively. Here $U$ is the maximum velocity of the base Poiseuille flow, and $d$ is the gap between the two plates. There are three important similarity groups in the problem, namely, the Reynolds number, Re, the Weissenberg number, We, and the solvent-to-solute viscosity ratio, $\varepsilon$, which are given, respectively, by

$$\mathrm{Re}=\frac{d^2\rho}{\eta_p\lambda},\quad \mathrm{We}=\frac{U\lambda}{d},\quad \varepsilon=\frac{\eta_s}{\eta_p}. \tag{1}$$

The continuity and conservation of momentum equations for a general incompressible viscoelastic fluid are given in dimensionless form as

$$\nabla\cdot\mathbf{u}=0,\quad \mathrm{Re}\frac{d\mathbf{u}}{dt}=-\mathrm{We}\nabla p+\nabla\cdot\boldsymbol{\tau}+\varepsilon\nabla^2\mathbf{u} \tag{2}$$

where $\mathbf{u}$ is the velocity vector, $p$ is the pressure, $\tau$ is the polymeric contribution of the stress tensor, $t$ is the time, $d/dt$ is the substantial derivative operator, and $\nabla$ is the gradient operator. The constitutive equation adopted in this study belongs to the *Oldroyd* class of incompressible viscoelastic fluids:

$$\frac{d\boldsymbol{\tau}}{dt}-\left(1-\frac{\zeta}{2}\right)[(\nabla\mathbf{u})^t\cdot\boldsymbol{\tau}+\boldsymbol{\tau}\cdot\nabla\mathbf{u}]+\frac{\zeta}{2}[\nabla\mathbf{u}\cdot\boldsymbol{\tau}+\boldsymbol{\tau}\cdot(\nabla\mathbf{u})^t]$$
$$=\nabla\mathbf{u}+(\nabla\mathbf{u})^t \tag{3}$$

where $(\nabla\mathbf{u})^t$ denotes the transpose of $\nabla\mathbf{u}$. Equation (3) includes both lower and upper-convective terms. It is often referred to as the Johnson-Segalman model ([19]). Here $\zeta\in[0,2]$, which is a dimensionless material (slip) parameter. The value of $\zeta$ is a measure of the contribution of nonaffine motion to the shear tensor. For $\zeta=0$, the motion is affine and the Oldroyd-B model is recovered, whereas for $\zeta=2$, the motion is completely nonaffine and the model is reduced to the Oldroyd-Jaumann model ([4]). When $\zeta=0$ and $\eta_s=0$, the upper-convected Maxwell model is recovered.

If the $x$-axis is taken to lie halfway between the two plates, and $y$ is the coordinate in the transverse direction, then the total shear stress corresponding to the base (Poiseuille) flow is given by

$$T_{xy}^b=\varepsilon\dot{\gamma}+\frac{\dot{\gamma}}{1+\zeta(2-\zeta)\dot{\gamma}^2}=\mathrm{We}\,y \tag{4}$$

where $\dot{\gamma}=du/dy$ is the shear rate and $u$ is the velocity in the $x$-direction. Note that We is the dimensionless driving pressure gradient. Equation (4) is perhaps the most revealing result of the JS model. It reflects the possibility of a nonmonotonic behavior for the stress/shear-rate relation. Indeed, Fig. 1 shows the behavior of the shear stress, $T_{xy}^b$, as a function of $\dot{\gamma}$ for $\varepsilon\in[0,1]$ and $\zeta$

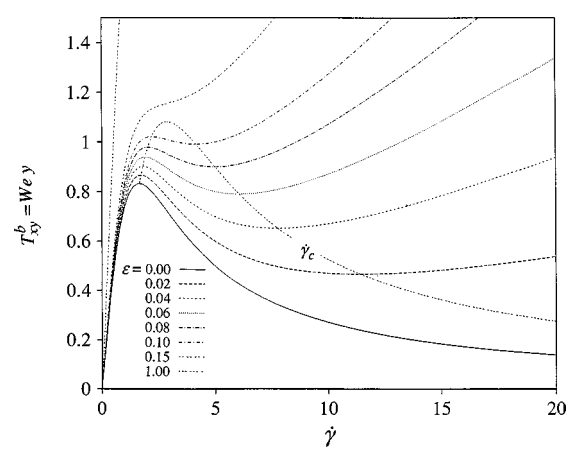



**Fig. 1 Steady-state shear stress versus shear-rate curves for $\zeta=0.2$ and $\varepsilon\in[0,1]$. The loci of the two extrema are also shown, which join into one curve denoted here by $\dot{\gamma}_c$. The curves in the figure resemble the pressure/stretch-ratio related to the inflation of a Mooney-Rivlin material (see Fig. 2 in [20]).**

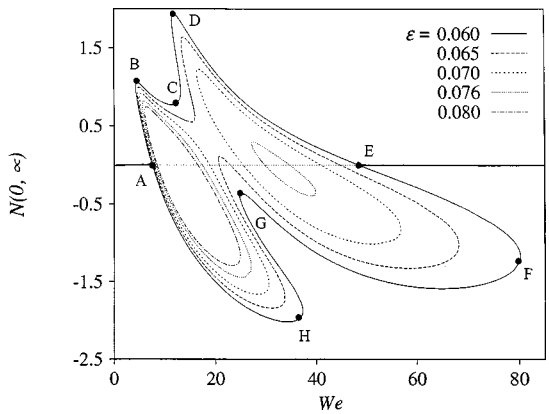



**Fig. 2 Bifurcation diagrams for the normal stress difference, $N(0,\infty)$, at the center of the channel as function of We for $\zeta=0.2$ and $\varepsilon\in[0.06,0.08]$. The smallest diagram corresponds to the highest viscosity ratio, $\varepsilon$. As $\varepsilon$ exceeds a critical level (in this case 1/8), the (closed) diagram reduces to zero, as the base flow is always stable. The branches AB, CD, EF, and GH of diagram $\varepsilon=0.06$ are unstable, whereas the branches BC, DE, EF, FG, and HA are stable.**

$=0.2$. The curve $\varepsilon=1$ corresponds essentially to the Newtonian limit. For this value of the viscosity ratio, the elastic contribution to $T^b_{xy}$ in expression (4) is negligible. In this case Newton's law of viscosity applies. The figure indicates that the stress curves lose monotonicity for $\varepsilon<1/8$. Two extrema (a maximum and a minimum) appear, which tend to merge as $\varepsilon$ increases, as indicated by the curve joining the loci of the extrema. The base flow is found to be unstable for the range of the stress curves with negative slope. This situation is reminiscent of the load/deformation behavior in elasticity. In the case of nonlinear inflation of a Mooney-Rivlin (hyperelastic) membrane, for instance, the pressure also exhibits a similar behavior as function of the stretch ratio for various Mooney constants ([20]). Upon comparison with the curves in Fig. 1, the curve $\varepsilon=0$ is comparable to that of a Neo-Hookean solid, while the curve for a Newtonian fluid ($\varepsilon=1$) is comparable to the curve of a Hookean solid (see Fig. 2 in [20]).

The solution of the system (1)–(2) is carried out using the Galerkin projection method. For one-dimensional disturbance along the channel ($x$-axis), the departure (from base flow) is reduced to the axial velocity, $u(y,t)$, normal stress difference, $N(y,t)$, and shear stress, $S(y,t)$. In this case, Eqs. (1)–(3) reduce to

$$\mathrm{Re}\, u_t=\varepsilon u_{yy}+S_y \tag{5a}$$

$$N_t=-N+2(\mathrm{We}\, S+Su_y+S^b u_y). \tag{5b}$$

$$S_t=-S+u_y+a(\mathrm{We}\, N+Nu_y+N^b u_y) \tag{5c}$$

where $a=\zeta(\zeta/2-1)$. Here $S^b=\dot\gamma/1+\zeta(2-\zeta)\dot\gamma^2$ is the non-Newtonian contribution of the shear stress of the base flow, and $N^b=2\dot\gamma^2/1+\zeta(2-\zeta)\dot\gamma^2$ is the corresponding first normal stress difference. Note that a subscript in Eqs. (5) denotes partial differentiation. The flow departure is represented by series of Chandrasekhar functions, which satisfy the homogeneous (no-slip) boundary conditions ([15]). A judicious selection process and truncation level is applied for the choice of the various modes in order to ensure the physical and mathematical coherence of the final model.

## 3 Bifurcation and Stability Picture

While the (linear) stability picture is somewhat predictable, the bifurcation picture is far from being intuitively obvious. The bifurcation diagrams depend strongly on $\varepsilon$ and $\zeta$. We thus monitor the influence of the viscosity ratio by fixing the parameter $\zeta$ to 0.2 and varying $\varepsilon$. Figure 2 displays the resulting bifurcation diagrams in the ($N$, We) plane for $\varepsilon\in[0.06,0.08]$. The figure shows the dependence of the steady-state normal stress difference, $N(0,\infty)$, at the center of the channel. Linear stability analysis asserts that, for large $\varepsilon$ value ($>1/8$), the base flow is stable (to small perturbations) for any value of We. This situation corresponds to a monotonic shear-stress/shear-rate curve in Fig. 1. As $\varepsilon$ decreases, two extrema appear in the stress curves in Fig. 1, entraining a loss of stability of the base flow in between. For each $\varepsilon>1/8$, a closed bifurcation diagram emerges as depicted in Fig. 2, which shows a widening of the unstable range of We values as the viscosity ratio decreases.

Although the case $\varepsilon=0.06$ will be discussed in detail below, we examine first the evolution of the bifurcation and stability picture as the flow deviates from close to the Newtonian limit (this limit is approached when the solvent-to-solute viscosity ratio, $\varepsilon$, is high). As $\varepsilon$ decreases below a critical value, two static bifurcations emerge at two critical values, $\mathrm{We}_{c1}$ and $\mathrm{We}_{c2}$, of the Weissenberg number as predicted by linear stability analysis. The two critical points coincide with points A and E for the $\varepsilon=0.06$ diagram. The two bifurcating branches join over the unstable range to form a closed diagram. This is clearly illustrated for $\varepsilon=0.08$; the closed diagram intersects the We axis at $\mathrm{We}_{c1}=8.48$ and $\mathrm{We}_{c2}=16.53$. As $\varepsilon$ decreases further, the (closed) diagram widens, and another closed diagram appears as depicted in Fig. 2 for $\varepsilon=0.076$. In this case, there are four critical values of the Weissenberg number that are present at 8.4, 17, 30, and 34.5. The second range of We values (30 to 34.5) corresponds to unstable base flow. A stable range exists between the two diagrams. As $\varepsilon$ decreases further, the two diagrams grow, come in contact with one another, and finally merge to form a simply connected closed diagram as shown in Fig. 2 for $\varepsilon=0.06$. In this case, the range of instability of the base flow becomes larger as it covers the values $7.54<\mathrm{We}<48.5$ (between A and E). The figure also indicates that the solution branch changes concavity, and presents regions alternating in stability.

In general, and as typically depicted by the $\varepsilon=0.06$ diagram, there is an exchange of stability at the two critical points $\mathrm{We}_{c1}$ (A) and $\mathrm{We}_{c2}$ (E), with the base flow losing its stability at $\mathrm{We}_{c1}$ and regaining it at $\mathrm{We}_{c2}$. However, the base flow is not always unconditionally stable for $\mathrm{We}<\mathrm{We}_{c1}$ and $\mathrm{We}>\mathrm{We}_{c2}$; simultaneously, the diagram $\varepsilon=0.06$ is not always unconditionally stable for $\mathrm{We}_{c1}<\mathrm{We}<\mathrm{We}_{c2}$. We have indicated in Fig. 2 the various branches of alternating stability of the $\varepsilon=0.06$ diagram. Thus, branches AB, CD, EF, and GH are unstable, while the branches BC, DE, FG, and HA are stable. Consequently, close to each critical point, just before $\mathrm{We}_{c1}$ and just after $\mathrm{We}_{c2}$, there is a branch, BC and FG, respectively, to which the flow can converge if the perturbation is not small, similarly to what occurs in the vicinity of transcritical and subcritical bifurcations. Although there are stable and unstable nontrivial branches in the range $\mathrm{We}_{c1}<\mathrm{We}<\mathrm{We}_{c2}$, there is total loss of stability of the base flow. In this range, only nonlinear velocity profiles are stable. The stability of the branches at the two critical points was established numerically since linear stability analysis cannot be applied in the vicinity of the critical (nonhyperbolic fixed) points.

It is, perhaps, at this stage that one begins to connect the stability and bifurcation picture to physical reality. It is well known that in real systems, physical instabilities are observed when the flow rate and/or the level of elasticity are high. These instabilities are believed to be potentially responsible for the onset of surface roughness in extrusion [10]. If we note that the flow rate is controlled by We, and the level of elasticity controlled by both We and $\varepsilon$, then we can clearly observe that the trend shown in Fig. 2 confirms that both the flow rate and fluid elasticity are the determining factors behind the destabilization of the base flow. It is also well known that instabilities are suspected to set in after the Weissenberg number has reached a certain value. This is also

inferred from Fig. 2, as the diagram corresponding to $\varepsilon = 0.06$, for instance, indicates that the base flow is practically unstable for the whole range We$>$7.5.

In summary, a nonlinear analysis is carried out to examine the onset of constitutive instability and bifurcation for polymer solutions. Only elastic effects, which lead to the well-known Weissenberg rod-climbing phenomenon, are responsible for the loss of stability of the base (Poiseuille) flow. The viscoelastic model used displays nonmonotonicity of the shear-stress/shear-rate curve, and belongs to the wider class of *Oldroyd* constitutive models that lead to the destabilization of channel flow. The bifurcation diagrams are obtained for the first time, and show how the secondary flow evolves as one deviates from the Newtonian limit. The bifurcation diagrams are always closed and widen in range as the solvent-to-solute viscosity decreases, thus reflecting the destabilization observed in practice as the level of elasticity increases. It is emphasized that the present stability and bifurcation picture corresponds to perturbations of infinite wavelength, which may not be the most dangerous modes. Only a higher-dimensional stability analysis can indicate whether the present findings are of physical relevance.


## Acknowledgments

We would like thank J. M. Floryan (UWO) for helpful comments on the manuscript. This work is funded by the Natural Sciences and Engineering Research Council of Canada.


## References


[1] Bird, R. B., Armstrong, R. C., and Hassager, O., 1987, *Dynamics of Polymeric Liquids*, Vol. 1, 2nd Ed., John Wiley and Sons, New York.

[2] Vinogradov, G. V., Malkin, A. Ya., Vanovskii, Yu G., Borisenkova, E. K., Yarlykov, B. V., and Berezheneya, G. V., 1972, J. Polym. Sci., Part A: Gen. Pap., **10**, p. 1061.

[3] Joseph, D. D., Renardy, M., and Saut, J. C., 1985, ''Hyperbolicity and Change of Type in the Flow of Viscoelastic Fluids,'' Arch. Ration. Mech. Anal., **87**, p. 213.

[4] Denn, M. M., 1990, ''Issues in Viscoelastic Fluid Mechanics,'' Annu. Rev. Fluid Mech., **22**, p. 13.

[5] Larson, R. G., 1992, ''Instabilities in Viscoelastic Flows,'' Rheol. Acta, **31**, p. 213.

[6] Kolkka, R. W., Malkus, D. S., Hansen, M. G., and Ierley, G. R., 1988, ''Spurt Phenomena of the Johnson-Segalman Fluid and Related Models,'' J. Non-Newtonian Fluid Mech., **29**, p. 303.

[7] Malkus, D. S., Nohel, J. A., and Plohr, B. J., 1990, ''Dynamics of Shear Flow of a Non-Newtonian Fluid,'' J. Comput. Phys., **87**, p. 464.

[8] Georgiou, G. C., and Vlassopoulos, D., 1998, ''On the Stability of the Simple Shear Flow of a Johnson-Segalman Fluid,'' J. Non-Newtonian Fluid Mech., **75**, p. 77.

[9] Sell, G. R., Foias, C., and Temam, R., 1993, *Turbulence in Fluid Flows: A Dynamical Systems Approach*, Springer-Verlag, New York.

[10] Khayat, R. E., 1994, ''Chaos and Overstability in the Thermal Convection of Viscoelastic Fluids,'' J. Non-Newtonian Fluid Mech., **53**, p. 227.

[11] Khayat, R. E., 1995, ''Nonlinear Overstability in the Thermal Convection of Viscoelastic Fluids,'' J. Non-Newtonian Fluid Mech., **58**, p. 331.

[12] Khayat, R. E., 1995, ''Fluid Elasticity and Transition of Chaos in Thermal Convection,'' Phys. Rev. E, **51**, p. 380.

[13] Avgousti, M., and Beris, A. N., 1993, ''Non-Axisymmetric Subcritical Bifurcations in Viscoelastic Taylor-Couette Flow,'' Proc. R. Soc. London, Ser. A, **A443**, p. 17.

[14] Khayat, R. E., 1995, ''Onset of Taylor Vortices and Chaos in Viscoelastic Fluids,'' Phys. Fluids A, **7**, p. 2191.

[15] Khayat, R. E., 1997, ''Low-Dimensional Approach to Nonlinear Overstability of Purely Elastic Taylor-Vortex Flow,'' Phys. Rev. Lett., **78**, p. 4918.

[16] Graham, M. D., 1998, ''Effect of Axial Flow on Viscoelastic Taylor-Couette Instability,'' J. Fluid Mech., **360**, p. 341.

[17] Ashrafi, N., and Khayat, R. E., 2000, ''Finite Amplitude Taylor-Vortex Flow of Weakly Shear-Thinning Fluids,'' Phys. Rev. E, **61**, p. 1455.

[18] Muller, S. J., Shaqfeh, E. S. J., and Larson, R. G., 1993, ''Experimental Study of the Onset of Oscillatory Instability in Viscoelastic Taylor-Couette Flow,'' J. Non-Newtonian Fluid Mech., **46**, p. 315.

[19] Johnson, M. W., and Segalman, D., 1977, ''A Model for Viscoelastic Fluid Behavior Which Allows Non-Affine Deformation,'' J. Non-Newtonian Fluid Mech., **2**, p. 278.

[20] Khayat, R. E., and Derdouri, A., 1994, ''Inflation of Hyperelastic Cylindrical Membranes as Applied to Blow Moulding, Part I. Axisymmetric Case,'' Int. J. Numer. Methods Eng., **37**, p. 3773.


# In-Plane Gravity Loading of a Circular Membrane


**R. O. Tejeda**
Research Assistant

**E. G. Lovell**
Professor, Mem. ASME

**R. L. Engelstad**
Professor, Mem. ASME

Department of Mechanical Engineering, University of Wisconsin-Madison, Madison, WI 53706-1572



*This paper develops the displacement field for a circular membrane which is statically loaded by gravity acting in its plane. Coupled to the displacements are the stress and strain distributions. The solution is applicable to the modeling of next generation lithographic masks, ion-beam projection lithography masks in particular.* [S0021-8936(00)00803-5]


## 1 Introduction

In most engineering applications, the displacement, stress, and strain fields induced by gravity are negligible. However, in next generation (nonoptical) lithography masks used for semiconductor device fabrication, it is critical to predict and compensate for distortions which could potentially alter the quality of the microcircuit that is to be manufactured. Typically, the allowable error in a lithographic mask is only a fraction of the microcircuit's minimum feature size ([1]). Since ion-beam projection lithography (IPL) is targeting the production of sub-100 nm devices, displacements due to gravity can be significant.

An IPL mask is composed of a circular membrane that is supported by a relatively stiff frame and held in a vertical orientation during exposure ([2]). It is typically made of silicon with a diameter on the order of 200 mm and 3.0-$\mu$m thickness. If the mask is modeled as a circular membrane that is constrained on its perimeter by a rigid ring and subjected to in-plane gravitational loading (as shown in Fig. 1), it can be considered as a plane stress problem and solved directly by traditional applied elasticity methods. To the best of the authors' knowledge, a solution to this problem has not been presented in the elasticity literature.

## 2 Solution Development

The position of an arbitrary point on the membrane is defined by the polar coordinates $(r,\theta)$ with the origin taken at the center. All translational displacement components are constrained at the outside radius, $R$. In general, radial $(u)$ and circumferential $(v)$ displacements which arise from the loading of the membrane are related to the radial strain $(\varepsilon_r)$, circumferential strain $(\varepsilon_\theta)$, and shear strain $(\gamma_{r\theta})$ by strain-displacement equations, and to radial normal stress $(\sigma_r)$, circumferential normal stress $(\sigma_\theta)$, and shear stress $(\tau_{r\theta})$ by Hooke's law, i.e.,

$$\varepsilon_r = \frac{\partial u}{\partial r} = \frac{1}{E}(\sigma_r - \nu\sigma_\theta) \tag{1}$$


Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Apr. 28, 1999; final revision, May 5, 2000. Associate Technical Editor: R. C. Benson.




**Copyright © 2000 by ASME**

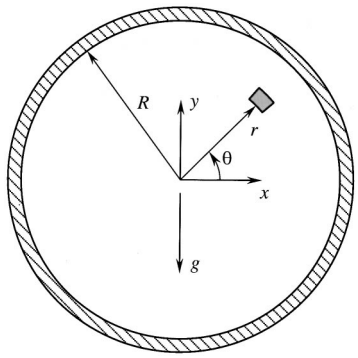



**Fig. 1 In-plane gravity loading of a circular membrane**

$$\varepsilon_\theta=\frac{u}{r}+\frac{1}{r}\frac{\partial v}{\partial\theta}=\frac{1}{E}(\sigma_\theta-\nu\sigma_r) \tag{2}$$

$$\gamma_{r\theta}=\frac{1}{r}\frac{\partial u}{\partial\theta}+\frac{\partial v}{\partial r}-\frac{v}{r}=\frac{2(1+\nu)}{E}\tau_{r\theta}. \tag{3}$$

Here, $E$ is Young's modulus and $\nu$ is Poisson's ratio. Equilibrium of the shaded element shown in Fig. 1 requires that

$$\frac{\sigma_r-\sigma_\theta}{r}+\frac{\partial\sigma_r}{\partial r}+\frac{1}{r}\frac{\partial\tau_{r\theta}}{\partial\theta}-\rho g\sin\theta=0 \tag{4}$$

$$\frac{1}{r}\frac{\partial\sigma_\theta}{\partial\theta}+\frac{\partial\tau_{r\theta}}{\partial r}+\frac{2}{r}\tau_{r\theta}-\rho g\cos\theta=0 \tag{5}$$

where $\rho$ and $g$ are the material mass density and gravitational acceleration, respectively.

A solution form which uses Airy's stress function, $\phi$, can be expressed as

$$\sigma_r=\frac{1}{r}\frac{\partial\phi}{\partial r}+\frac{1}{r^2}\frac{\partial^2\phi}{\partial\theta^2}+F(r,\theta) \tag{6}$$

$$\sigma_\theta=\frac{\partial^2\phi}{\partial r^2}+G(r,\theta) \tag{7}$$

$$\tau_{r\theta}=-\frac{\partial}{\partial r}\left(\frac{1}{r}\frac{\partial\phi}{\partial\theta}\right) \tag{8}$$

where $F(r,\theta)$ and $G(r,\theta)$ are arbitrary functions. Substituting Eqs. (6)–(8) into the equilibrium Eqs. (4) and (5) and subsequent integration yields

$$G(r,\theta)=\rho g r\sin\theta+G_2(r) \tag{9}$$

$$F(r,\theta)=\rho g r\sin\theta+\frac{F_2(r)}{r}+\frac{1}{r}\int G_2(r). \tag{10}$$

To preclude singular stresses, and without loss of generality, let $F_2(\theta)=G_2(r)=0$. Therefore,

$$\sigma_r=\frac{1}{r}\frac{\partial\phi}{\partial r}+\frac{1}{r^2}\frac{\partial^2\phi}{\partial\theta^2}+\rho g r\sin\theta \tag{11}$$

$$\sigma_\theta=\frac{\partial^2\phi}{\partial r^2}+\rho g r\sin\theta. \tag{12}$$

General forms of $\phi$ satisfying compatibility were originally given by Michell [3] for problems described in polar coordinates. From Timoshenko's ([4]) summary of this,

$$\begin{aligned}\phi=&a_0\log r+b_0r^2+c_0r^2\log r+d_0r^2\theta+a_0'\theta+\frac{a_1}{2}r\theta\sin\theta\\&+(b_1r^3+a_1'r^{-1}+b_1'r\log r)\cos\theta-\frac{c_1}{2}r\theta\cos\theta\\&+(d_1r^3+c_1'r^{-1}+d_1'r\log r)\sin\theta\\&+\sum_{n=2}^{\infty}(a_nr^n+b_nr^{n+2}+a_n'r^{-n}+b_n'r^{-n+2})\cos n\theta\\&+\sum_{n=2}^{\infty}(c_nr^n+d_nr^{n+2}+c_n'r^{-n}+d_n'r^{-n+2})\sin n\theta.\end{aligned} \tag{13}$$

A stress function which gives nonperiodic or singular stresses is not admissible. Therefore, all the terms except those with coefficients of $b_0$, $b_1$, $d_1$, $a_n$, $b_n$, $c_n$, and $d_n$ will be dropped. So, for $n=2,3,4,\ldots$

$$\phi\propto r^n(\sin n\theta,\cos n\theta) \tag{14}$$

$$(\sigma_r,\sigma_\theta)\propto r^{n-2}(\sin n\theta,\cos n\theta) \tag{15}$$

$$\tau_{r\theta}\propto r^{n-2}(\cos n\theta,\sin n\theta). \tag{16}$$

On a circular edge, Eqs. (14)–(16) describe normal and shear stresses which vary harmonically. They are not necessary to solve the basic problem. Indeed, they may lead to displacements which cannot be zero at the boundary $r=R$. Thus, a solution may be constructed using relevant $\phi$'s established by Michell, plus additional terms associated with the gravitational body forces:

$$\phi=b_0r^2+b_1r^3\cos\theta+d_1r^3\sin\theta. \tag{17}$$

By the definition established in Eq. (8),

$$\tau_{r\theta}=2b_1r\sin\theta-2d_1r\cos\theta. \tag{18}$$

Due to the symmetric nature of the problem, $\tau_{r\theta}(r,\pi/2)=0$. Therefore, $b_1=0$. Similarly, Eqs. (11) and (12) show that

$$\sigma_r=2b_0+2d_1r\sin\theta+\rho g r\sin\theta \tag{19}$$

$$\sigma_\theta=2b_0+6d_1r\sin\theta+\rho g r\sin\theta. \tag{20}$$

Because it is associated with a hydrostatic loading such as uniform membrane prestress, $b_0$ will also be discarded. This leaves one unknown, $d_1$, which can be identified by noting that the circumferential normal strain is zero on the boundary, i.e., $\varepsilon_\theta(R,\theta)=0$. By using Eqs. (19), (20), and (2), this yields

$$d_1=\frac{-\rho g}{2}\frac{(1-\nu)}{(3-\nu)} \tag{21}$$

and the stresses can now be written in the form

$$\sigma_r=\frac{2\rho g}{(3-\nu)}r\sin\theta \tag{22}$$

$$\sigma_\theta=\frac{2\rho g\nu}{(3-\nu)}r\sin\theta \tag{23}$$

$$\tau_{r\theta}=\rho g\frac{(1-\nu)}{(3-\nu)}r\cos\theta. \tag{24}$$

Employing Eqs. (1)–(3), boundary conditions, and symmetry conditions gives the following strains and displacements in the membrane:

$$\varepsilon_r(r,\theta)=\frac{2\rho g}{E}\frac{(1-\nu^2)}{(3-\nu)}r\sin\theta \tag{25}$$

$$\varepsilon_\theta(r,\theta)=0 \tag{26}$$

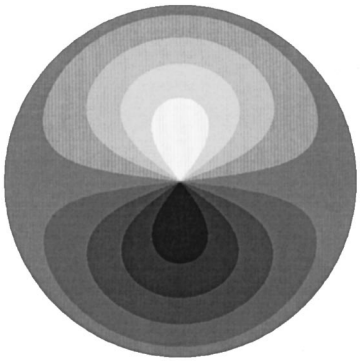


**Fig. 2 Radial displacement contours. Range is $\pm\rho g(1-\nu^2)R^2/E(3-\nu)$.**

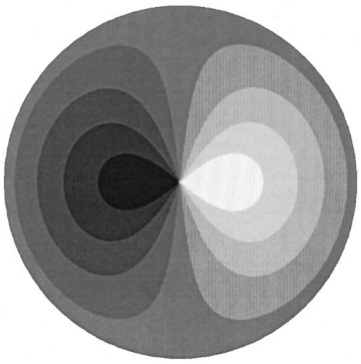


**Fig. 3 Circumferential displacement contours. Range is $\pm\rho g(1-\nu^2)R^2/E(3-\nu)$.**

$$\gamma_{r\theta}(r,\theta)=\frac{2\rho g}{E}\frac{(1-\nu^2)}{(3-\nu)}r\cos\theta \tag{27}$$

$$u=\frac{\rho g}{E}\frac{(1-\nu^2)}{(3-\nu)}(r^2-R^2)\sin\theta \tag{28}$$

$$v=\frac{\rho g}{E}\frac{(1-\nu^2)}{(3-\nu)}(r^2-R^2)\cos\theta. \tag{29}$$

The displacement component distributions given by Eqs. (28) and (29) are illustrated in Figs. 2 and 3.

Transformation of the stress components into an $x$-$y$ coordinate system yields

$$\sigma_x=\frac{2\rho g\nu}{(3-\nu)}y \tag{30}$$

$$\sigma_y=\frac{2\rho g}{(3-\nu)}y \tag{31}$$

$$\tau_{xy}=\frac{\rho g(1-\nu)}{(3-\nu)}x; \tag{32}$$

where $x$ and $y$ are Cartesian coordinates, $\sigma_x$ is the normal stress in the $x$-direction, $\sigma_y$ is the normal stress in the $y$-direction, and $\tau_{xy}$ is the in-plane shear stress. As required, these stresses satisfy equilibrium equations in $x$ and $y$, and the associated displacement components vanish on a circular boundary.

## 3 Conclusion

An elasticity solution for the in-plane gravitational loading of a circular membrane has been established. Stress, strain, and displacement fields satisfy symmetry, boundary conditions, equilibrium, compatibility for a linear elastic material, and are singularity-free.

The solution is useful for predicting the distortion of masks used in next generation lithography processes. It is particularly appropriate to modeling the ion-beam lithography mask, which consists of a large circular membrane supported by a frame or wafer ring while it is in a vertical orientation.


## Acknowledgments

This research was supported by International SEMATECH, the Semiconductor Research Corporation (SRC), and the National Science Foundation Graduate Research Fellowship Program.


## References


[1] Gross, G., 1997, ''Ion Projection Lithography: Next Generation Technology?,'' J. Vac. Sci. Technol., B**15**, No. 6, pp. 2136–2138.

[2] Tejeda, R., Engelstad, R., Lovell, E., and Berry, I., 1998, ''Analysis, Design, and Optimization of Ion-Beam Lithography Masks,'' Proc. SPIE, Emerging Lithographic Technologies II, **3331**, pp. 621–628.

[3] Michell, J. H., 1899, ''On the Direct Determination of Stress in an Elastic Solid, With Application to the Theory of Plates,'' Proc. London Math. Soc., **31**, pp. 100–124.

[4] Timoshenko, S. P., and Goodier, J. N., 1970, *Theory of Elasticity*, 3rd Ed., McGraw-Hill, New York, pp. 132–135.


# Free Vibration of a Spinning Stepped Timoshenko Beam


**S. D. Yu**
Assistant Professor, Mem. ASME, Department of Mechanical Engineering, Ryerson Polytechnic University, 350 Victoria Street, Toronto, ON M5B 2K3, Canada

**W. L. Cleghorn**
Professor, Department of Mechanical and Industrial Engineering, University of Toronto, Five King's College Road, Toronto, ON M5S 3G8, Canada



*The finite element method is employed in this paper to investigate free-vibration problems of a spinning stepped Timoshenko beam consisting of a series of uniform segments. Each uniform segment is considered a substructure which may be modeled using beam finite elements of uniform cross section. Assembly of global equation of motion of the entire beam is achieved using Lagrange's multiplier method. The natural frequencies and mode shapes are subsequently reduced with the help of linear transformations to a standard eigenvalue problem for which a set of natural frequencies and mode shapes may be easily obtained. Numerical results for an overhung stepped beam consisting of three uniform segments are obtained and presented as an illustrative example.*
[S00021-8936(01)00101-5]


## 1 Introduction

In studying machine tool vibration during a turning cutting process, it is often necessary to conduct free-vibration analysis of


Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received and accepted by the ASME Applied Mechanics Division, Sept. 1, 1999; final revision, Apr. 10, 2000. Associate Technical Editor: N. C. Perkins.




**Copyright © 2000 by ASME**

spinning stepped shaft. Bauer [1] presented an analytical study of a rotating uniform Euler-Bernoulli beam with various combinations of simple boundary conditions. Lee et al. [2] studied the free vibration of a rotating Rayleigh shaft using the modal analysis approach and Galerkin's method. Katz et al. [3] investigated the dynamic responses of a uniform rotating shaft subjected to a moving load in the axial direction using both Rayleigh and Timoshenko beam theories. Zu and Han [4] presented analytical solutions for free vibration of a spinning uniform Timoshenko beam with all combinations of the three classical boundary conditions.

In this paper, free vibration of a rotating stepped Timoshenko beam is investigated using the finite element method. To enhance the accuracy of the computed eigenvalues and mode shapes, a three-node beam element, which permits the use of quintic polynomials as the interpolation function for both lateral displacements and bending angles, is utilized. For each of field variable, six nodal quantities—the variable and its derivative with respect to the axial coordinate at all three nodes are introduced to the element displacement vector. Two lateral deflections and two bending angles at each axial location need be defined in beam flexural vibration. Therefore, a three-node beam element has 24 nodal variables.

Use of the finite element method makes it possible to reduce the free-vibration problem of a spinning beam to a standard eigenvalue problem for which all eigenvalues and eigenvectors may be determined simultaneously. This is one advantage over the use of an analytical method in which eigenvalues are determined by searching for the roots of a characteristic equation. As an illustrative example, natural frequencies of an overhung stepped are obtained for several different spin rates.

## 2 Mathematical Procedure

In this section, the equations of motion of a stepped beam of circular cross section, as shown in Fig. 1, are presented using an inertial coordinate system. Parameters defining each segment are length, diameter, and axial coordinate of the left-most plane.

**2.1 Governing Differential Equations for a Substructure.** In modeling the stepped beam, each segment is considered a substructure. Within a substructure, the equations of motion may be written as ([3])

$$\left[\begin{bmatrix}\mathbf{L}_2 & \mathbf{0}\\ \mathbf{0} & \mathbf{L}_2\end{bmatrix}\frac{\partial^2}{\partial t^2}+2\Omega\begin{bmatrix}\mathbf{0} & \mathbf{L}_1\\ -\mathbf{L}_1 & \mathbf{0}\end{bmatrix}\frac{\partial}{\partial t}+\begin{bmatrix}\mathbf{L}_0 & \mathbf{0}\\ \mathbf{0} & \mathbf{L}_0\end{bmatrix}\right]\begin{Bmatrix}u_x\\ \psi_x\\ u_y\\ \psi_y\end{Bmatrix}=\begin{Bmatrix}0\\0\\0\\0\end{Bmatrix} \tag{1}$$

where $u_x$ and $u_y$ are lateral displacements of the beam centroid in the $x$ and $y$ directions, respectively; $\psi_x$ and $\psi_y$ are angles of rotation of the plane normal to the beam centroid, measured in the $xoz$ and $yoz$ coordinate planes, respectively; the three operator matrices are defined as

$$\mathbf{L}_2=\begin{bmatrix}\rho A & \mathbf{0}\\ \mathbf{0} & \rho I\end{bmatrix},\quad \mathbf{L}_1=\begin{bmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & \rho I\end{bmatrix},$$
$$\mathbf{L}_0=\begin{bmatrix}-\kappa GA\dfrac{\partial^2}{\partial z^2} & \kappa GA\dfrac{\partial}{\partial z}\\ -\kappa GA\dfrac{\partial}{\partial z} & \kappa GA-EI\dfrac{\partial^2}{\partial z^2}\end{bmatrix}. \tag{2}$$

In Eq. (2), $\rho$ is the volume mass density of the beam material; $G$ is the shear modulus; $E$ is the modulus of elasticity; $A$ is the cross-sectional area; $I$ is the second moment of area; $\kappa$ is the shear correction factor (0.9 for solid circular cross section, [4]); $\Omega$ is the spin rate.

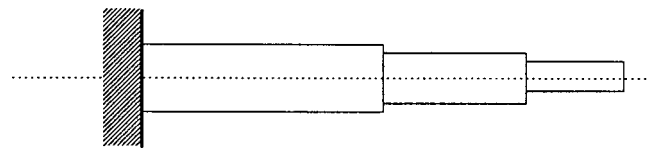


**Fig. 1 An overhung stepped shaft**

Assume that a uniform segment $k$ is modeled using $N_{e,k}$ ($k=1,2,\ldots,N_s$) three-node beam elements. Within each element, the displacement vector $\{\mathbf{u}_e\}_k$ is related to the nodal displacement vector $\{\mathbf{q}_e\}_k$ by

$$\{\mathbf{u}_e\}_k=[\mathbf{N}_e(\xi)]\{\mathbf{q}_e\}_k \quad (0\leqslant\xi\leqslant l_e) \tag{3}$$

where $\xi$ is the local coordinate; $\mathbf{N}_e(\xi)$ is the shape function matrix. The equations of motion of a uniform segment in terms of the nodal coordinate vector may be easily derived using the minimum potential energy principle or the Galerkin principle.

**2.2 Assembly of Equations of Motion for the Stepped Beam.** The global equations of motion for the entire beam may be formulated by enforcing continuity conditions across each interface between two adjacent substructures. The procedure is illustrated here for a single interface at node $j$ with $z=z_j$. The four displacement and four force continuity conditions may be written in terms of the nodal displacements as

$$\begin{Bmatrix}u_x\\ \theta_x\\ \psi_x\\ k_x\\ u_y\\ \theta_y\\ \psi_y\\ k_y\end{Bmatrix}_{z=z_j^-}=\begin{bmatrix}1&0&0&0&0&0&0&0\\0&\alpha&1-\alpha&0&0&0&0&0\\0&0&1&0&0&0&0&0\\0&0&0&\beta&0&0&0&0\\0&0&0&1&0&0&0&0\\0&0&0&0&0&\alpha&1-\alpha&0\\0&0&0&0&0&0&1&0\\0&0&0&0&0&0&0&\beta\end{bmatrix}$$
$$\times\begin{Bmatrix}u_x\\ \theta_x\\ \psi_x\\ k_x\\ u_y\\ \theta_y\\ \psi_y\\ k_y\end{Bmatrix}_{z=z_j^+}=\mathbf{C}_j\begin{Bmatrix}u_x\\ \theta_x\\ \psi_x\\ k_x\\ u_y\\ \theta_y\\ \psi_y\\ k_y\end{Bmatrix}_{z=z_j^+} \tag{4}$$

where $\alpha$ and $\beta$ are the shear rigidity ratio and the bending rigidity ratio, defined as

$$\alpha=\frac{(\kappa GA)_1}{(\kappa GA)_2},\quad \beta=\frac{(EI)_1}{(EI)_2}. \tag{5}$$

The continuity conditions may be implemented into the global equations of motion using the Lagrange multiplier method ([5]). The final equations of motion of a stepped beam may be written as

$$[\mathbf{M}]\{\ddot{\mathbf{q}}\}+[\mathbf{G}]\{\dot{\mathbf{q}}\}+[\mathbf{K}]\{\mathbf{q}\}=\mathbf{0} \tag{6}$$

where the global mass, gyroscopic and stiffness matrices for the entire stepped beam may be formulated from the corresponding matrices for substructures and constraint matrices $\mathbf{C}_j$. For example, the global mass matrix for a three-segment stepped beam is

$$[\mathbf{M}]=\begin{bmatrix} \mathbf{M}_{11}^1 & \mathbf{M}_{12}^1 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{M}_{21}^1 & \mathbf{M}_{22}^1+\mathbf{C}_1^T\mathbf{M}_{11}^2\mathbf{C}_1 & \mathbf{C}_1^T\mathbf{M}_{12}^2 & \mathbf{M}_{13}^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_{21}^2\mathbf{C}_1 & \mathbf{M}_{22}^2 & \mathbf{M}_{23}^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_{31}^2 & \mathbf{M}_{32}^2 & \mathbf{M}_{33}^2+\mathbf{C}_2^T\mathbf{M}_{11}^3\mathbf{C}_2 & \mathbf{C}_2^T\mathbf{M}_{12}^3 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{M}_{21}^3\mathbf{C}_2 & \mathbf{M}_{22}^3 \end{bmatrix}.$$

**2.3 Boundary Conditions.** The global equations of motion of a stepped beam may be readily modified to account for boundary conditions at the ends using the penalty approach. In this paper, an overhung shaft clamped at one end and free at the other is investigated.

## 3 Standard Eigenvalue Problem

One of the most important tasks in free-vibration analysis is to determine the natural frequencies and mode shapes described by a system of homogeneous second-order differential equations and boundary conditions at the ends of a beam. To utilize commercial routines for eigenanalysis, the second-order system may be replaced by the following equivalent first-order system

$$\mathbf{A}\dot{\mathbf{x}}=\mathbf{B}\mathbf{x} \tag{7}$$

where coefficient matrices **A** and **B** may be obtained from matrices **M**, **K**, and **G**.

Natural frequencies and mode shapes of a spinning Timoshenko beam may then be reduced to seeking $\lambda$ and **X** satisfying

$$\mathbf{A}\mathbf{X}=\frac{1}{\lambda}\mathbf{B}\mathbf{X}. \tag{8}$$

## 4 Numerical Results

Values of the material and geometric properties of a three segment stepped beam under investigation are given in Table 1. Table 2 provides values of natural frequencies of an overhung three segment stepped beam shown in Fig. 1. Five different spin rates are considered. For the case of zero spin rate, natural frequencies have matched pairs. For a nonzero spin rate, there are two natural frequencies developed around each at-rest value. One is associated with the forward precession mode; the other is associated with the backward precession mode.

**Table 1 Material and geometric properties of a stepped spinning beam**

| | Geometric Properties | | | Material Properties | | |
|---|---|---|---|---|---|---|
| | $L_k$ (m) | $d_k$ (m) | $\kappa_k$ | $E_k$ (GPa) | $G_k$ (GPa) | $\rho_k$ (kg/m$^3$) |
| Segment 1 | 0.3 | 0.191 | 0.9 | 207 | 77.6 | 7,700 |
| Segment 2 | 0.3 | 0.150 | 0.9 | 207 | 77.6 | 7,700 |
| Segment 3 | 0.4 | 0.100 | 0.9 | 207 | 77.6 | 7,700 |

**Table 2 The first six natural frequencies of an overhung stepped beam**

| | Rotational Speed (rad/s) | | | | |
|---|---|---|---|---|---|
| Modes | 0 | 2,500 | 5,000 | 7,500 | 10,000 |
| 1a | 1242.2 | 1224.2 | 1206.3 | 1188.7 | 1171.3 |
| 1b | 1242.2 | 1260.4 | 1278.9 | 1297.5 | 1316.4 |
| 2a | 3932.0 | 3867.3 | 3803.1 | 3739.3 | 3676.2 |
| 2b | 3932.0 | 3997.1 | 4062.6 | 4128.4 | 4194.4 |
| 3a | 9362.8 | 9216.9 | 9071.9 | 8928.1 | 8785.4 |
| 3b | 9362.8 | 9509.5 | 9656.7 | 9804.5 | 9952.4 |

## 5 Conclusions

This paper presents a finite element analysis of free vibration of a spinning stepped Timoshenko beam. Through use of the Lagrange multipliers method, both displacement continuity and force equilibrium conditions are satisfied at the interfaces of joining substructures. Because of the use of higher-order beam finite elements for each substructure, highly accurate natural frequencies of a stepped beam of any desired mode may be obtained.

## References


[1] Bauer, H. F., 1980, ''Vibration of a Rotating Uniform Beam, Part 1: Orientation in the Axis of Rotation,'' J. Sound Vib., **72**, No. 2, pp. 177–189.

[2] Lee, C. W., Katz, R., Ulsoy, A. G., and Scott, R. A., 1988, ''Modal Analysis of a Distributed Parameter Rotating Beam,'' J. Sound Vib., **122**, No. 1, pp. 119–130.

[3] Katz, R., Lee, C. W., Ulsoy, A. G., and Scott, R. A., 1988, ''Dynamic Responses of a Rotating Shaft Subjected to a Moving Load,'' J. Sound Vib., **122**, No. 1, pp. 131–148.

[4] Zu, J. W., and Han, R. P. S., 1992, ''Natural Frequencies and Normal Modes of a Spinning Timoshenko Beam With General Boundary Conditions,'' ASME J. Appl. Mech., **59**, pp. S197–S203.

[5] Tabarrok, B., and Rimrott, F. P. J., 1994, *Variational Methods and Complementary Formulations in Dynamics*, Kluwer, Dordrecht, The Netherlands.


# Penetration Experiments With Limestone Targets and Ogive-Nose Steel Projectiles


**D. J. Frew**
Assoc. Mem. ASME, Waterways Experiment Station,
Vicksburg, MS 39180-6199

**M. J. Forrestal**
Fellow ASME, Sandia National Laboratories,
Albuquerque, NM 87185-0303

**S. J. Hanchak**
University of Dayton Research Institute,
Dayton, OH 45469-0182



*We conducted three sets of depth-of-penetration experiments with limestone targets and 3.0 caliber-radius-head (CRH), ogive-nose steel rod projectiles. The ogive-nose rod projectiles with length-to-diameter ratios of ten were machined from 4340 $R_c$ 45 and Aer Met 100 $R_c$ 53 steel, round stock and had diameters and masses of 7.1 mm, 0.020 kg; 12.7 mm, 0.117 kg; and 25.4 mm, 0.931 kg.*



Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Nov. 29, 1999; final revision, May 10, 2000. Associate Technical Editor: I. M. Daniel.




**Copyright © 2000 by ASME**

*Powder guns or a two-stage, light-gas gun launched the projectiles at normal impacts to striking velocities between 0.4 and 1.9 km/s. In addition, we present an analytical penetration equation that described the target resistance by its density and a strength parameter determined from depth of penetration versus striking velocity data.* [S0021-8936(01)00201-X]

## Introduction

Several authors have written review articles that discuss the many analytical, computational, and experimental methods used to study the broad field of penetration mechanics ([1–5]). The responses of the projectiles and targets depend strongly on the problem geometry, materials, and impact conditions. Because many penetration mechanisms are possible, experimental observations usually precede and guide analytical or computational models. For this study, post-test observations showed a conical entry crater with a depth of two or more projectile diameters followed by a circular penetration channel or tunnel with nearly the projectile diameter. For recent penetration studies with 6061-T6511 aluminum targets ([6,7]) we could obtain post-test radiographs of the penetration channels. However, we could obtain post-test observations of the limestone penetration channels only after the targets were split with the techniques used by stone masons. The projectiles recovered from the targets had small mass losses caused by abrasion, but the overall nose shapes before and after penetration looked very similar. We previously observed similar post-test, target channels and abraded projectiles with our studies on concrete targets ([8,9]). However, the concrete targets abraded the nose tips more severely than the limestone targets.

Based on penetration data sets with three projectile scales, we present an analytical penetration equation that describes target resistance by a parameter determined from penetration depth versus striking velocity data. Our limestone penetration equation is similar to that previously derived for concrete targets ([10]). However, for limestone targets, we observed a noticeable decrease in the target resistance parameter as the projectile diameters increased. So for the limestone targets, the target resistance parameter is equal to a constant term plus a term that depends on the projectile diameter. We hypothesize that the penetration equation will be reasonably accurate for larger scale projectiles, but data from much more expensive field tests must be obtained to confirm our hypothesis.

The limestone targets were quarried and cut by the Elliot Stone Company.[1] In the rock mechanics literature, this particular limestone is often called Salem, Indiana, or Bedford limestone. For this study, we conducted unconfined compression tests and some triaxial compression tests on samples cored from representative blocks and from individual targets before the penetration tests. Material properties from our targets are nearly the same as those reported by Fossum, Senseny, Pfeifle, and Mellegard [11].

As previously mentioned, our penetration equation contains a target strength constant that is determined from penetration depth versus striking velocity data. While this methodology provides an accurate and convenient engineering equation, the detailed response mechanisms for the target are not modeled. The authors are not aware of any rigorous target models for rock penetration problems, but Lagrangian computational models that use adaptive meshing techniques have shown promise for brittle ceramic targets ([12,13]). Detailed computational approaches that model target responses also require a broad array of quasi-static and dynamic material properties data. For limestone, examples of some materials experiments and data include (1) quasi-static, triaxial compression experiments ([11]), (2) split Hopkinson bar experiments ([14–16]), (3) shock wave studies ([17]), (4) dynamic tensile failure with planar-impact techniques ([18,19]), and (5) compression-shear loading with plate impact experiments ([20]). Data from other experimental techniques may also be required for a careful target analysis.

In the next sections, we present the penetration model, describe the experiments, and present our results and conclusions.

## Penetration Model

For both limestone and concrete targets, post-test observations showed a conical entry crater with a depth of two or more projectile diameters followed by a circular channel or tunnel with nearly the projectile diameter. The limestone penetration equations are similar to the previously published concrete penetration equations. From Forrestal, Altman, Cargile, and Hanchak [10], depth of penetration $P$ for an ogive-nose projectile and a concrete target is given by

$$P=\frac{m}{2\pi a^2\rho N}\ln\left(1+\frac{N\rho V_1^2}{R}\right)+4a,\quad P>4a \tag{1a}$$

$$N=\frac{8\psi-1}{24\psi^2},\quad V_1^2=\frac{mV_s^2-4\pi a^3R}{m+4\pi a^3N\rho} \tag{1b}$$

in which the ogive-nose rod projectile is described by mass $m$, shank radius $a$, and caliber-radius-head $\psi$. The target is described by density $\rho$ and the target strength constant $R$. The strength constant is determined from

$$R=\frac{N\rho V_s^2}{\left(1+\frac{4\pi a^3N\rho}{m}\right)\exp\left[\frac{2\pi a^2(P-4a)N\rho}{m}\right]-1} \tag{2}$$

where $V_s$ is striking velocity. For a set of experiments, we hold all parameters constant and vary striking velocity. From each experiment, we measure striking velocity $V_s$ and penetration depth $P$, so $R$ can be determined from (2) for each experiment. We then take the average value of $R$ from the data set and compare the prediction from (1) with the measured values of $V_s$ and $P$. For this study with limestone targets, ogive-nose steel rod projectiles with 7.1, 12.7, and 25.4-mm diameters and a length-to-diameter ratio of ten have values of $R=913$, 787, and 693 MPa, respectively. Thus, the target resistance decreases as the projectile shank diameter increases. We found that for these limestone targets

$$R=K+k(2a_o/2a) \tag{3}$$

in which $K$ and $k$ are constants obtained from data fits, $2a_o$ is a reference projectile diameter, and $2a$ is the projectile diameter. We show later that with $K=607$ MPa, $k=86$ MPa, and $2a_o=25.4$ mm, we accurately recover the measured values of $R$ for each of the three data sets.

In summary, the procedure used to calculate $R$ from penetration depth data for a fixed projectile is the same for concrete or limestone targets. However, for limestone targets, $R$ depends on the projectile shank diameter. Thus, we can use the penetration equations (1*a*) and (1*b*) for limestone when $R$ is given by (3).

## Experiments

We conducted three sets of penetration experiments (a total of 30 experiments) with ogive-nose steel rod projectiles and limestone targets. All projectiles had a total length-to-diameter ratio of ten and 3.0 caliber-radius-head (CRH) nose shapes. The shank diameters and masses for each of the three sets of experiments were 7.1 mm, 0.020 kg; 12.7 mm, 0.117 kg; and 25.4 mm, 0.931 kg.

**Limestone Targets.** The limestone targets were quarried and cut by the Elliot Stone Company of Bedford, IN. We obtained the targets in three batches from nearby sites. Nominal material properties for the three target batches are given in Table 1 and show minimal variations among the batches. In addition, we conducted

[1]Elliot Stone Company, 3326 Mitchell Road, Bedford, IN 47421.

**Table 1 Nominal material properties for the limestone targets**

| Batch number | Density $\rho$ (Mg/m$^3$) | Water content w (%) | Porosity n (%) | Compressive strength $\sigma_{cf}$ (MPa) |
|---|---|---|---|---|
| 1 | 2.32 | 0.16 | 14.4 | 58 |
| 3 | 2.30 | 0.16 | 15.1 | 63 |
| 4 | 2.31 | 0.11 | 14.8 | 61 |

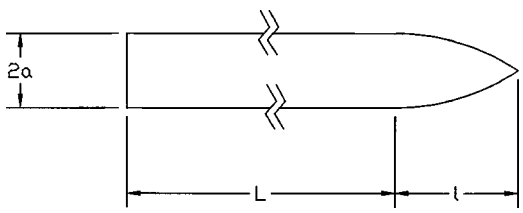



**Fig. 1 Projectile geometry**

several triaxial compression tests ([21,22]) on samples from the three target batches and found the brittle to ductile transition confining pressure to be about 50 MPa.

**The 12.7-mm-Diameter, 0.117 kg, 3.0 CRH Projectiles.** Our first set of experiments was conducted with steel projectiles machined from 4340 $R_c$ 45 ([23]) round stock. Figure 1 shows the projectile geometry, and for this set of experiments $2a=12.7$ mm, $L=106$ mm, and $l=21$ mm. The target impact surface was 0.51 m square, and the target lengths and other data are given in Table 2. The sides and bottom of the targets were surrounded by 0.10-m-thick concrete placed between a steel form and the limestone. Six unconfined compression tests were conducted with 51-mm-diameter, 108-mm-long samples cored from a representative limestone block from Batch 1 and the average strength was $\sigma_{cf}=58$ MPa.

A 32-mm-diameter powder gun launched the 0.117 kg projectiles to the striking velocities recorded in Table 2. An additional experiment was conducted at $V_s=1605$ m/s, but the trajectory was

**Table 2 Penetration data for the 12.7-mm-diameter 4340 $R_c$ 44–45, 0.117 kg, 3.0 CRH projectiles. For ptich and yaw: *D*=down, *U*=up, *R*=right, *L*=left.**

| Shot, batch numbers | Target length (m) | Striking velocity $V_s$ (m/s) | Pitch, yaw (deg.) | Penetration depth P(m) | Projectile mass loss (%) | R (MPa) |
|---|---|---|---|---|---|---|
| 1-0370, 1 | 0.61 | 459 | 1.9D, 3.0L | 0.141 | 0.16 | 733 |
| 1-0371, 1 | 0.61 | 608 | 1.0U, 1.0R | 0.232 | 0.24 | 734 |
| 1-0372, 1 | 0.91 | 853 | 0, 0 | 0.362 | 0.96 | 875 |
| 1-0377, 1 | 0.91 | 956 | 0, 0.3L | 0.523 | 1.29 | 719 |
| 1-0373, 1 | 0.91 | 1134 | 1.0U, 1.9R | 0.562 | 2.59 | 930 |
| 1-0376, 1 | 1.22 | 1269 | 0, 0 | 0.812 | 3.87 | 745 |
| 1-0375, 1 | 1.22 | 1404 | 0, 0.7R | 0.924 | 5.25 | 775 |
| 1-0374, 1 | 1.22 | 1502 | 0.7U, 0 | 1.017 | 5.90 | 783 |

**Table 3 Penetration data for the 7.1-mm-diameter 4340 $R_c$ 44–46 or Aer Met 100 $R_c$ 53 (shots 4-1847 and 4-1846), 0.0205 kg, 3.0 CRH projectiles. For pitch and yaw: *D*=down, *U*=up, *R*=right, *L*=left.**

| Shot, batch numbers | Target length (m) | Striking velocity $V_s$ (m/s) | Pitch, yaw (deg.) | Penetration depth P(m) | Projectile mass loss (%) | R (MPa) | Compressive strength $\sigma_{cf}$ (MPa) |
|---|---|---|---|---|---|---|---|
| 1-0422, 3 | 0.30 | 497 | 0, 0.75R | 0.067 | 0.34 | 1033 | |
| 1-0418, 3 | 0.61 | 597 | 0.5U, 0 | 0.105 | 0.34 | 895 | 64.7 |
| 1-0419, 3 | 0.61 | 787 | 0, 1.25R | 0.165 | 0.78 | 937 | |
| 1-0420, 3 | 0.61 | 1037 | 0.75U, 0.5R | 0.271 | 2.11 | 927 | |
| 4-1835, 3 | 1.02 | 1365 | 2.0D, 1.6R | 0.430 | 4.90 | 926 | 58.2 |
| 4-1836, 3 | 1.02 | 1516 | 0.9U, 1.1R | 0.516 | 7.34 | 903 | |
| 1-0439, 1 | 0.46 | 795 | 0.5U, 0 | 0.178 | 0.83 | 877 | 60.1 |
| 1-0440, 1 | 0.61 | 1060 | 0.25U, 0.5R | 0.294 | 2.49 | 877 | 60.3 |
| 1-0441, 1 | 0.61 | 1230 | 0, 0.25R | 0.392 | 3.76 | 837 | 55.7 |
| 4-1845, 1 | 1.02 | 1340 | 1.75D, 1.0L | 0.437 | 4.84 | 870 | 57.5 |
| 4-1847, 3 | 1.02 | 1266 | 1.0D, 1.5R | 0.379 | 4.59 | 925 | 63.8 |
| 4-1846, 3 | 1.02 | 1674 | 0.5U, 0 | 0.581 | 10.85 | 944 | 63.9 |

**Table 4 Penetration data for the 25.4-mm-diameter 4340 $R_c$ 45–46, 0.931 kg, 3.0 CRH projectiles. For pitch and yaw: *D*=down, *U*=up, *R*=right, *L*=left.**

| Shot, batch numbers | Target length (m) | Striking velocity $V_s$ (m/s) | Pitch, yaw (deg.) | Penetration depth P(m) | Projectile mass loss (%) | R (MPa) | Compressive strength $\sigma_{cf}$ (MPa) |
|---|---|---|---|---|---|---|---|
| LROD99-1, 4 | 0.61 | 407 | 0.4D, 0.5L | 0.26 | 0 | 617 | 56.2 |
| LROD99-2, 4 | 0.76 | 566 | 0.6U, 0 | 0.39 | 0.1 | 769 | 61.9 |
| LROD99-3, 4 | 1.07 | 800 | 0.6D, 0.1L | 0.79 | 0.3 | 693 | 66.7 |
| LROD99-6, 4 | 1.52 | 917 | —, 0.3R | 1.02 | 1.1 | 681 | 59.8 |
| LROD99-7, 4 | 1.98 | 1177 | 0, 0.3R | 1.50 | 2.8 | 705 | 61.2 |

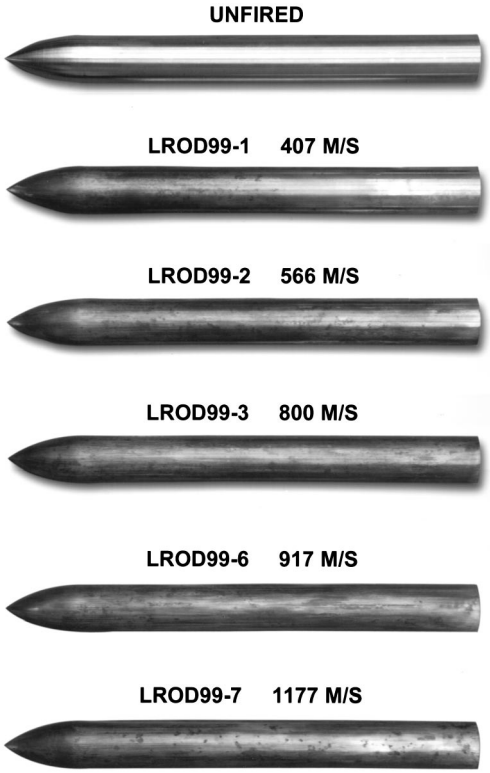



**Fig. 2 Post-test photographs of the 25.4-mm-diameter projectiles**

curved and the projectile exited the side of the target at a depth of about 0.65 m. The projectiles were fitted with sabots and obturators that separated from the projectiles before impact. Four laser diode systems measured striking velocities and orthogonal radiographs measured pitch and yaw angles. The target resistance $R$ was calculated from (2) for each experiment and recorded in Table 2. The average target resistance parameter for this set of experiments is $R = 787$ MPa.

**The 7.1-mm-Diameter 0.020 kg, 3.0 CRH Projectiles.** Our second set of experiments was conducted with steel projectiles machined from both 4340 $R_c$ 45 and Aer Met 100 $R_c$ 53 ([24]) round stock. Figure 1 shows the projectile geometry, and for this set of experiments $2a = 7.11$ mm, $L = 59.3$ mm, and $l = 11.8$ mm. A 20-mm-powder gun launched the 0.020 kg projectiles to striking velocities of 1230 m/s. For the larger striking velocities recorded in Table 3, a two-stage 50/20 mm light-gas gun launched the projectiles. The same target geometries and ballistics measurements as those described for the 12.7-mm-diameter 0.117 kg projectiles were used for this set of experiments.

For this second set of experiments, we performed unconfined compression tests on two samples cored from the targets. The compressive strengths listed in Table 3 are the average of two unconfined compression tests conducted with 50-mm-diameter 108-mm-long cores. In addition, we conducted penetration experiments from both Batch 1 and Batch 3 limestone targets to compare results from the two batches. We show later negligible differences in the penetration data from both batches. The average target resistance parameter of this set of experiments is $R = 913$ MPa.

We also conducted an experiment with a 4340 $R_c$ 45 projectile at $V_s = 1649$ m/s. That projectile severely bent and turned within the target. Table 3 shows two experiments conducted with Aer Met 100 $R_c$ 53 projectiles. Shot 4-1846 with a striking velocity of 1674 m/s had a nearly straight trajectory. We then conducted experiments at $V_s = 1749$, 1826, and 1863 m/s with Aer Met 100 $R_c$ 53 projectiles and these projectiles severely bent and turned within the targets. Piekutowski, Forrestal, Poormon, and Warren [7] discusses in detail the better performance of the Aer Met 100 $R_c$ 53 projectiles.

**The 25.4-mm-diameter 0.931 kg, 3.0 CRH Projectiles.** Our third set of experiments was conducted with steel projectiles ma-

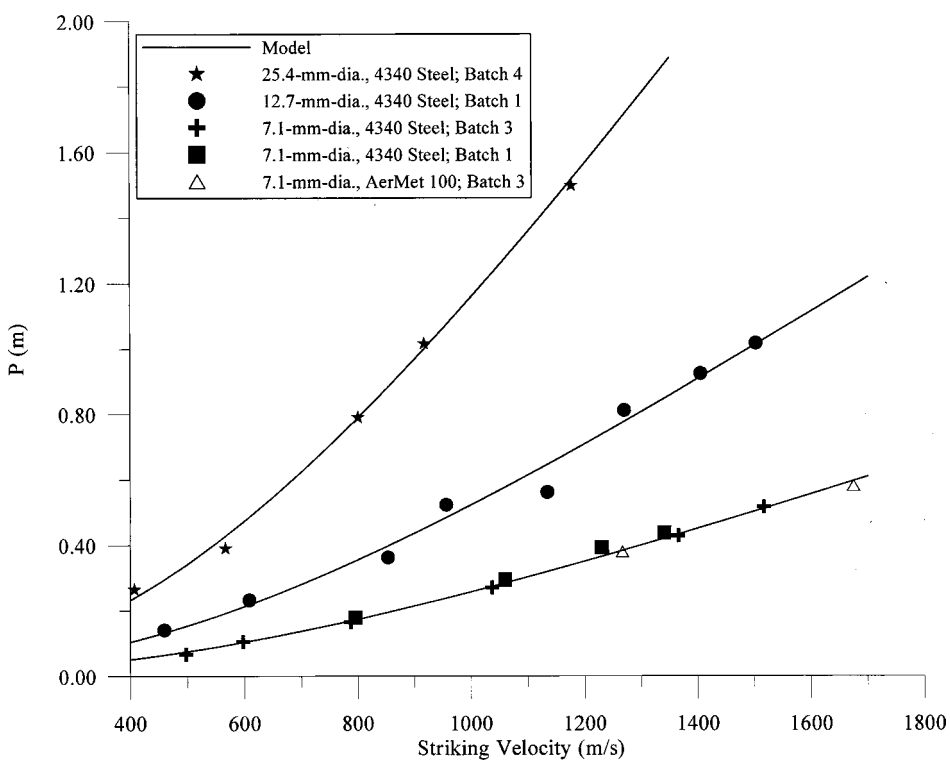



**Fig. 3 Data and model predictions for the limestone targets**

chined from 4340 $R_c$ 45 round stock. Figure 1 shows the projectile geometry, and for this set of experiments $2a=25.4$ mm, $L=212$ mm, and $l=42$ mm. The target impact surface was 1.02-m square and the target lengths are given in Table 4. The sides and bottom of the targets were surrounded by 0.10-m-thick concrete placed between a steel form and the limestone. An 83-mm powder gun launched the 0.931 kg projectiles to the striking velocities recorded in Table 4. The same experimental methods used for the other data sets were also used to obtain the results given in Table 4. Data in Table 4 were limited to $V_s=1177$ m/s because of the size and mass of the targets. The average value of $R$ for this set of experiments was $R=693$ MPa.

## Results and Discussion

As previously discussed, post-test target observations for the experiments summarized in Tables 2, 3, and 4 showed a conical entry crater with a depth of two or more projectile diameters followed by a circular penetration tunnel with nearly the projectile diameter. However, we made post-test observations only after the targets were split with the techniques used by stone masons. Figure 2 shows post-test photographs of the 25.4-mm-diameter projectiles, and mass losses are given in Table 4. These recovered projectiles lost mass caused by abrasion, but the nose shapes before and after penetration have similar shapes. Post-test projectiles from the other data sets (Tables 2 and 3) had shapes similar to those shown in Fig. 2.

Figure 3 shows depth of penetration $P$ versus striking velocity $V_s$ for the data in Tables 2, 3, and 4. For the model, we use the average value of $R$ calculated from (2) for the experiments in each data set. For the 7.1, 12.7, and 25.4-mm-diameter projectiles, the average target resistance parameter is $R=913$, 787, and 693 MPa, respectively. For projectile diameters not tested in this study, we recommend (3) be used to calculate $R$. Equation (3) with $K=607$ MPa, $k=86$ MPa, and $2a_o=25.4$ mm accurately recover the measured values of $R$ from each of the three data sets. We hypothesize that the penetration equations ((1) and (3)) for this limestone target will be reasonably accurate for larger scale projectiles, but data from much more expensive field tests with larger diameter projectiles must be obtained to confirm our hypothesis.

## Summary

We conducted sets of penetration experiments into limestone targets with three scales of geometrically similar projectiles. The ogive-nose rod projectiles with a length-to-diameter ratio of ten were machined from 4340 $R_c$ 45 and Aer Met 100 $R_c$ 53 steel, round stock and had diameters and masses of 7.1 mm, 0.020 kg; 12.7 mm, 0.117 kg; and 25.4 mm, 0.931 kg. For the 4340 $R_c$ 45 and Aer Met 100 $R_c$ 53 steel projectiles, penetration depth increased as striking velocity increased to 1500 and 1700 m/s, respectively. For larger striking velocities, the projectiles bent during penetration without nose erosion, deviated from the shot line, and exited the sides of the target or turned severely in the target.

Based on data sets with these three projectile scales, we present an analytical penetration equation based on a target strength parameter that is determined from penetration depth versus striking velocity data. We show that the target resistance parameter depends on the projectile shank diameter and present an equation that describes this diameter dependence. We hypothesize that the penetration equations for this limestone target will give accurate predictions for larger scale projectile, but data from much more expensive field tests with larger diameter projectiles must be obtained to confirm our hypothesis.


## Acknowledgments

This work was supported by the United States Department of Energy under Contract DE-AC04-94AL85000, the Joint DoD/DOE Munitions Technology Development Program, and the U.S. Army Corps of Engineers Hardened Structures Research Programs (AT22 and AT40).


## References


[1] Backman, M. E., and Goldsmith, W., 1978, ''The Mechanics of Penetration of Projectiles Into Targets,'' Int. J. Eng. Sci., **16**, pp. 1–99.

[2] Anderson, C. E., and Bodner, S. R., 1988, ''Ballistic Impact: The Status of Analytical and Numerical Modeling,'' Int. J. Impact Eng., **7**, pp. 9–35.

[3] Hohler, V., and Stilp, A. J., 1990, ''Long-Rod Penetration Mechanics,'' *High Velocity Impact Dynamics*, J. A. Zukas, ed., John Wiley and Sons, New York, Chapter 5.

[4] Recht, R., 1990, ''High Velocity Impact Dynamics: Analytical Modeling of Plate Penetration Dynamics,'' *High Velocity Impact Dynamics*, J. A. Zukas, ed., John Wiley and Sons, New York, Chapter 7.

[5] Corbett, G. G., Reid, S. R., and Johnson, W., 1996, ''Impact Loading of Plates and Shells by Free-Flying Projectiles: A Review,'' Int. J. Impact Eng., **18**, pp. 141–230.

[6] Warren, T. L., and Forrestal, M. J., 1998, ''Effects of Strain Hardening and Strain-Rate Sensitivity on the Penetration of Aluminum Targets With Spherical-Nosed Rods,'' Int. J. Solids Struct., **35**, Nos. 28–29, pp. 3737–3753.

[7] Piekutowski, A. J., Forrestal, M. J., Poormon, K. L., and Warren, T. L., 1999, ''Penetration of 6061-T6511 Aluminum Targets by Ogive-Nose Steel Projectiles With Striking Velocities Between 0.5 and 3.0 km/s,'' Int. J. Impact Eng., **23**, pp. 723–734.

[8] Forrestal, M. J., Frew, D. J., Hanchak, S. J., and Brar, N. S., 1996, ''Penetration of Grout and Concrete Targets With Ogive-Nose Steel Projectiles,'' Int. J. Impact Eng., **18**, No. 5, pp. 465–476.

[9] Frew, D. J., Hanchak, S. J., Green, M. L., and Forrestal, M. J., 1998, ''Penetration of Concrete Targets With Ogive-Nose Steel Rods,'' Int. J. Impact Eng., **21**, No. 6, pp. 489–497.

[10] Forrestal, M. J., Altman, B. S., Cargile, J. D., and Hanchak, S. J., 1994, ''An Empirical Equation for Penetration Depth of Ogive-Nose Projectiles Into Concrete Targets,'' Int. J. Impact Eng., **15**, No. 4, pp. 395–405.

[11] Fossum, A. F., Senseny, P. E., Pfeifle, T. W., and Mellegard, K. D., 1995, ''Experimental Determination of Probability Distributions for Parameters of a Salem Limestone Cap Plasticity Model,'' Mech. Mater., **21**, pp. 119–137.

[12] Ortiz, M., 1996, ''Computational Micromechanics,'' Comput. Mech., **18**, pp. 321–338.

[13] Camacho, G. T., and Ortiz, M., 1997, ''Adaptive Lagrangian Modelling of Ballistic Penetration of Metallic Targets,'' Comput. Meth., Appl. Eng., **142**, pp. 269–301.

[14] Green, S. J., and Perkins, R. D., 1972, ''Uniaxial Compression Tests at Varying Strain Rates on Three Geologic Materials,'' *Proc. 10th Symp. Rock Mechanics, Inst. Mining Metallurgical and Petroleum Engrs.*, K. E. Gray, ed., AIME, New York, pp. 35–54.

[15] Lindholm, U. S., Yeakley, L. M., and Nagy, A., 1974, ''The Dynamic Strength and Fracture Properties of Dresser Basalt,'' Int. J. Rock Mech. Min. Sci. Geomech. Abstr., **II**, pp. 181–191.

[16] Lipkin, J., Grady, D. E., and Campbell, J. D., 1977, ''Dynamic Flow and Fracture of Rock in Pure Shear,'' *Proc. 18th U.S. Symp. Rock Mechanics*, F. Wang and G. B. Clark, eds., Colorado School of Mines Press, Golden, Co, pp. 3 B2–1, 3 B2–7.

[17] Larson, D. B., and Anderson, G. D., 1979, ''Plane Shock Wave Studies of Porous Geological Media,'' J. Geophys. Res., **84**, No. B9, pp. 4592–4600.

[18] Grady, D. E., and Hollenbach, R. E., 1979, ''Dynamic Fracture Strength of Rock,'' Geophys. Res. Lett., **6**, No. 2, pp. 73–76.

[19] Ahrens, T. J., and Rubin, A. M., 1993, ''Impact-Induced Tensional Failure in Rock,'' J. Geophys. Res., **98**, No. E1, pp. 1185–1203.

[20] Aidun, J. B., and Gupta, Y. M., 1995, ''Shear and Compression Waves in Shocked Calcium Carbonate,'' J. Geophys. Res., **100**, No. B2, pp. 1955–1980.

[21] Farmer, I., 1983, *Engineering Behavior of Rocks*, 2nd Ed., Chapman and Hall, New York.

[22] Jaeger, J. C., and Cook, N. G. W., 1979, *Fundamentals of Rock Mechanics*, 3rd Ed., Chapman and Hall, New York.

[23] Brown, W. F., Mindlin, H., and Ho, C. Y., 1996, *Aerospace Structural Metals Handbook*, Vol. 1, Code 1206, CINDAS/USAF CRDA Handbooks Operation, Purdue University, West Lafayette, IN.

[24] Dahl, J., 1991, ''Aer Met 100—An Advanced Steel for the Aerospace Industry,'' Adv. Mat. Tech. Int., pp. 40–45.

# Eccentric Crack in a Piezoelectric Strip Under Electromechanical Loading

**Jeong Woo Shin**
Researcher, Korea Aerospace Research Institute,
52 Oun-dong, Yusong-gu, Taejon 306-333, Korea

**Soon Man Kwon**
Graduate Student

**Kang Yong Lee**[1]
Professor

Department of Mechanical Engineering, Graduate School,
Yonsei University, Seoul 120-749, Korea

## 1 Introduction

With an increasingly wide application media in engineering, the study on the crack problem in piezoelectric media has received much interest. In this paper, we apply the theory of linear piezoelectricity to the electroelastic problem of a finite eccentric crack off the center line in a piezoelectric ceramic strip under antiplane shear loading. The continuous crack boundary condition is adopted. By using integral transform techniques, the problem is reduced to the solution of a Fredholm integral equation of the second kind, which is obtained from two pairs of dual integral equations. Numerical results for the energy release rate are shown graphically.

## 2 Problem Statement and Method of Solution

Consider a piezoelectric medium in the form of an infinitely long strip containing a finite eccentric crack off the center line parallel to the edges subjected to the combined antiplane mechanical and in-plane electric loads. A set of Cartesian coordinates $(x,y,z)$ is attached to the center of the crack. The total thickness of the strip is $2h(=h_1+h_2)$, where $h_1(=h-e)$ is the distance between the top surface of the strip and the crack plane and $h_2(=h+e)$ is that between the bottom surface of the strip and the crack plane. $e$ denotes the distance from the center line of the strip to the crack plane. The crack is situated along the virtual interface $(-a\leqslant x\leqslant a,\ y=0)$. Because of the assumed symmetry in geometry and loading, it is sufficient to consider the problem for $0\leqslant x<\infty$ only. Antiplane governing equations are simplified to

$$c_{44}\nabla^2 w^{(i)}+e_{15}\nabla^2\phi^{(i)}=0, \tag{1}$$

$$e_{15}\nabla^2 w^{(i)}-d_{11}\nabla^2\phi^{(i)}=0, \tag{2}$$

where $\nabla^2=\partial^2/\partial x^2+\partial^2/\partial y^2$ is the two-dimensional Laplace operator, $w^{(i)}$, $\phi^{(i)}$ $(i=1,2)$, $c_{44}$, $d_{11}$, and $e_{15}$ are the out-of-plane displacement, the electric potential, the elastic modulus measured in a constant electric field, the dielectric permittivity measured at a constant strain, and the piezoelectric constant, respectively. Superscript $i$ $(i=1,2)$ stands for upper and lower regions, respectively. The boundary conditions are written as follows:

$$\sigma_{yz}^{(i)}(x,0)=0,\quad (0\leqslant x<a), \tag{3}$$

$$w^{(1)}(x,0^+)=w^{(2)}(x,0^-),\quad (a\leqslant x<\infty), \tag{4}$$

$$D_y^{(1)}(x,0^+)=D_y^{(2)}(x,0^-),\quad (0\leqslant x<a), \tag{5}$$

$$E_x^{(1)}(x,0^+)=E_x^{(2)}(x,0^-),\quad (0\leqslant x<a), \tag{6}$$

$$\phi^{(1)}(x,0^+)=\phi^{(2)}(x,0^-),\quad (a\leqslant x<\infty), \tag{7}$$

$$\begin{aligned}&\sigma_{yz}^{(1)}(x,0^+)=\sigma_{yz}^{(2)}(x,0^-)\quad (a\leqslant x<\infty),\\ &D_y^{(1)}(x,0^+)=D_y^{(2)}(x,0^-),\quad (a\leqslant x<\infty).\end{aligned} \tag{8}$$

$$\text{Case 1:}\quad \begin{aligned}&\sigma_{yz}^{(1)}(x,h_1)=\sigma_{yz}^{(2)}(x,-h_2)=\tau_0,\\ &D_y^{(1)}(x,h_1)=D_y^{(2)}(x,-h_2)=D_0,\end{aligned} \tag{9}$$

$$\text{Case 2:}\quad \begin{aligned}&\gamma_{yz}^{(1)}(x,h_1)=\gamma_{yz}^{(2)}(x,-h_2)=\gamma_0,\\ &E_y^{(1)}(x,h_1)=E_y^{(2)}(x,-h_2)=E_0,\end{aligned} \tag{10}$$

$$\text{Case 3:}\quad \begin{aligned}&\sigma_{yz}^{(1)}(x,h_1)=\sigma_{yz}^{(2)}(x,-h_2)=\tau_0,\\ &E_y^{(1)}(x,h_1)=E_y^{(2)}(x,-h_2)=E_0,\end{aligned} \tag{11}$$

$$\text{Case 4:}\quad \begin{aligned}&\gamma_{yz}^{(1)}(x,h_1)=\gamma_{yz}^{(2)}(x,-h_2)=\gamma_0,\\ &D_y^{(1)}(x,h_1)=D_y^{(2)}(x,-h_2)=D_0,\end{aligned} \tag{12}$$

where $\sigma_{yz}^{(i)}$, $\gamma_{yz}^{(i)}$, $D_y^{(i)}$, and $E_j^{(i)}$ $(j=x,y)$ are the shear stress, the shear strain, the electric displacements, and the electric fields, respectively, and $\tau_0$, $D_0$, $\gamma_0$, and $E_0$ are a uniform shear stress, displacement, shear strain, and electric field, respectively. A Fourier transform is applied to Eqs. (1) and (2), and the results are

$$w^{(i)}(x,y)=\frac{2}{\pi}\int_0^\infty\{A_1^{(i)}(s)\cosh(sy)+A_2^{(i)}(s)\sinh(sy)\}\times\cos(sx)ds+a_0y, \tag{13}$$

$$\phi^{(i)}(x,y)=\frac{2}{\pi}\int_0^\infty\{B_1^{(i)}(s)\cosh(sy)+B_2^{(i)}(s)\sinh(sy)\}\times\cos(sx)ds-b_0y, \tag{14}$$

where $A_j^{(i)}$, $B_j^{(i)}$ $(j=1,2)$ are the unknowns to be solved. $a_0,b_0$ are real constants, which will be determined from the edge loading conditions. Applying Eqs. (8)–(12), and considering Eqs. (3)–(7), we obtain the following two simultaneous dual integral equations:

$$\int_0^\infty sF(s)\left[M_A(s)+\frac{e_{15}}{c_{44}}M_B(s)\right]\cos(sx)ds=\frac{\pi}{2}\frac{c_0}{c_{44}},\quad (0\leqslant x<a), \tag{15}$$

$$\begin{aligned}&\int_0^\infty M_A(s)\cos(sx)ds=0,\qquad (a\leqslant x<\infty),\\ &\int_0^\infty sM_B(s)\sin(sx)ds=0,\qquad (0\leqslant x<a),\\ &\int_0^\infty M_B(s)\cos(sx)ds=0,\qquad (a\leqslant x<\infty),\end{aligned} \tag{16}$$

where

$$2M_A(s)=A_1^{(1)}(s)-A_1^{(2)}(s),\quad 2M_B(s)=B_1^{(1)}(s)-B_1^{(2)}(s), \tag{17}$$

$$F(s)=\tanh(sh)-\frac{2\sinh^2(se)}{\sinh(2sh)}, \tag{18}$$

[1]To whom correspondence should be addressed.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Jan. 19, 2000; final revision, Apr. 20, 2000. Associate Technical Editor: D. Kouris.

 **Copyright © 2000 by ASME** 

$$c_0 = c_{44}a_0 - e_{15}b_0, \tag{19}$$

$$= \tau_0, \quad \text{Cases 1 and 3}, \tag{20a}$$

$$= c_{44}\gamma_0 - e_{15}E_0, \quad \text{Case 2}, \tag{20b}$$

$$= \frac{(c_{44}d_{11} + e_{15}^2)\gamma_0 - e_{15}D_0}{d_{11}}, \quad \text{Case 4}. \tag{20c}$$

To solve the two sets of dual integral equations, we define $M_A(s)$ and $M_B(s)$ as follows:

$$M_A(s) \equiv \frac{\pi}{2}\frac{c_0 a^2}{c_{44}}\int_0^1 \sqrt{\xi}\Psi(\xi)J_0(sa\xi)d\xi,$$
$$M_B(s) \equiv \int_0^1 \sqrt{\xi}\Phi(\xi)J_0(sa\xi)d\xi, \tag{21}$$

where $J_0(sa\xi)$ is the zero-order Bessel function of the first kind. Inserting Eq. (21) into Eqs. (15) and (16), we can find that $\Phi(\xi)=0$ and $\Psi(\xi)$ is given by a Fredholm integral equation of the second kind in the form

$$\Psi(\xi) + \int_0^1 K(\xi,\eta)\Psi(\eta)d\eta = \sqrt{\xi}, \tag{22}$$

where

$$K(\xi,\eta) = \sqrt{\xi\eta}\int_0^\infty s\{F(s/a)-1\}J_0(s\eta)J_0(s\xi)ds. \tag{23}$$

Extending the traditional concept of stress intensity factor to other field variables ([1]), we have

$$K^T = c_0\sqrt{\pi a}\Psi(1), \quad K^S = K^T/c_{44}, \quad K^D = e_{15}K^T/c_{44}, \quad K^E = 0, \tag{24}$$

where $K^T$, $K^S$, $K^D$, and $K^E$ are stress intensity, strain intensity, electric displacement intensity, and electric field intensity factor, respectively. Evaluating the energy release rate $G$ obtained by Pak [1], we obtain

$$G = \frac{K^T K^S - K^D K^E}{2} = \frac{\pi a}{2c_{44}}c_0^2\Psi^2(1). \tag{25}$$

## 3 Numerical Results and Discussion

From Eq. (24), it is noted that the uniform electric load has no influence on the field singularities, and the electric displacement intensity factor depends on the material constants and the applied mechanical load, $\tau_0$ but not on the applied electric loads, $D_0$ and $E_0$. The results coincide with the ones of Kwon and Lee [2] and Gao and Fan [3]. Figure 1 displays the normalized energy release rate $G/G_\infty$ versus the $a/h$ with various $e/h$ values. $G/G_\infty$ increases with increase of the $a/h$ and $e/h$ ratios. The normalized energy release rate $G/G_{cr}$ of a PZT-5H ceramic is shown in Fig. 2 with the variation of the applied electric field $E_0$ and the $a/h$ ratio for a crack length of $2a = 0.02$ m and $\gamma_0 = 9.5\times10^{-5}$ (Case 2) in cases of $e/h = 0.0$ and $e/h = 0.5$. As the magnitude of electric field increases from zero, $G$ increases or decreases according to the direction of the load. But once the minimum value of $G$ is reached, further increase of the electrical load increase $G$ continuously. $G$ increase with increase of the $e/h$ ratio. It can be shown that similar results are obtained in Case 4.

## 4 Conclusions

The normalized energy release rate increases when the $a/h$ and $e/h$ ratios increase. The energy release rate is dependent on the electric loading only under constant strain loading and independent of it under constant stress loading. In constant strain loadings, the minimum normalized energy release rate can exist with the variation of electrical load but has always positive values.

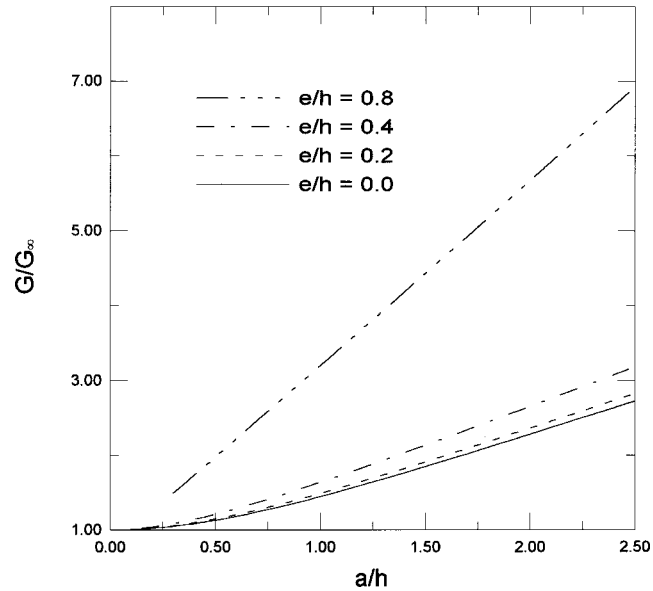



**Fig. 1 Energy release rate $G/G_\infty$ versus $a/h$ with various $e/h$ values**

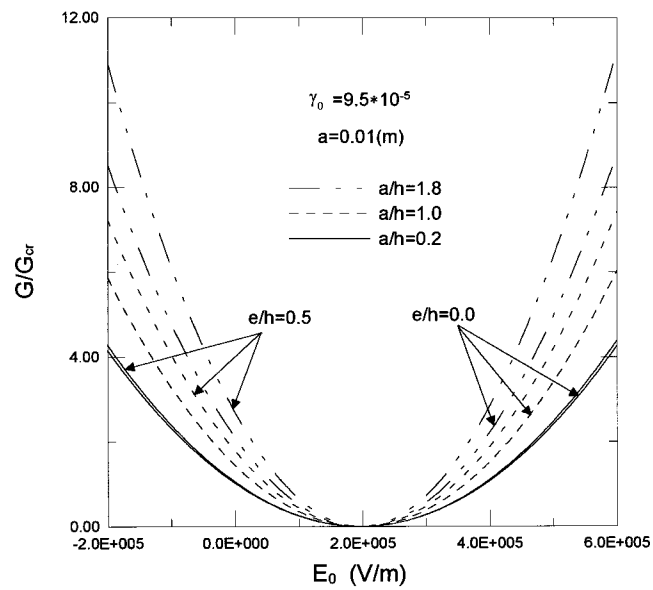



**Fig. 2 Energy release rate of PZT-5H (Case 2) with various $e/h$ values**


## Acknowledgments

The authors are grateful for the support provided by the Brain Korea 21 Project and a grant from the Korea Science & Engineering Foundation (KOSEF) and Safety and Structural Integrity Research Center at the Sungkyunkwan University.


## References


[1] Pak, Y. E., 1990, ''Crack Extension Force in a Piezoelectric Materials,'' ASME J. Appl. Mech., **57**, pp. 647–653.

[2] Kwon, S. M., and Lee, K. Y., 2000, ''Analysis of Stress and Electric Fields in a Rectangular Piezoelectric Body With a Center Crack Under Anti-plane Shear Loading,'' Int. J. Solids Struct., **37**, pp. 4859–4869.

[3] Gao, C. F., and Fan, W. X., 1999, ''A General Solution for the Plane Problem in Piezoelectric Media With Collinear Crack,'' Int. J. Eng. Sci., **37**, pp. 347–363.