

# Transactions of the ASME®

Editor, LEWIS T. WHEELER  
APPLIED MECHANICS DIVISION

Executive Committee  
(Chair) D. KRAJCIKOVIC  
S. KYRIAKIDES  
P. D. SPANOS  
M. C. BOYCE  
W.-K. LIU

Associate Editors  
E. ARRUDA (2004)  
J. R. BARBER (2003)  
R. C. BENSON (2003)  
A. A. FERRI (2003)  
H. GAO (2003)  
V. K. KINRA (2002)  
D. A. KOURIS (2002)  
A. K. MAL (2004)  
B. M. MORAN (2002)  
A. NEEDLEMAN (2004)  
O. O'REILLY (2004)  
N. C. PERKINS (2002)  
M.-J. PINDER (2003)  
K. R. RAJAGOPAL (2003)  
K. T. RAMESH (2003)  
K. RAVI-CHANDAR (2003)  
W. S. SARIC (2003)  
D. A. SIGINER (2003)  
T. E. TEZDUYAR (2003)  
N. TRIANTAFYLIDIS (2003)

## BOARD ON COMMUNICATIONS

Chair and Vice-President  
OZDEN OCHOA

## OFFICERS OF THE ASME

President, W. A. WEIBLEN  
Executive Director, D. L. BELDEN  
Treasurer, R. E. NICKELL

## PUBLISHING STAFF

Managing Director, Engineering  
THOMAS G. LOUGHLIN  
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 bimonthly  
(Jan., Mar., May, July, Sept., Nov.)

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 © 2001 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.

# Journal of Applied Mechanics

Published Bimonthly by The American Society of Mechanical Engineers

VOLUME 68 • NUMBER 5 • SEPTEMBER 2001

## TECHNICAL PAPERS

- 697 Linear Thermoelastic Higher-Order Theory for Periodic Multiphase Materials  
J. Aboudi, M.-J. Pindera, and S. M. Arnold
- 708 Stress Analysis of Layered Elastic Solids With Cracks Using the Fast Fourier Transform and Conjugate Gradient Techniques  
I. A. Polonsky and L. M. Keer
- 715 On Plastic Collapse of Media With Random Yield Strength  
A. P.-D. Ku and R. P. Nordgren
- 725 Mode II Edge Delamination of Compressed Thin Films  
D. S. Balint and J. W. Hutchinson
- 731 A Complex Potential-Variational Method for Stress Analysis of Unsymmetric Laminates With an Elliptical Cutout  
E. Madenci, A. Barut, and M. P. Nemeth
- 740 Asymptotic Solutions of Penny-Shaped Inhomogeneities in Global Eshelby's Tensor  
Q. Yang, W. Y. Zhou, and G. Swoboda
- 751 On the Simulation of Discontinuous Functions  
J. D. Turner
- 758 Relationship Between the External Force and the Specific Work of Fracture  $R$  for Steady-State Tearing and Peeling of Elastoplastic Materials  
J. H. Liu, A. G. Atkins, and G. Jeronimidis
- 766 On the Steady Motions of a Rotating Elastic Rod  
N. M. Kinkaid, O. M. O'Reilly, and J. S. Turcotte
- 772 Size Effects in a Slowly Sheared Granular Media  
S. J. Antony and M. Ghadiri
- 776 A New Method for Nonlinear Two-Point Boundary Value Problems in Solid Mechanics  
L. S. Ramachandra and D. Roy
- 787 A Nonlinear Generalized Maxwell Fluid Model for Viscoelastic Materials  
D. T. Corr, M. J. Starr, R. Vanderby Jr., and T. M. Best
- 791 Structural Modification for the Assignment of Zeros Using Measured Receptances  
J. E. Mottershead
- 799 On Some Issues in Shakedown Analysis  
G. Maier

## BRIEF NOTES

- 809 On the Existence of a Solution for a Solid Circular Plate Bilaterally Supported Along Two Antipodal Boundary Arcs and Loaded by a Central Transverse Concentrated Force  
G. Monegato and A. Strozzi
- 812 A Note on Wavelet-Based Method for Damage Detection  
L. Zhang, S. T. Quek, and Q. Wang
- 814 A Novel Technique in the Solution of Axisymmetric Large Deflection Analysis of a Circular Plate  
L. S. Ramachandra and D. Roy

(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.

816 On Characteristic Times in Generalized Thermoelasticity  
D. V. Strunin

817 Closed-Form Solutions for the Stress Fields Induced by Blunt Wedge-Shaped Indenters in Elastic Half-Planes  
C. E. Truman and A. Sackfield

## **ANNOUNCEMENTS AND SPECIAL NOTES**

820 Information for Authors

821 Preparing and Submitting a Manuscript for Journal Production and Publication

822 Preparation of Graphics for ASME Journal Production and Publication

823 New Reference Format

824 34th Solid Mechanics Conference—Announcement

# Linear Thermoelastic Higher-Order Theory for Periodic Multiphase Materials

**J. Aboudi**

Tel-Aviv University,  
Ramat-Aviv 69978, Israel

**M.-J. Pindera**

Civil Engineering Department  
University of Virginia,  
Charlottesville, VA 22903  
Mem. ASME

**S. M. Arnold**

NASA Glenn Research Center,  
Cleveland, OH 44135

*A new micromechanics model is presented which is capable of accurately estimating both the effective elastic constants of a periodic multiphase composite and the local stress and strain fields in the individual phases. The model is presently limited to materials characterized by constituent phases that are continuous in one direction, but arbitrarily distributed within the repeating unit cell which characterizes the material's periodic microstructure. The model's analytical framework is based on the homogenization technique for periodic media, but the method of solution for the local displacement and stress fields borrows concepts previously employed by the authors in constructing the higher-order theory for functionally graded materials, in contrast with the standard finite element solution method typically used in conjunction with the homogenization technique. The present approach produces a closed-form macroscopic constitutive equation for a periodic multiphase material valid for both uniaxial and multiaxial loading which, in turn, can be incorporated into a structural analysis computer code. The model's predictive accuracy is demonstrated by comparison with reported results of detailed finite element analyses of periodic composites as well as with the classical elasticity solution for an inclusion in an infinite matrix. [DOI: 10.1115/1.1381005]*

## 1 Introduction

Micromechanical modeling of multiphase materials continues to be an important area in both the mechanics and materials communities due to the need for the development of new materials for use in wide range of modern applications. These include aerospace, aircraft, biomedical, electronic, and recreational industry applications. Many advanced applications require the use of engineered material microstructures which necessarily involve the use of multiphase material concepts. Micromechanical modeling techniques offer an efficient approach to developing an understanding of how different microstructural details affect the average and local responses of multiphase materials. Thus they can be employed by both the mechanics and materials communities as design and diagnostic tools in developing and analyzing different microstructural concepts for a given application. The fact that new approaches are continuously being proposed attests to the importance of this area of research.

A large body of literature exists which deals with the micromechanical modeling techniques for heterogeneous materials developed during the past several decades, and thus a comprehensive review of these techniques is beyond the scope of this article. These include use of simple Voigt and Reuss hypotheses, self-consistent schemes and their generalizations, differential schemes, concentric cylinder models, bounding techniques, and approximate or numerical analyses of periodic arrays of inclusions or fibers in the surrounding matrix phase. The various approaches may be divided into three broad categories: those based on the direct calculation of Hill's stress or strain concentration matrices, Hill [1], and those based on the concepts of a representative volume element (RVE) and a repeating unit cell (RUC). A major difference between the latter two categories lies in the boundary conditions applied to a material subvolume considered to repre-

sent the average response of the entire composite. RVE-based models employ homogeneous displacement or traction boundary conditions, whereas symmetry or periodic boundary conditions are employed in RUC-based models. Typically, RVE-based approaches employ simplified geometric representations of the entire composite for which analytical solutions under homogeneous displacement or traction boundary conditions are readily available. RUC-based approaches, on the other hand, model the actual microstructure of a material's subvolume which is assumed to repeat itself, and therefore often (but not always, cf., Walker et al. [2]) rely on numerical solutions of the governing field equations using finite difference, finite element, or boundary element methods. Symmetry boundary conditions are typically employed in those situations where the RUC possesses planes of material symmetry which are not altered under loading. Periodic boundary conditions must be applied to the RUC in the absence of planes of material symmetry or symmetry-preserving loading. These different boundary conditions ensure that the deformation of an RUC with a specific microstructure remains compatible with the deformation of its neighbors. The use of periodic boundary conditions in conjunction with a multiscale asymptotic expansion of the displacement and stress fields in the RUC forms the basis of the so-called homogenization methods for estimating the effective properties of periodic materials. Reviews and comparisons of the different approaches have been provided by Christensen [3], Aboudi [4], Hollister and Kikuchi [5], Nemat-Nasser and Horii [6], Parton and Kudryavtsev [7], Arnold et al. [8], and Kalamkarov and Kolpakov [9], among others.

In the case of periodic multiphase materials characterized by an arbitrary distribution of different phases within the RUC, the homogenization theory (cf. references [7] and [9] for comprehensive reviews of the theory) is an effective tool for determining the material's effective moduli as it provides the correct periodic boundary conditions that must be applied to the RUC under loading along different directions, thereby avoiding ad hoc assumptions on the deformation of the RUC's bounding surface. This, in turn, allows the determination of all the effective moduli for applications involving multiaxial loading in a rational manner. However, with the exception of simple RUC architectures amenable to analytical solutions, the evaluation of the effective moduli is per-

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 21, 2000; final revision, February 12, 2001. Editor: L. T. Wheeler. Discussion on the paper should be addressed to the Editor, Prof. 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.

formed by conducting a numerical analysis of the RUC, typically based on the finite element approach. For example, the homogenization technique has been employed in conjunction with the finite element procedure by Banks-Sills et al. [10] to analyze composites in which the repeating unit cell has no planes of material symmetry.

Herein, a new analytical model for the determination of the effective moduli of periodic multiphase materials, characterized by constituent phases continuous along one direction, is presented that circumvents the use of a finite element solution for the displacement and stress fields within an RUC. The model's analytical framework is based on the homogenization technique, but the method of solution for the local displacement and stress fields within an RUC borrows concepts previously employed by the authors in constructing the higher-order theory for functionally graded materials, (Aboudi et al. [11]). It is shown that the model is capable of accurately estimating both the effective elastic moduli of periodic composites and the local stress and strain fields in the individual constituents. The fully analytical nature results in closed-form expressions for all the effective moduli of a periodic multiphase material under multiaxial loading which, in turn, can be incorporated into a finite element analysis of a structural component. Incorporation of the model into a structural analysis code is further facilitated by the ease of an RUC's construction. Periodic multiphase materials with arbitrary phase distributions in the plane normal to the direction of continuous reinforcement can be accommodated, whose response in this plane is fully anisotropic.

## 2 Theoretical Framework

The homogenization approach is employed herein to construct the displacement field approximation at the local microstructural level of a multiphase periodic material in a consistent fashion, and to derive the governing field equations and the boundary conditions that the displacement field must satisfy. The elements of the homogenization technique employed in constructing our theory are outlined first in Section 2.1. Section 2.2 defines the strain concentration tensor employed in the calculation of the effective stiffness tensor of a multiphase periodic material discussed in Section 2.3. The method of determining the displacement, and thus strain and stress, fields within the repeating unit cell is described in Section 2.4. Section 2.5 outlines how the local strain concentration tensors are obtained once the displacement field within the repeating unit cell is known, and how these tensors are used to construct the macroscopic constitutive equation for the given unit cell in terms of the effective elastic stiffness tensor.

**2.1 Elements of the Homogenization Theory.** Consider a multiphase composite wherein the microstructure is periodically distributed in the plane  $x_2$ – $x_3$  defined by the global coordinates  $(x_2, x_3)$ ; see Fig. 1 where the repeating unit cell used to construct the periodic array is highlighted. In the framework of the homogenization method the displacements are asymptotically expanded as follows:

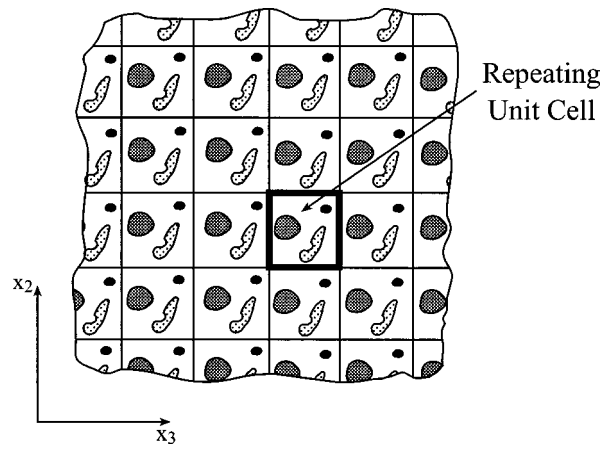
$$u_i(\mathbf{x}, \mathbf{y}) = u_{0i}(\mathbf{x}, \mathbf{y}) + \delta u_{1i}(\mathbf{x}, \mathbf{y}) + \delta^2 u_{2i}(\mathbf{x}, \mathbf{y}) + \dots \quad (1)$$

where  $\mathbf{x} = (x_1, x_2, x_3)$  are the macroscopic (global) coordinates, and  $\mathbf{y} = (y_1, y_2, y_3)$  are the microscopic (local) coordinates that are defined with respect to the repeating unit cell. The size of the unit cell is further assumed to be much smaller than the size of the body so that the relation between the global and local systems is

$$y_i = \frac{x_i}{\delta} \quad (2)$$

where  $\delta$  is a small scaling parameter characterizing the size of the unit cell. This implies that a movement of order unity on the local scale corresponds to a very small movement on the global scale.

The material's periodicity imposes the following constraint on the different-order terms  $u_{\alpha i}$  ( $\alpha = 1, 2, \dots$ ) in Eq. (1):



**Fig. 1 A multiphase composite with a periodic microstructure in the  $x_2$ – $x_3$  plane characterized by a repeating unit cell (high-lighted)**

$$u_{\alpha i}(\mathbf{x}, \mathbf{y}) = u_{\alpha i}(\mathbf{x}, \mathbf{y} + \mathbf{d}) \quad (3)$$

where  $\mathbf{d}$  is the local length scale characterizing the material's periodicity.

The displacement  $u_{0i}$  is the average value of  $u_i$  and hence it is not a function of  $y_i$ . Indeed, it can be formally shown that  $u_{0i}$  is a function of  $x_i$  only by employing the equilibrium equations in conjunction with the Hooke's law (cf. Kalamkarov and Kolpakov [9]).

Let

$$u_{0i} = u_{0i}(\mathbf{x}) \equiv \bar{u}_i \quad (4)$$

and

$$u_{1i} \equiv \tilde{u}_i(\mathbf{x}, \mathbf{y}) \quad (5)$$

where the first-order terms are the fluctuating displacements which are unknown periodic functions. These displacements arise due to the heterogeneity of the medium.

Due to the change of coordinates from the global to the local systems the following relation must be employed in evaluating the derivative of a field quantity:

$$\frac{\partial}{\partial x_i} \rightarrow \frac{\partial}{\partial x_i} + \frac{1}{\delta} \frac{\partial}{\partial y_i} \quad (6)$$

Therefore, upon employing Eq. (6), the strain components are determined from the displacement expansion (1) in the following form:

$$\epsilon_{ij} = \bar{\epsilon}_{ij}(\mathbf{x}) + \tilde{\epsilon}_{ij}(\mathbf{x}, \mathbf{y}) + O(\delta) \quad (7)$$

where

$$\bar{\epsilon}_{ij}(\mathbf{x}) = \frac{1}{2} \left( \frac{\partial \bar{u}_i}{\partial x_j} + \frac{\partial \bar{u}_j}{\partial x_i} \right) \quad (8)$$

and

$$\tilde{\epsilon}_{ij}(\mathbf{x}, \mathbf{y}) = \frac{1}{2} \left( \frac{\partial \tilde{u}_i}{\partial y_j} + \frac{\partial \tilde{u}_j}{\partial y_i} \right) \quad (9)$$

This shows that the strain components can be represented as a sum of the average strain  $\bar{\epsilon}_{ij}(\mathbf{x})$  in the composite and a fluctuating strain  $\tilde{\epsilon}_{ij}(\mathbf{x}, \mathbf{y})$ . It can be easily shown that

$$\frac{1}{V_y} \int \epsilon_{ij} dV_y = \frac{1}{V_y} \int (\bar{\epsilon}_{ij} + \tilde{\epsilon}_{ij}) dV_y = \bar{\epsilon}_{ij}$$

where  $V_y$  is the volume of the repeating unit cell. This follows directly from the periodicity of the fluctuating strain, implying that the average of the fluctuating strain taken over the unit re-

peating cell vanishes. For a homogeneous material it is obvious that the fluctuating displacements and strains vanish.

Using (7), one can readily represent the displacements in the form

$$u_i(\mathbf{x}, \mathbf{y}) = \bar{\epsilon}_{ij}x_j + \tilde{u}_i + O(\delta^2). \quad (10)$$

This representation will be employed in constructing an approximate displacement field for the solution of the cell problem discussed in Section 2.4.

For an elastic material the stresses are related to the strains according to the Hooke's law as follows:

$$\sigma_{ij} = C_{ijkl}\epsilon_{kl} \quad (11)$$

where  $C_{ijkl}(\mathbf{x})$  are the components of the stiffness tensor of the composite's phases. The stiffness tensor forms a periodic function that is defined in the unit repeating cell in terms of the local coordinates  $\mathbf{y}$  such that

$$C_{ijkl}(\mathbf{x}) = C_{ijkl}(\mathbf{y}). \quad (12)$$

Substituting (7) into (11) and differentiating with respect to the microvariable coordinates  $y_i$  leads to

$$\frac{\partial}{\partial y_j} C_{ijkl}(\mathbf{y}) [\bar{\epsilon}_{kl}(\mathbf{x}) + \tilde{\epsilon}_{kl}(\mathbf{x}, \mathbf{y})] = 0, \quad (13)$$

Let us define the following stress quantities:

$$\sigma_{ij}^0 = C_{ijkl}(\mathbf{y}) \bar{\epsilon}_{kl}(\mathbf{x}) \quad (14)$$

$$\sigma_{ij}^1 = C_{ijkl}(\mathbf{y}) \tilde{\epsilon}_{kl}(\mathbf{x}, \mathbf{y}) \quad (15)$$

with the latter being the fluctuating stresses. It follows then that

$$\frac{\partial \sigma_{ij}^1}{\partial y_j} + \frac{\partial \sigma_{ij}^0}{\partial y_j} = 0 \quad (16)$$

which is the strong form of the equilibrium equations. It is readily seen that the first term in (16) involves the unknown fluctuating periodic displacements  $\tilde{u}_i$ , while the second term produces pseudo-body forces whose derivatives are actually zero everywhere except at the interfaces between the phases.

For given values of the average strains  $\bar{\epsilon}_{kl}$ , the unknown fluctuating displacements are governed by (16) subject to periodic boundary conditions that are prescribed at the boundaries of the repeating unit cell. In addition to these boundary conditions one needs to impose the continuity of displacements and tractions at the internal interfaces between the phases that fill the repeating unit cell.

Suppose that the repeating unit cell is given in the plane  $y_2 - y_3$  (with fibers oriented in the 1-direction) by a rectangle defined with respect to the local coordinates by  $0 \leq y_2 \leq H, 0 \leq y_3 \leq L$ , Fig. 2(a). Consequently, the periodic boundary conditions are given by

$$\begin{aligned} \tilde{u}_i(y_2=0) &= \tilde{u}_i(y_2=H) \\ \sigma_{2i}(y_2=0) &= \sigma_{2i}(y_2=H) \end{aligned} \quad (17)$$

and

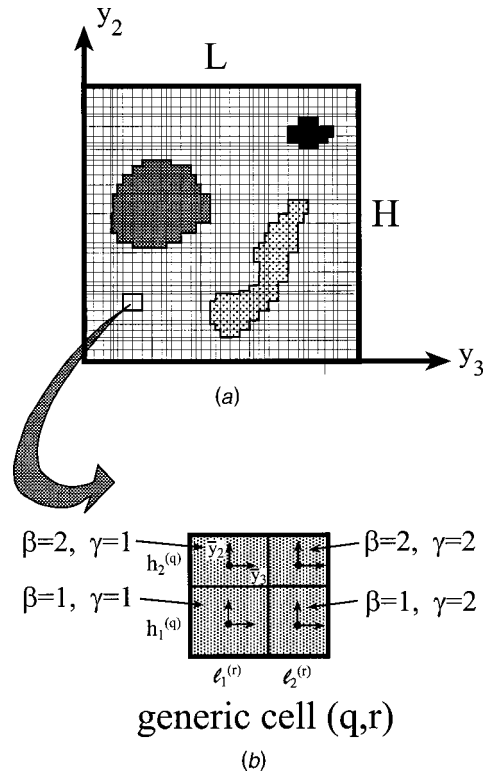
$$\begin{aligned} \tilde{u}_i(y_3=0) &= \tilde{u}_i(y_3=L) \\ \sigma_{3i}(y_3=0) &= \sigma_{3i}(y_3=L) \end{aligned} \quad (18)$$

where the total stress (which is given by (11)) is expressed as

$$\sigma_{ij} = \sigma_{ij}^0 + \sigma_{ij}^1. \quad (19)$$

It is also necessary to fix the displacement field at a point in the repeating unit cell.

**2.2 Strain Concentration Tensor.** Once the solution of (16), subject to the internal interfacial continuity conditions and periodic boundary conditions (17)–(18) has been established, one can proceed to determine the strain concentration tensor associ-



**Fig. 2 (a) Volume discretization of the repeating unit cell employed in the present model, (b) generic cell within the repeating unit cell**

ated with the defined repeating unit cell. This tensor expresses the local strain in the cell in terms of the external average strain. To this end, let us define the fourth-order tensor  $\tilde{\mathbf{A}}$  which relates the fluctuating strain to the average strain

$$\tilde{\epsilon} = \tilde{\mathbf{A}}(\mathbf{y}) \bar{\epsilon}. \quad (20)$$

Using Eq. (7), we readily obtain the required strain concentration tensor  $\mathbf{A}(\mathbf{y})$  as follows:

$$\epsilon = \bar{\epsilon} + \tilde{\mathbf{A}}(\mathbf{y}) \bar{\epsilon} = [\mathbf{I}_4 + \tilde{\mathbf{A}}(\mathbf{y})] \bar{\epsilon} \equiv \mathbf{A}(\mathbf{y}) \bar{\epsilon} \quad (21)$$

where  $\mathbf{I}_4$  is the fourth-order identity tensor:

$$(I_4)_{ijkl} = \frac{1}{2} (\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk})$$

with  $\delta_{ij}$  being the Kronecker delta.

To obtain the strain concentration tensor  $\mathbf{A}(\mathbf{y})$  a series of problems must be solved as follows. Solve Eqs. (16) in conjunction with the internal interfacial and periodic boundary conditions with  $\bar{\epsilon}_{11} = 1$  and all other components set to zero. The solution of (16) readily provides  $A_{ij11}$  for  $i, j = 1, 2, 3$ . This procedure is repeated with  $\bar{\epsilon}_{22} = 1$  and all other components set to zero, which provides  $A_{ij22}$ , and so on.

**2.3 Effective Stiffness Tensor.** Once the strain concentration tensor  $\mathbf{A}(\mathbf{y})$  has been determined, it is possible to compute the effective stiffness tensor of the multiphase composite as follows. Substitution of  $\epsilon$  given by (21) in (11) yields

$$\sigma = \mathbf{C}(\mathbf{y}) \mathbf{A}(\mathbf{y}) \bar{\epsilon}. \quad (22)$$

Taking the average of both sides of Eq. (22) over the repeating unit cell yields the average stress in the composite in terms of the average strain via the effective elastic stiffness tensor  $\mathbf{C}^*$ , namely

$$\bar{\sigma} = \mathbf{C}^* \bar{\epsilon}$$



where

$$\mathbf{C}^* = \frac{1}{V_y} \int \mathbf{C}(\mathbf{y}) \mathbf{A}(\mathbf{y}) dV \quad (23)$$

**2.4 Method of Solution for the Cell Problem.** In this section we present a solution methodology of Eqs. (16) for two-dimensional multiphase elastic composites. In this case the repeating unit cell extends over  $0 \leq y_2 \leq H$ ,  $0 \leq y_3 \leq L$  given in terms of the local coordinates  $(y_2, y_3)$  as stated previously. The microstructure in the  $y_2$ – $y_3$  plane of the composite is modeled by discretizing the cross section of the repeating unit cell into  $N_q$  and  $N_r$  internal or generic cells in the intervals  $0 \leq y_2 \leq H$  and  $0 \leq y_3 \leq L$ , respectively. Figure 2(a) illustrates how the repeating unit cell highlighted in Fig. 1 could be discretized. In addition, every generic cell consists of four subcells designated by the pair  $(\beta\gamma)$  where each index takes the values 1 or 2 which indicate the relative position of the given subcell along the  $y_2$  and  $y_3$ -axis, respectively, see Fig. 2(b). The indices  $q$  and  $r$ , whose ranges are  $q = 1, 2, \dots, N_q$  and  $r = 1, 2, \dots, N_r$ , identify the generic cell in the  $y_2$ – $y_3$  plane. The dimensions of the generic cell along the  $y_2$  and  $y_3$ -axes are  $h_1^{(q)}$ ,  $h_2^{(q)}$  and  $l_1^{(r)}$ ,  $l_2^{(r)}$ , such that

$$H = \sum_{q=1}^{N_q} (h_1^{(q)} + h_2^{(q)}), \quad L = \sum_{r=1}^{N_r} (l_1^{(r)} + l_2^{(r)}).$$

This manner of discretizing a periodic material's microstructure has also been employed in constructing the "generalized method of cells" (Paley and Aboudi [12]). The construction of the higher-order theory for functionally graded materials characterized by spatially variable microstructures without a definable repeating unit cell is also based on such volume discretization (Aboudi et al. [11]).

Given an applied macroscopic loading, an approximate solution for the displacements field is constructed based on volumetric averaging of the field equations together with the imposition of periodic boundary conditions, and both displacement and traction continuity conditions, in an average sense between the cells and subcells used to characterize the material's microstructure. This is accomplished by approximating the fluctuating displacements in each subcell using a quadratic expansion in terms of local coordinates  $(\bar{y}_2^{(\beta)}, \bar{y}_3^{(\gamma)})$  centered at the subcell's midpoint. A higher-order representation of the fluctuating displacement field is necessary in order to capture the local effects created by the field gradients and the microstructure of the composite. This is in sharp contrast with the generalized method of cells where the displacement expansion was linear as a result of which the coupling between the local normal and shear effects was lost.

In the current development, we include spatially uniform thermal loading characterized by the temperature deviation  $\Delta T$  from a reference temperature, in addition to mechanical loading. Further, the phases corresponding to the individual subcells are taken as orthotropic. Thus the constitutive relation of the material filling subcell  $(\beta, \gamma)$  is obtained by generalizing Eq. (11) as follows:

$$\sigma_{ij}^{(\beta\gamma)} = C_{ijkl}^{(\beta\gamma)} (\epsilon_{kl}^{(\beta\gamma)} - \epsilon_{kl}^{T(\beta\gamma)}) \quad (24)$$

where  $C_{ijkl}^{(\beta\gamma)}$  are the elements of the stiffness tensor of the phase filling subcell  $(\beta\gamma)$  with nine independent constants,  $\epsilon_{kl}^{(\beta\gamma)}$  are the total strains which are given by (7), and  $\epsilon_{kl}^{T(\beta\gamma)}$  are the thermal strains in these subcells, with no summation implied by repeated Greek letters in the above and henceforth. Equation (24) can be rewritten in the form

$$\sigma_{ij}^{(\beta\gamma)} = C_{ijkl}^{(\beta\gamma)} \epsilon_{kl}^{(\beta\gamma)} - \sigma_{ij}^{T(\beta\gamma)} \quad (25)$$

where the term  $\sigma_{ij}^{T(\beta\gamma)}$ , henceforth referred to as thermal stress, stands for the thermal contribution

$$\sigma_{ij}^{T(\beta\gamma)} = \Gamma_{ij}^{(\beta\gamma)} \Delta T \quad (26)$$

where  $\Gamma_{ij}^{(\beta\gamma)}$  are the thermal stress coefficients.

The equilibrium Eqs. (16) in conjunction with (19) for the material occupying the subcell  $(\beta\gamma)$  in the region  $|\bar{y}_2^{(\beta)}| \leq h_\beta^{(q)}/2$ ,  $|\bar{y}_3^{(\gamma)}| \leq l_\gamma^{(r)}/2$  can be written in the form

$$\partial_2 \sigma_{2j}^{(\beta\gamma)} + \partial_3 \sigma_{3j}^{(\beta\gamma)} = 0 \quad j = 1, 2, 3 \quad (27)$$

where  $\partial_2 = \partial/\partial \bar{y}_2^{(\beta)}$  and  $\partial_3 = \partial/\partial \bar{y}_3^{(\gamma)}$ .

As stated before, the fluctuating displacement field in the subcell  $(\beta\gamma)$  of the  $(q, r)$ th cell is approximated by a second-order expansion in the local coordinate system. Consequently, according to Eq. (10) the displacements in the subcell can be represented in the form (omitting the cell label  $(q, r)$ )

$$\begin{aligned} u_1^{(\beta\gamma)} &= \bar{\epsilon}_{1j} x_j + W_{1(00)}^{(\beta\gamma)} + \bar{y}_2^{(\beta)} W_{1(10)}^{(\beta\gamma)} + \bar{y}_3^{(\gamma)} W_{1(01)}^{(\beta\gamma)} \\ &\quad + \frac{1}{2} \left( 3\bar{y}_2^{(\beta)2} - \frac{h_\beta^{(q)2}}{4} \right) W_{1(20)}^{(\beta\gamma)} \\ &\quad + \frac{1}{2} \left( 3\bar{y}_3^{(\gamma)2} - \frac{l_\gamma^{(r)2}}{4} \right) W_{1(02)}^{(\beta\gamma)} \end{aligned} \quad (28)$$

$$\begin{aligned} u_2^{(\beta\gamma)} &= \bar{\epsilon}_{2j} x_j + W_{2(00)}^{(\beta\gamma)} + \bar{y}_2^{(\beta)} W_{2(10)}^{(\beta\gamma)} + \bar{y}_3^{(\gamma)} W_{2(01)}^{(\beta\gamma)} \\ &\quad + \frac{1}{2} \left( 3\bar{y}_2^{(\beta)2} - \frac{h_\beta^{(q)2}}{4} \right) W_{2(20)}^{(\beta\gamma)} \\ &\quad + \frac{1}{2} \left( 3\bar{y}_3^{(\gamma)2} - \frac{l_\gamma^{(r)2}}{4} \right) W_{2(02)}^{(\beta\gamma)} \end{aligned} \quad (29)$$

$$\begin{aligned} u_3^{(\beta\gamma)} &= \bar{\epsilon}_{3j} x_j + W_{3(00)}^{(\beta\gamma)} + \bar{y}_2^{(\beta)} W_{3(10)}^{(\beta\gamma)} + \bar{y}_3^{(\gamma)} W_{3(01)}^{(\beta\gamma)} \\ &\quad + \frac{1}{2} \left( 3\bar{y}_2^{(\beta)2} - \frac{h_\beta^{(q)2}}{4} \right) W_{3(20)}^{(\beta\gamma)} \\ &\quad + \frac{1}{2} \left( 3\bar{y}_3^{(\gamma)2} - \frac{l_\gamma^{(r)2}}{4} \right) W_{3(02)}^{(\beta\gamma)} \end{aligned} \quad (30)$$

where  $W_{i(00)}^{(\beta\gamma)}$ , which are the volume-averaged fluctuating displacements and the higher-order terms  $W_{i(mn)}^{(\beta\gamma)}$  ( $i = 1, 2, 3$ ) must be determined, as shown below, from the equilibrium Eqs. (27) as well as the periodic boundary conditions (17)–(18) that the fluctuating displacements must fulfill, as well as the interfacial continuity conditions between subcells. The number of unknowns that describe the fluctuating displacements in the cell  $(q, r)$  is 60.

The total strains in the subcell  $(\beta\gamma)$  are given by (7) in conjunction with (8) and (9), namely

$$\epsilon_{ij}^{(\beta\gamma)} = \bar{\epsilon}_{ij} + \frac{1}{2} (\partial_i u_j^{(\beta\gamma)} + \partial_j u_i^{(\beta\gamma)}) \quad (31)$$

where  $\partial_1 = 0$  and  $\partial_2, \partial_3$  have been defined previously. Therefore, the volume-averaged strains  $\bar{\epsilon}_{ij(m,n)}^{(\beta\gamma)}$  in the subcell  $(\beta\gamma)$  of the cell  $(q, r)$  are obtained in terms of the displacement field (28)–(30) as follows (omitting  $(q, r)$ ):

$$\begin{aligned} \bar{\epsilon}_{11}^{(\beta\gamma)} &= \bar{\epsilon}_{11} \\ \bar{\epsilon}_{22}^{(\beta\gamma)} &= \bar{\epsilon}_{22} + W_{2(10)}^{(\beta\gamma)} \\ \bar{\epsilon}_{33}^{(\beta\gamma)} &= \bar{\epsilon}_{33} + W_{3(01)}^{(\beta\gamma)} \\ \bar{\epsilon}_{23}^{(\beta\gamma)} &= \bar{\epsilon}_{23} + \frac{1}{2} (W_{2(01)}^{(\beta\gamma)} + W_{3(10)}^{(\beta\gamma)}) \\ \bar{\epsilon}_{13}^{(\beta\gamma)} &= \bar{\epsilon}_{13} + \frac{1}{2} W_{1(01)}^{(\beta\gamma)} \\ \bar{\epsilon}_{12}^{(\beta\gamma)} &= \bar{\epsilon}_{12} + \frac{1}{2} W_{1(10)}^{(\beta\gamma)}. \end{aligned} \quad (32)$$

In the course of satisfying the equilibrium equations in a volumetric sense, it is convenient to define the following stress quantities:

$$[S_{ij(m,n)}^{(\beta\gamma)}]^{(q,r)} = \frac{1}{h_\beta^{(q)} l_\gamma^{(r)}} \int_{-h_\beta^{(q)}/2}^{h_\beta^{(q)}/2} \int_{-l_\gamma^{(r)}/2}^{l_\gamma^{(r)}/2} (\bar{y}_2^{(\beta)})^m (\bar{y}_3^{(\gamma)})^n \sigma_{ij}^{(\beta\gamma)} d\bar{y}_2^{(\beta)} d\bar{y}_3^{(\gamma)}. \quad (33)$$

For  $m=n=0$ , Eq. (33) provides the average stresses in the subcell, whereas for other values of  $(m, n)$  higher-order stresses are obtained that are needed to describe the governing field equations of the continuum. These stress quantities can be evaluated explicitly in terms of the unknown coefficients  $W_{i(mn)}^{(\beta\gamma)}$  by performing the required volume integration upon substituting Eqs. (25), (31), and (28)–(30) in Eq. (33). This yields the following nonvanishing zeroth-order and first-order stress-components in terms of the unknown coefficients in the displacement field expansion (omitting  $(q, r)$ ):

$$S_{11(0,0)}^{(\beta\gamma)} = C_{11}^{(\beta\gamma)} \bar{\epsilon}_{11} + C_{12}^{(\beta\gamma)} (W_{2(10)}^{(\beta\gamma)} + \bar{\epsilon}_{22}) + C_{13}^{(\beta\gamma)} (W_{3(01)}^{(\beta\gamma)} + \bar{\epsilon}_{33}) - \Gamma_1^{(\beta\gamma)} \Delta T \quad (34)$$

$$S_{11(1,0)}^{(\beta\gamma)} = \frac{1}{4} h_\beta^{(q)2} C_{12}^{(\beta\gamma)} W_{2(20)}^{(\beta\gamma)} \quad (35)$$

$$S_{11(0,1)}^{(\beta\gamma)} = \frac{1}{4} l_\gamma^{(r)2} C_{13}^{(\beta\gamma)} W_{3(02)}^{(\beta\gamma)} \quad (36)$$

with similar expressions for the other normal stress components, and

$$S_{23(0,0)}^{(\beta\gamma)} = C_{44}^{(\beta\gamma)} (2\bar{\epsilon}_{23} + W_{2(01)}^{(\beta\gamma)} + W_{3(10)}^{(\beta\gamma)}) \quad (37)$$

$$S_{23(1,0)}^{(\beta\gamma)} = \frac{1}{4} h_\beta^{(q)2} C_{44}^{(\beta\gamma)} W_{3(20)}^{(\beta\gamma)} \quad (38)$$

$$S_{23(0,1)}^{(\beta\gamma)} = \frac{1}{4} l_\gamma^{(r)2} C_{44}^{(\beta\gamma)} W_{2(02)}^{(\beta\gamma)} \quad (39)$$

$$S_{13(0,0)}^{(\beta\gamma)} = C_{55}^{(\beta\gamma)} (2\bar{\epsilon}_{13} + W_{1(01)}^{(\beta\gamma)}) \quad (40)$$

$$S_{13(0,1)}^{(\beta\gamma)} = \frac{1}{4} l_\gamma^{(r)2} C_{55}^{(\beta\gamma)} W_{1(02)}^{(\beta\gamma)} \quad (41)$$

$$S_{12(0,0)}^{(\beta\gamma)} = C_{66}^{(\beta\gamma)} (2\bar{\epsilon}_{12} + W_{1(10)}^{(\beta\gamma)}) \quad (42)$$

$$S_{12(1,0)}^{(\beta\gamma)} = \frac{1}{4} h_\beta^{(q)2} C_{66}^{(\beta\gamma)} W_{1(20)}^{(\beta\gamma)} \quad (43)$$

where contracted notation has been employed for the stiffness elements  $C_{ijkl}^{(\beta\gamma)}$ .

Subsequently, satisfaction of the zeroth, first, and second moments of the equilibrium Eqs. (27) results in the following 12 relations among the volume-averaged first-order stresses  $S_{ij(m,n)}^{(\beta\gamma)}$  in the different subcells  $(\beta\gamma)$  of the  $(q, r)$  cell, after lengthy algebraic manipulations

$$[S_{2j(1,0)}^{(\beta\gamma)} h_\beta^2 + S_{3j(0,1)}^{(\beta\gamma)} l_\gamma^2]^{(q,r)} = 0 \quad j=1,2,3. \quad (44)$$

The continuity of tractions at the subcell interfaces and between adjacent cells, imposed in an average sense, can be shown to be ensured by the following relations:

$$[-12S_{2j(1,0)}^{(1\gamma)}/h_1 + S_{2j(0,0)}^{(2\gamma)} - 6S_{2j(1,0)}^{(2\gamma)}/h_2]^{(q,r)} - [S_{2j(0,0)}^{(2\gamma)} + 6S_{2j(1,0)}^{(2\gamma)}/h_2]^{(q-1,r)} = 0 \quad (45)$$

$$\left[ -S_{2j(0,0)}^{(1\gamma)} + \frac{1}{2} S_{2j(0,0)}^{(2\gamma)} - 3S_{2j(1,0)}^{(2\gamma)}/h_2 \right]^{(q,r)} + \frac{1}{2} [S_{2j(0,0)}^{(2\gamma)} + 6S_{2j(1,0)}^{(2\gamma)}/h_2]^{(q-1,r)} = 0 \quad (46)$$

$$[-12S_{3j(0,1)}^{(\beta 1)}/l_1 + S_{3j(0,0)}^{(\beta 1)} - 6S_{3j(0,1)}^{(\beta 2)}/l_2]^{(q,r)} - [S_{3j(0,0)}^{(\beta 2)} + 6S_{3j(0,1)}^{(\beta 2)}/l_2]^{(q,r-1)} = 0 \quad (47)$$

$$\left[ -S_{3j(0,0)}^{(\beta 1)} + \frac{1}{2} S_{3j(0,0)}^{(\beta 2)} - 3S_{3j(0,1)}^{(\beta 2)}/l_2 \right]^{(q,r)} + \frac{1}{2} [S_{3j(0,0)}^{(\beta 2)} + 6S_{3j(0,1)}^{(\beta 2)}/l_2]^{(q,r-1)} = 0 \quad (48)$$

where  $j=1, 2$ , and  $3$ . The details of derivation of equations similar to Eqs. (44)–(48) have been provided by Aboudi et al. [13] for a more general case of a microstructure containing periodic inclusions in the out-of-plane direction.

Equations (45)–(48) provide 24 additional relations among the zeroth-order and first-order stresses. These relations together with Eq. (44), can be expressed in terms of the unknown coefficients  $W_{i(mn)}^{(\beta\gamma)}$  by making use of Eqs. (34)–(43), providing a total of 36 of the required 60 equations necessary for the determination of these coefficients in the cell  $(q, r)$ .

The additional 24 relations necessary to determine the unknown coefficients in the displacement field expansion are subsequently obtained by imposing displacement continuity conditions on an average basis at each subcell and cell interface. This produces

$$\left[ W_{j(00)}^{(1\gamma)} + \frac{1}{2} h_1 W_{j(10)}^{(1\gamma)} + \frac{1}{4} h_1^2 W_{j(20)}^{(1\gamma)} \right]^{(q,r)} = \left[ W_{j(00)}^{(2\gamma)} - \frac{1}{2} h_2 W_{j(10)}^{(2\gamma)} + \frac{1}{4} h_2^2 W_{j(20)}^{(2\gamma)} \right]^{(q,r)} \quad (49)$$

$$\left[ W_{j(00)}^{(2\gamma)} + \frac{1}{2} h_2 W_{j(10)}^{(2\gamma)} + \frac{1}{4} h_2^2 W_{j(20)}^{(2\gamma)} \right]^{(q,r)} = \left[ W_{j(00)}^{(1\gamma)} - \frac{1}{2} h_1 W_{j(10)}^{(1\gamma)} + \frac{1}{4} h_1^2 W_{j(20)}^{(1\gamma)} \right]^{(q+1,r)} \quad (50)$$

$$\left[ W_{j(00)}^{(\beta 1)} + \frac{1}{2} l_1 W_{j(01)}^{(\beta 1)} + \frac{1}{4} l_1^2 W_{j(02)}^{(\beta 1)} \right]^{(q,r)} = \left[ W_{j(00)}^{(\beta 2)} - \frac{1}{2} l_2 W_{j(01)}^{(\beta 2)} + \frac{1}{4} l_2^2 W_{j(02)}^{(\beta 2)} \right]^{(q,r)} \quad (51)$$

$$\left[ W_{j(00)}^{(\beta 2)} + \frac{1}{2} l_2 W_{j(01)}^{(\beta 2)} + \frac{1}{4} l_2^2 W_{j(02)}^{(\beta 2)} \right]^{(q,r)} = \left[ W_{j(00)}^{(\beta 1)} - \frac{1}{2} l_1 W_{j(01)}^{(\beta 1)} + \frac{1}{4} l_1^2 W_{j(02)}^{(\beta 1)} \right]^{(q,r+1)} \quad (52)$$

where  $j=1, 2$ , and  $3$ , which comprise the required additional 24 relation.

The equilibrium relations, Eqs. (44), together with the traction and displacement continuity conditions Eqs. (45)–(48) and (49)–(52), respectively, form 60 equations in the 60 unknowns  $W_{i(mn)}^{(\beta\gamma)}$  which govern the equilibrium of a subcell  $(\beta\gamma)$  within an interior cell  $(q, r)$ ;  $q=2, \dots, N_q-1$ ,  $r=2, \dots, N_r-1$ . For the boundary cells  $q=1, N_q$  and  $r=1, N_r$  a different treatment must be applied.

For cell  $(1, r)$ , the above relations are operative, except Eqs. (45) and (46), which follow from the continuity of tractions between a given cell and the preceding one. These 12 equations must be replaced by the conditions of continuity of tractions at the interior interfaces of the cell  $(1, r)$  (imposed in the average sense), and by

the conditions that the fluctuating displacements are periodic. It follows, by imposing the first relations in (17) on an average basis, that

$$\left[ W_{j(00)}^{(1\gamma)} - \frac{1}{2} h_1 W_{j(10)}^{(1\gamma)} + \frac{1}{4} h_1^2 W_{j(20)}^{(1\gamma)} \right]^{(1,r)} = \left[ W_{j(00)}^{(2\gamma)} + \frac{1}{2} h_2 W_{j(10)}^{(2\gamma)} + \frac{1}{4} h_2^2 W_{j(20)}^{(2\gamma)} \right]^{(N_q, r)} \quad (53)$$

where  $j=1, 2$ , and  $3$ . Both conditions provide the required 12 relations to be used for cell  $(1, r)$ .

For cell  $(N_q, r)$ , the previously derived governing equations are operative except for the six relations given by Eqs. (50), which are obviously not applicable. These are replaced by the conditions that the tractions are periodic. Thus the second relations in (17), imposed on an average basis, provide the six equations to be used for cell  $(N_q, r)$  as follows:

$$[\sigma_{2j}^{(1\gamma)}]_{\bar{y}_2^{(1)} = -h_1/2}^{(1,r)} = [\sigma_{2j}^{(2\gamma)}]_{\bar{y}_2^{(2)} = h_2/2}^{(N_q, r)} \quad (54)$$

where the stresses  $\sigma_{ij}^{(\beta\gamma)}$  are given by Eq. (25).

Similar treatments hold for boundary cells  $(q, 1)$  and  $(q, N_r)$ . Thus the 12 equations (47)–(48) are obviously not applicable in cell  $(q, 1)$  and should be replaced by the conditions of continuity of tractions at the interior interfaces of this cell (imposed in the average sense), and by the conditions that the fluctuating displacements are periodic. The latter yield according to the first equation in (18),

$$\left[ W_{j(00)}^{(\beta 1)} - \frac{1}{2} l_1 W_{j(01)}^{(\beta 1)} + \frac{1}{4} l_1^2 W_{j(02)}^{(\beta 1)} \right]^{(q, 1)} = \left[ W_{j(00)}^{(\beta 2)} + \frac{1}{2} l_2 W_{j(01)}^{(\beta 2)} + \frac{1}{4} l_2^2 W_{j(02)}^{(\beta 2)} \right]^{(q, N_r)} \quad (55)$$

For boundary cell  $(q, N_r)$ , Eqs. (52) are not operative and they should be replaced by the periodicity of tractions which is given by the second relation in (18) that is imposed on the average basis:

$$[\sigma_{3j}^{(\beta 1)}]_{\bar{y}_3^{(1)} = -l_1/2}^{(q, 1)} = [\sigma_{3j}^{(\beta 2)}]_{\bar{y}_3^{(2)} = l_2/2}^{(q, N_r)} \quad (56)$$

Consequently, the governing equations for the interior and boundary cells form a system of  $60N_q N_r$  algebraic equations in the unknown coefficients  $W_{i(mn)}^{(\beta\gamma)}$ . The final form of this system of equations can be symbolically represented by

$$\mathbf{KU} = \mathbf{f} \quad (57)$$

where the structural stiffness matrix  $\mathbf{K}$  contains information on the geometry and thermomechanical properties of the materials within the individual subcells  $(\beta\gamma)$  of the cells comprising the multiphase periodic composite. The displacement vector  $\mathbf{U}$  contains the unknown displacement coefficients in each subcell, i.e.,

$$\mathbf{U} = [\mathbf{U}_{11}^{(11)}, \dots, \mathbf{U}_{N_q N_r}^{(22)}] \quad (58)$$

where in subcell  $(\beta\gamma)$  of cell  $(q, r)$  these coefficients are

$$\mathbf{U}_{qr}^{(\beta\gamma)} = [W_{i(00)}, W_{i(10)}, W_{i(01)}, W_{i(20)}, W_{i(02)}]_{qr}^{(\beta\gamma)} \quad i=1, 2, 3.$$

The mechanical force  $\mathbf{f}$  contains information on the applied average strains  $\bar{\epsilon}_{ij}$  and the imposed temperature deviation  $\Delta T$ .

A careful check of the preceding equations reveals that the equations that govern the normal and in-plane (2–3) shear deformations are coupled, thus providing the necessary shear coupling effects. On the other hand, these equations are not coupled to the axial shear deformations (1–2 and 1–3). Thus the above system, Eq. (57), can be decoupled in practical applications and solved for the normal and transverse shear deformations (with  $40N_q N_r$  algebraic equations) separately from the axial shear deformations (with  $20N_q N_r$  algebraic equations).

**2.5 Global Constitutive Relations.** Once the solution  $\mathbf{U}$  for a given set of average strains  $\bar{\epsilon}$  has been established, we can determine, in particular, the average strains  $[\bar{\epsilon}^{(\beta\gamma)}]^{(q, r)}$  in subcell  $(\beta\gamma)$  of the cell  $(q, r)$  given by (32).

The average stress components  $[S_{ij(0,0)}^{(\beta\gamma)}]^{(q, r)}$  in subcell  $(\beta\gamma)$  of the cell  $(q, r)$  are given by Eqs. (33), (36), (39), and (41). They can be assembled in a compact form as follows:

$$[S_{(0,0)}^{(\beta\gamma)}]^{(q, r)} = [\mathbf{C}^{(\beta\gamma)} \bar{\epsilon}^{(\beta\gamma)} - \mathbf{\Gamma}^{(\beta\gamma)} \Delta T]^{(q, r)}. \quad (59)$$

Let us generalize the localization Eq. (21) (that was given in the elastic case) to the present situation which involves thermoelastic effects and write it in the form

$$[\bar{\epsilon}^{(\beta\gamma)}]^{(q, r)} = [\mathbf{A}^{(\beta\gamma)} \bar{\epsilon} + \mathbf{D}^{(\beta\gamma)}]^{(q, r)} \quad (60)$$

where  $[\mathbf{A}^{(\beta\gamma)}]^{(q, r)}$  is the mechanical strain concentration matrix of the subcell  $(\beta\gamma)$ , and  $[\mathbf{D}^{(\beta\gamma)}]^{(q, r)}$  is a vector that involves thermal effects in the subcell. In the absence of thermal effects this vector vanishes, and we can readily determine from (60) the mechanical strain concentration matrix  $[\mathbf{A}^{(\beta\gamma)}]^{(q, r)}$  by solving the system (57) six consecutive times upon imposing a single nonzero component of  $\bar{\epsilon}$  one at a time.

The thermal analysis is performed by imposing a spatially uniform temperature  $\Delta T$ . Thus for a given value of applied thermal load, the average strains  $[\bar{\epsilon}^{(\beta\gamma)}]^{(q, r)}$  in the subcell are also obtained from the solution of Eq. (57), and hence the matrix  $[\mathbf{D}^{(\beta\gamma)}]^{(q, r)}$  from (60) in the absence of  $\bar{\epsilon}$ .

Substitution of (60) into (59) yields

$$[S_{(0,0)}^{(\beta\gamma)}]^{(q, r)} = [\mathbf{C}^{(\beta\gamma)} (\mathbf{A}^{(\beta\gamma)} \bar{\epsilon} + \mathbf{D}^{(\beta\gamma)}) - \mathbf{\Gamma}^{(\beta\gamma)} \Delta T]^{(q, r)}. \quad (61)$$

The average stress in the multiphase periodic composite is determined from

$$\bar{\sigma} = \frac{1}{HL} \sum_{q=1}^{N_q} \sum_{r=1}^{N_r} \sum_{\beta, \gamma=1}^2 h_{\beta}^{(q)} l_{\gamma}^{(r)} [S_{(0,0)}^{(\beta\gamma)}]^{(q, r)}. \quad (62)$$

Consequently, Eqs. (61)–(62) establish the effective constitutive law of the multiphase thermoelastic composite in the form

$$\bar{\sigma} = \mathbf{C}^* \bar{\epsilon} - \bar{\sigma}^T \quad (63)$$

where  $\mathbf{C}^*$  is the effective elastic stiffness matrix of the composite which is given by (see also Eq. (23))

$$\mathbf{C}^* = \frac{1}{HL} \sum_{q=1}^{N_q} \sum_{r=1}^{N_r} \sum_{\beta, \gamma=1}^2 h_{\beta}^{(q)} l_{\gamma}^{(r)} [\mathbf{C}^{(\beta\gamma)} \mathbf{A}^{(\beta\gamma)}]^{(q, r)} \quad (64)$$

and  $\bar{\sigma}^T$  denote the overall (macroscopic) thermal stresses in the composite given by

$$\bar{\sigma}^T = \frac{-1}{HL} \sum_{q=1}^{N_q} \sum_{r=1}^{N_r} \sum_{\beta, \gamma=1}^2 h_{\beta}^{(q)} l_{\gamma}^{(r)} [\mathbf{C}^{(\beta\gamma)} \mathbf{D}^{(\beta\gamma)} - \mathbf{\Gamma}^{(\beta\gamma)} \Delta T]^{(q, r)}. \quad (65)$$

Alternatively, the results independently obtained by Levin [14] and Schapery [15] can be employed to express the global thermal stress in terms of the mechanical strain concentration matrices for the individual subcells. The global thermal stress in the multiphase composite  $\bar{\sigma}^T = \mathbf{\Gamma}^* \Delta T$  ( $\mathbf{\Gamma}^*$  is related to the effective coefficients of thermal expansion  $\alpha^*$  of the composite by  $\mathbf{\Gamma}^* = \mathbf{C}^* \alpha^*$ ) is given in accordance with the Levin-Schapery formula in terms of the mechanical strain concentration matrices and the thermal stress vector in the individual phases by

$$\bar{\sigma}^T = \frac{\Delta T}{HL} \sum_{q=1}^{N_q} \sum_{r=1}^{N_r} \sum_{\beta, \gamma=1}^2 h_{\beta}^{(q)} l_{\gamma}^{(r)} [\mathbf{A}^{tr(\beta\gamma)} \mathbf{\Gamma}^{(\beta\gamma)}]^{(q, r)} \quad (66)$$



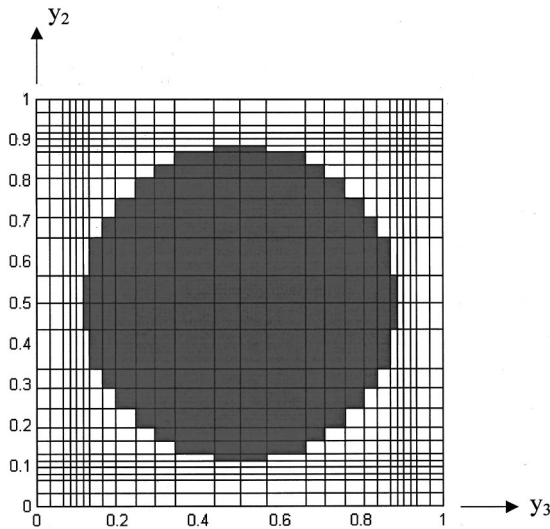
**Table 1 Elastic moduli of constituent fiber and matrix materials**

Material	$E_{11}$ (GPa)	$E_{22}$ (GPa)	$G_{12}$ (GPa)	$\nu_{12}$	$\nu_{23}$	$\alpha_{11}$ ( $10^{-6}/\text{C}$ )	$\alpha_{22}$ ( $10^{-6}/\text{C}$ )
Boron fiber	379.3	379.3	172.41	0.10	0.10	8.1	8.1
Aluminum matrix	68.3	68.3	26.3	0.30	0.30	23.0	23.0
Graphite fiber	235.0	14.0	28.0	0.20	0.25	-	-
Epoxy matrix	4.8	4.8	1.8	0.34	0.34	-	-
Glass fiber	69.0	69.0	28.75	0.20	0.20	-	-

where  $[\mathbf{A}^{tr(\beta\gamma)}]^{(q,r)}$  is the transpose of the mechanical strain concentration matrix  $[\mathbf{A}^{(\beta\gamma)}]^{(q,r)}$  of subcell  $(\beta\gamma)$  within cell  $(q,r)$ . This provides an additional check on the consistency of the proposed approach.

### 3 Numerical Results and Discussion

To illustrate the predictive capability of the developed theory, we first compare the predictions for the effective moduli of unidirectional continuous fiber-reinforced composites with the finite element results reported by Sun and Vaidya [16] and Tamma and Avila [17]. Symmetry and periodic boundary conditions were imposed on the deformation of the repeating unit cell in these investigations depending on the direction of the applied load. The results have been generated for two material systems with substantially different fiber/matrix elastic moduli mismatch. These materials systems are boron/aluminum and graphite/epoxy. Table 1 gives the elastic moduli of the constituent phases for these two unidirectional composites. The fiber volume fractions of the boron/aluminum and graphite/epoxy composites are 0.47 and 0.60, respectively. Figure 3 shows the volume discretization of the repeating unit cell used in the analysis of the boron/aluminum system. It contains  $26 \times 26$  subcells appropriately dimensioned to approximate the circular fiber shape sufficiently well. This cell generates a square array of circular fibers in the matrix phase. The



**Fig. 3 Volume discretization of the repeating unit cell employed in the analysis of a boron/aluminum unidirectional composite with a fiber volume fraction of 0.47**

repeating unit cell for the graphite/epoxy system (not shown) is similar to that shown in Fig. 3 with the only difference being the dimensions of the subcells in the matrix phase to account for the higher volume fraction.

Tables 2, 3, and 4 present comparison between the effective moduli and thermal expansion coefficients predicted by the present theory and the finite-element results for the boron/aluminum and graphite/epoxy systems obtained from the above two references. Included in the tables are the predictions of the generalized method of cells obtained using the same repeating unit cell geometry as shown in Fig. 3. In the case of the boron/aluminum system, Table 2, very good agreement is observed between the present theory's predictions for the effective elastic moduli and the results reported by Sun and Vaidya with the exception of the shear modulus  $G_{12}$  and the Poisson's ratio  $\nu_{23}$  which are somewhat larger. However, the values for these moduli reported by Tamma and Avila are very close to those predicted by the present theory. The generalized method of cells' predictions are generally lower (with the exception of the two Poisson's ratios) than the present theory's predictions, although the differences are quite acceptable. As will be discussed in the sequel, the generalized method of cells' estimate of the subcell stresses is not as accurate as the effective moduli estimates due to the absence of coupling between the normal and shear stresses in the plane  $x_2 - x_3$ . The effective thermal expansion coefficients obtained from the present theory, Table 3, also agree quite well with the results reported by Tamma and Avila. It is remarkable that the generalized method of cells' estimates of the effective thermal expansion coefficients are even closer to the finite element results.

The agreement between the present theory's estimates of the effective moduli and the finite element results of Sun and Vaidya in the case of the graphite/epoxy system, Table 4, is even better than in the preceding case for all moduli. In addition to the higher fiber volume fraction for this case relative to the preceding case, the material property mismatch in the  $x_2 - x_3$  plane is lower, which could explain the better agreement. However, the shear modulus mismatch in the out-of-plane direction (noting that the graphite fiber is transversely isotropic) is now higher. Despite this greater mismatch, the estimates of  $G_{12}$  are very close. The prediction of the generalized method of cells for  $G_{12}$  also compares favorably with both the finite element and the present theory's results.

Despite the established predictive capability of the generalized method of cells to accurately estimate the macroscopic response of unidirectional composites in terms of both the effective moduli and the thermal expansion coefficients, demonstrated above and in a number of previous investigations (see the recent review paper by Aboudi [18] for instance, which also includes inelastic effects), its predictive capability of the local subcell stresses is not as good. As mentioned previously, this is due to the absence of coupling

**Table 2 Comparison of predicted effective elastic moduli of a boron/aluminum unidirectional composite ( $\nu_f = 0.47$ )**

Effective Elastic Moduli	$E_{11}$ (GPa)	$E_{22}$ (GPa)	$G_{12}$ (GPa)	$G_{23}$ (GPa)	$\nu_{12}$	$\nu_{23}$
Present model	215.4	144.0	54.34	45.83	0.195	0.255
Sun and Vaidya [16]	215.0	144.0	57.20	45.90	0.190	0.290
Tamma and Avila [17]	214.7	144.7	54.30	45.60	0.195	0.249
Generalized method of cells [12]	215.0	141.0	51.20	43.70	0.197	0.261

**Table 3 Comparison of predicted effective thermal expansion coefficients of a boron/aluminum unidirectional composite ( $v_f = 0.47$ )**

Effective Thermal Expansion Coefficients	$\alpha_{11}$ ( $10^{-6}/\text{C}$ )	$\alpha_{22}$ ( $10^{-6}/\text{C}$ )
Present model	11.0	16.7
Tamma and Avila [17]	10.77	17.34
Generalized method of cells [12]	10.91	16.94

between the normal and shear strains (or stresses) in the  $x_2$ - $x_3$  plane which, in turn, can be directly traced to the first-order expansion of the displacement field within each subcell and the manner in which the interfacial traction continuity conditions are applied in an average sense. This produces piecewise uniform stress and strain fields within each subcell as a result of which the relations between the macroscopic normal (inplane shear) strains, represented by the strain concentration matrix, are related to the average subcell normal (inplane shear) strains only. As demonstrated next, this problem has been circumvented in the present theory, and in fact was the motivating factor that led to the theory's development.

To demonstrate the proposed theory's accuracy in estimating the local subcell stress fields, we consider the classical problem of an isotropic circular fiber embedded in an infinite isotropic matrix subjected to the uniform far-field stress  $\sigma_{22}^\infty$  (the so-called Eshelby problem). The exact analytical solution to this problem is given in compact form in terms of two sets of complex potentials  $\phi$  and  $\psi$ , with each pair corresponding to the fiber and matrix phases. The stresses in the  $x_2$ - $x_3$  plane are obtained from the formulas

$$\begin{aligned}\sigma_{22} &= 2 \operatorname{Re} \phi' + \operatorname{Re}(\bar{z} \phi'' + \psi'), \\ \sigma_{33} &= 2 \operatorname{Re} \phi' - \operatorname{Re}(\bar{z} \phi'' + \psi'), \\ \sigma_{23} &= \operatorname{Im}(\bar{z} \phi'' + \psi')\end{aligned}\quad (67)$$

where the prime denotes the derivative with respect to the complex variable  $z = x + iy$ , and  $\operatorname{Re}$  and  $\operatorname{Im}$  denote the real and imaginary parts of the expressions within the parentheses, respectively. For the fiber phase, the two complex potentials are given by

$$\phi_f = a_1 z, \quad \psi_f = p_1 z \quad (68)$$

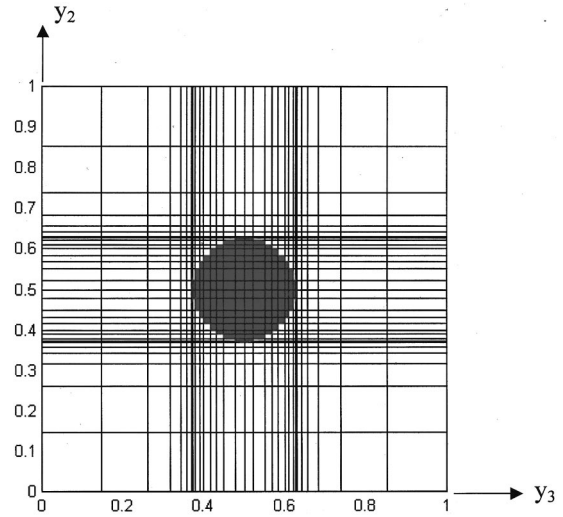
and for the matrix phase they are

$$\phi_m = \bar{a}_{-1} z^{-1} + \bar{a}_1 z, \quad \psi_m = \bar{p}_{-3} z^{-3} + \bar{p}_{-1} z^{-1} + \bar{p}_1 z. \quad (69)$$

Application of the interfacial displacement and traction continuity conditions, and the far-field boundary conditions gives the following expressions for the coefficients appearing in the above equations in the case of plane-strain loading ( $\bar{\epsilon}_{11} = 0$ )

$$a_1 = \frac{1}{4} \sigma_{22}^\infty \frac{(\kappa_m + 1) \mu_f}{2 \mu_f + (\kappa_f - 1) \mu_m}, \quad p_1 = \frac{1}{2} \sigma_{22}^\infty \left[ 1 + \frac{\mu_f - \mu_m}{\mu_m + \kappa_m \mu_f} \right] \quad (70)$$

$$\begin{aligned}\bar{a}_{-1} &= \frac{1}{2} \sigma_{22}^\infty \left[ \frac{\mu_f - \mu_m}{\mu_m + \kappa_m \mu_f} \right], \quad \bar{a}_1 = \frac{1}{4} \sigma_{22}^\infty, \quad \bar{p}_{-3} = \bar{a}_{-1}, \\ \bar{p}_{-1} &= 2a_1 - \frac{1}{2} \sigma_{22}^\infty, \quad \bar{p}_1 = \frac{1}{2} \sigma_{22}^\infty\end{aligned}\quad (71)$$



**Fig. 4 Volume discretization of the repeating unit cell employed in the analysis of a glass/epoxy unidirectional composite with a fiber volume fraction of 0.05**

where the subscripts  $f$  and  $m$  denote fiber and matrix phases, respectively, and the corresponding  $\kappa$ 's are related to the Lamé's constants  $\lambda$  and  $\mu$  as follows:

$$\kappa = \frac{\lambda + 3\mu}{\lambda + \mu}.$$

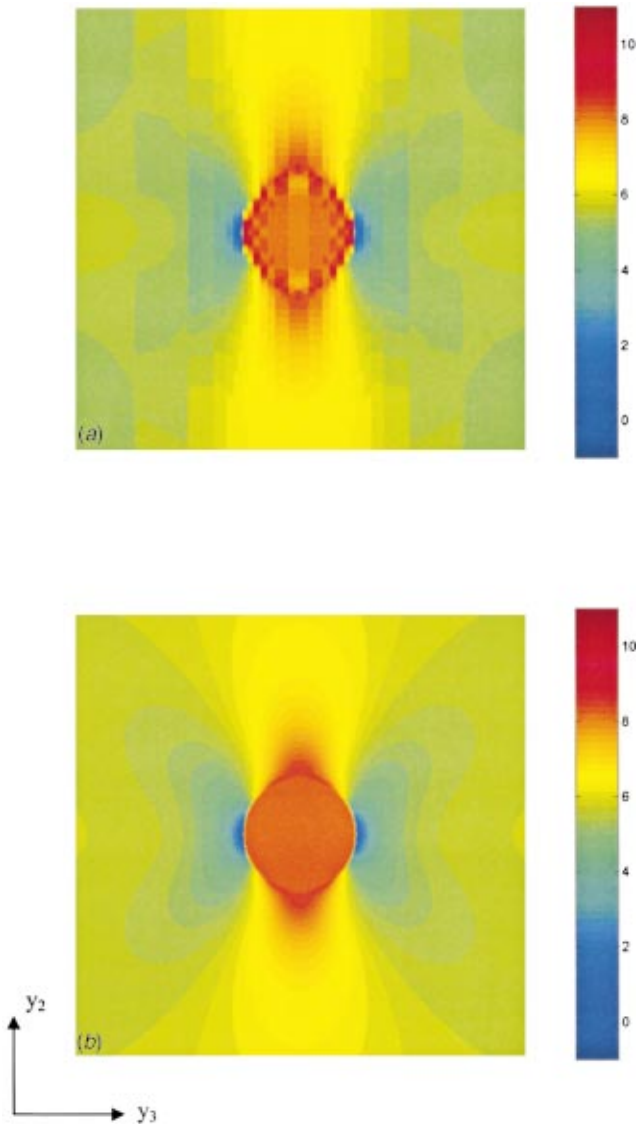
In the fiber, the in-plane normal stresses  $\sigma_{22}$  and  $\sigma_{33}$  are constant while the inplane shear stress  $\sigma_{23}$  vanishes. In the matrix, the stress field is quite complex with all stress components present, but approaches the uniform far-field stress  $\sigma_{22}^\infty$  with increasing distance from the fiber. These features of the exact analytical solution to the considered problem provide a convenient basis for the validation of the present theory.

The discretization of the repeating unit cell for the considered problem is shown in Fig. 4. The cell contains  $30 \times 30$  subcells with the fiber subcells having the same relative dimensions as those in Fig. 3. Thus the approximation of the fiber shape is the same as in the preceding case. The fiber volume fraction for this repeating unit cell is 0.05 which is sufficiently small to be considered dilute, and thus appropriate for comparison with the exact solution given above. The fiber elastic moduli are those of glass included in Table 1, while the matrix is the previously employed epoxy. The choice of these materials produces a high elastic moduli mismatch and thus a significant disturbance in the stress field in the vicinity of the embedded fiber.

The results obtained from the proposed theory have been generated for an applied macroscopic strain  $\bar{\epsilon}_{22}$  of 0.1 percent under the constraint  $\bar{\epsilon}_{11} = 0$ , simulating the plane strain condition employed in the exact analytical solution. The resulting macroscopic stress  $\bar{\sigma}_{22}$  obtained from the present theory ( $\bar{\sigma}_{22} = 5.83$  MPa) was then taken to be the uniform far-field stress  $\sigma_{22}^\infty$  employed in the exact analytical solution. Figure 5 compares  $\sigma_{22}$  stress contours in the repeating unit cell generated by the present theory with the

**Table 4 Comparison of predicted effective elastic moduli of a graphite/epoxy unidirectional composite ( $v_f = 0.60$ )**

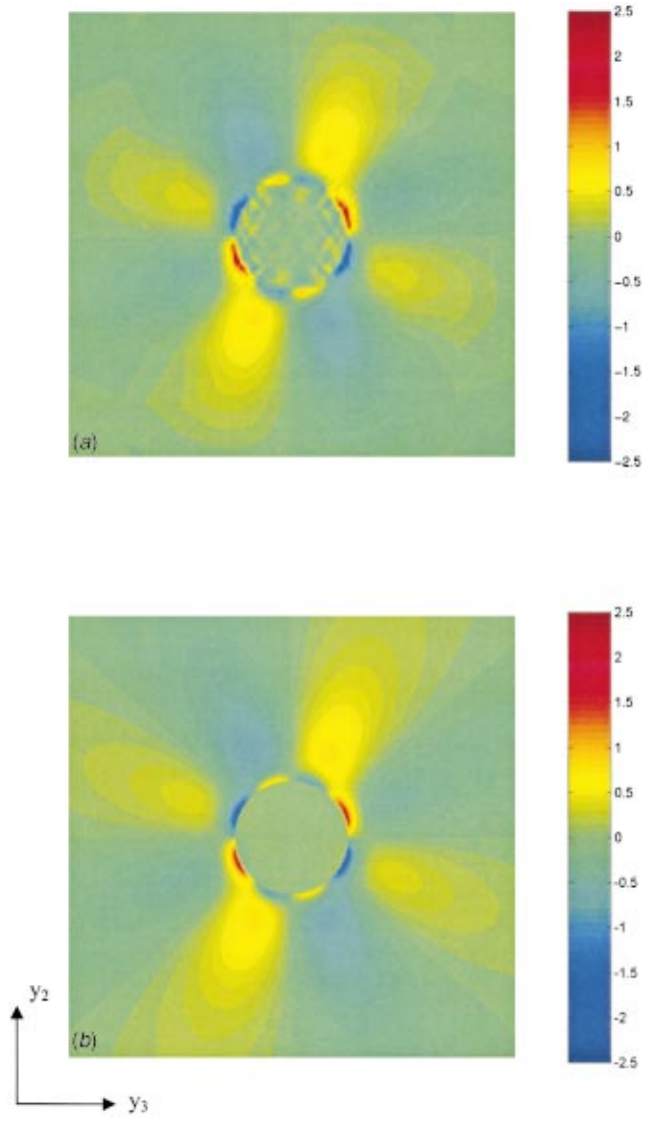
Effective Elastic Moduli	$E_{11}$ (GPa)	$E_{22}$ (GPa)	$G_{12}$ (GPa)	$G_{23}$ (GPa)	$\nu_{12}$	$\nu_{23}$
Present model	142.9	9.61	6.09	3.10	0.252	0.350
Sun and Vaidya [16]	142.6	9.60	6.00	3.10	0.250	0.350
Generalized method of cells [12]	143.0	9.47	5.68	3.03	0.253	0.358



**Fig. 5 Comparison of the  $\sigma_{22}$  stress contours in the repeating unit cell of a glass/epoxy unidirectional composite generated by the present theory (a) and the exact analytical solution (b), (colorbar scale in MPa)**

corresponding contours obtained from Eqs. (67)–(71), calculated in the same region as that occupied by the repeating unit cell. Both qualitative and quantitative agreement is observed between the present model and the exact analytical solution, despite some interaction effects due to the presence of adjacent cells which can be eliminated by further decreasing the fiber volume fraction. In particular, the  $\sigma_{22}$  field predicted by the present model within the fiber is nearly uniform, and the features of the stress field in the matrix phase in the fiber's vicinity are the same as those of the exact analytical solution. Similar results have been obtained for the  $\sigma_{33}$  stress field. In contrast, the  $\sigma_{22}$  stress distribution predicted by the generalized method of cells (not shown) is uniform within each column of subcells along the  $x_2$ -direction, due to the first-order displacement representation within each subcell together with the imposition of traction continuity conditions at the subcell interfaces in an average sense. Clearly, the higher-order representation of the displacement field employed in the present theory accurately captures both the qualitative and quantitative features of the actual stress field.

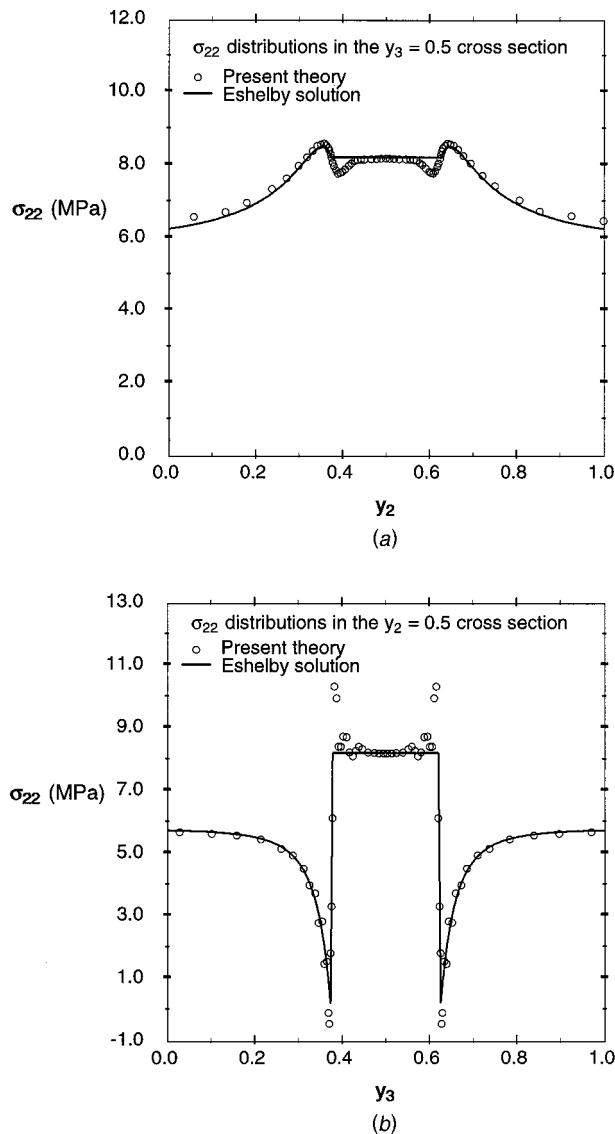
Equally important is the agreement for the  $\sigma_{23}$  stress field



**Fig. 6 Comparison of the  $\sigma_{23}$  stress contours in the repeating unit cell of a glass/epoxy unidirectional composite generated by the present theory (a) and the exact analytical solution (b), (colorbar scale in MPa)**

shown in Fig. 6. In particular, the in-plane shear stress predicted by the present theory is nearly zero within the fiber, and exhibits the same qualitative and quantitative characteristics in the matrix region as that of the exact analytical solution, including the anti-symmetric character about the unit cell's axes of symmetry (along which the shear stress vanishes) and the symmetric character through the unit cell's center. This is in stark contrast with the generalized method of cells which, due to the absence of coupling between the normal and shear effects, predicts zero shear stress everywhere within the repeating unit cell for the given loading. This, too, is a direct result of the first-order displacement representation in the individual subcells employed by the generalized method of cells.

Finally, Fig. 7 compares the  $\sigma_{22}$  distributions predicted by the proposed theory and the exact analytical solution along the cross sections that define the unit cell's material axes of symmetry, i.e., along the lines  $y_2=0.5$  and  $y_3=0.5$ . The quantitative agreement between the two approaches is clearly evident, as is the continuity of the normal traction component at the fiber/matrix interface in the  $y_3=0.5$  cross section (coincident with the loading direction). The sharp  $\sigma_{22}$  discontinuity at the fiber/matrix interface in the



**Fig. 7 Comparison of the  $\sigma_{22}$  stress distributions in the  $y_3 = 0.5$  (a) and  $y_2 = 0.5$  (b) cross section of the repeating unit cell of a glass/epoxy unidirectional composite generated by the present theory and the exact analytical solution**

$y_2 = 0.5$  cross section and the high stress gradient are also very well captured by the proposed theory, as is the asymptotic behavior towards  $\sigma_{22}^\infty$  with increasing  $y_3$ . The overshoot at the fiber/matrix interface within the fiber itself is most likely caused by the rectangular approximation of the fiber shape in this region (see Fig. 4) which, in principle, can be further refined.

All of the above results were generated on a SUN ULTRA 10 work station. The CPU time for the two repeating unit cells containing  $26 \times 26$  subcells, used for comparison with the finite element results, was 80 seconds. The analysis of the repeating unit cell containing  $30 \times 30$  subcells, used for comparison with the exact analytical solution, required 278 CPU seconds due to the corresponding increase in the size of the structural stiffness matrix  $\mathbf{K}$  in Eq. (57). The issue of the local stress field accuracy as a function of the repeating unit cell discretization, as well as the computational speed enhancement for repeating unit cells containing large numbers of subcells, will be addressed in separate communications. These issues can be efficiently investigated owing to

the ease of the repeating unit cell's construction, which also facilitates incorporation of the proposed micromechanics approach into structural analysis computer codes.

#### 4 Summary and Conclusions

A new theory for periodic multiphase materials has been presented and demonstrated to be capable of estimating with sufficient accuracy both the effective moduli and local stress fields in unidirectional composites with continuous reinforcement. The outlined theoretical framework combines elements of the homogenization technique, which provides a basis for consistent approximation of the displacement field at the local level, with the higher-order theory for functionally graded materials developed previously by the authors. The higher-order displacement field approximation at the local level employed in the present approach provides the necessary coupling between the local normal and inplane shear stress fields and the macroscopically applied loading. This coupling dramatically improves the accuracy of estimating the local stress fields relative to the generalized method of cells which is based on a first-order displacement approximation at the local level.

Closed-form expressions for the effective moduli were provided that are valid for microstructures characterized by repeating unit cells with arbitrary reinforcement distributions under multiaxial macroscopic loading due to the use of periodic boundary conditions that follow from the homogenization approach's framework. Thus all the effective moduli can be generated irrespective of whether or not a repeating unit cell possesses planes of material symmetry. The repeating unit cell's construction is simple due to the employed volume discretization that produces a rectangular grid whose subcells are appropriately assigned different material properties and dimensions so as to mimic a multiphase material's actual microstructure. Further, the computational speed with which effective elastic moduli and local stress fields are generated is sufficiently fast for reasonably detailed volume discretizations of a repeating unit cell. These features of the presented theory facilitate investigations of the impact of different materials architectures on both the macroscopic and local responses in an efficient and accurate manner. They also make it straightforward to incorporate the theory into a structural analysis computer code as a subroutine.

#### References

- [1] Hill, R., 1963, "Elastic Properties of Reinforced Solids: Some Theoretical Principles," *J. Mech. Phys. Solids*, **11**, pp. 357–372.
- [2] Walker, K. P., Freed, A. D., and Jordan, E. H., 1991, "Microstress Analysis of Periodic Composites," *Composites Eng.*, **1**, pp. 29–40.
- [3] Christensen, R. M., 1979, *Mechanics of Composite Materials*, John Wiley and Sons, New York.
- [4] Aboudi, J., 1991, *Mechanics of Composite Materials: A Unified Micromechanical Approach*, Elsevier, Amsterdam.
- [5] Hollister, S. J., and Kikuchi, N., 1992, "A Comparison of Homogenization and Standard Mechanics Analyses for Periodic Porous Composites," *Computational Mech.*, **10**, pp. 73–95.
- [6] Nemat-Nasser, S., and Horii, M., 1993, *Micromechanics: Overall Properties of Heterogeneous Materials*, North-Holland, New York.
- [7] Parton, V. Z., and Kudryavtsev, B. A., 1993, *Engineering Mechanics of Composite Structures*, CRC Press, Boca Raton, FL.
- [8] Arnold, S. M., Pindera, M.-J., and Wilt, T. E., 1996, "Influence of Fiber Architecture on the Inelastic Response of Metal Matrix Composites," *Int. J. Plast.*, **12**, No. 4, pp. 507–545.
- [9] Kalamkarov, A. L., and Kolpakov, A. G., 1997, *Analysis, Design and Optimization of Composite Structures*, John Wiley and Sons, New York.
- [10] Banks-Sills, L., Leiderman, V., and Fang, D., 1997, "On the Effect of Particle Shape and Orientation on Elastic Properties of Metal Matrix Composites," *Composites, Part B*, **28**, No. 4, pp. 465–481.
- [11] Aboudi, J., Pindera, M.-J., and Arnold, S. M., 1999, "Higher-Order Theory for Functionally Graded Materials," *Composites, Part B*, **30**, No. 8, pp. 777–832.
- [12] Paley, M., and Aboudi, J., 1992, "Micromechanical Analysis of Composites by the Generalized Method of Cells," *Mech. Mater.*, **14**, pp. 127–139.
- [13] Aboudi, J., Pindera, M.-J., and Arnold, S. M., 1996, "Thermoelastic Theory for the Response of Materials Functionally Graded in Two Directions," *Int. J. Solids Struct.*, **33**, No. 7, pp. 931–966.



- [14] Levin, V. M., 1967, "On the Coefficients of Thermal Expansion of Heterogeneous Materials," *Mekh. Tverd. Tela*, **1**, pp. 88, in Russian.
- [15] Schapery, R. A., 1968, "Thermal Expansion Coefficients of Composite Materials Based on Energy Principles," *J. Compos. Mater.*, **2**, pp. 380.
- [16] Sun, C. T., and Vaidya, R. S., 1996, "Prediction of Composite Properties From a Representative Volume Element," *Compos. Sci. Technol.*, **56**, pp. 171–179.
- [17] Tamma, K. K., and Avila, A. F., 1999, "An Integrated Micro/Macro Modeling and Computational Methodology for High Temperature Composites," *Thermal Stresses 5*, R. B. Hetnarski, ed., Lastran Corporation, Rochester, NY, pp. 143–256.
- [18] Aboudi, J., 1996, "Micromechanical Analysis of Composites by the Method of Cells—Update," *Appl. Mech. Rev.*, **49**, No. 10, Part 2 pp. S83–S91.

# Stress Analysis of Layered Elastic Solids With Cracks Using the Fast Fourier Transform and Conjugate Gradient Techniques

I. A. Polonsky

L. M. Keer

Department of Civil Engineering,  
Northwestern University,  
Evanston, IL 60208-3109

*The fast Fourier transform (FFT) technique has recently been applied to stress analyses of layered elastic solids, with a great deal of success. However, the existing FFT-based methods are limited to intact solids. This paper explores the possibility of using FFT for stress analyses of layered elastic solids containing cracks. A new numerical approach is developed by combining three-dimensional FFT with the theory of periodic eigenstrain and the conjugate gradient method. The new method is primarily designed for analyzing complex three-dimensional crack patterns in layered solids, such as those produced in thin protective coatings by roughness-induced contact stresses. The method should be particularly advantageous for studying crack propagation in coatings, as it does not require remeshing when the crack shape changes. Numerical examples illustrating advantages as well as limitations of the method are presented. Some unexpected results that were obtained for multiple cracks in a thin coating are discussed.*

[DOI: 10.1115/1.1381394]

## 1 Introduction

Protective coatings have been widely used in many tribological applications, including cutting tools, engine components, and magnetic data storage, and their importance can only be expected to increase in the near future ([1,2]). To understand the mechanisms underlying coating performance in various applications, it is essential to analyze stresses generated in the coated system under contact load as well as the coating material response to such stresses.

A remarkable breakthrough in the stress analysis of layered elastic solids has been achieved in the last several years. It was brought about by the application of the fast Fourier transform (FFT) technique to contact mechanics (Ju and Farris [3], Polonsky et al. [4], Nogi and Kato [5], and Polonsky and Keer [6]). The FFT-based approach offers two important advantages. First, the elastic response functions of the layered solid are only required in the wave number domain, where closed-form analytical expressions for such functions are available. Conversely, in the classical approaches using basis functions (Chen and Engel [7]) or influence coefficients (Chiu and Hartnett [8], Cole and Sayles [9]), inverse Fourier or Hankel transforms need to be computed by numerical integration, which is a very time-consuming operation. Second, the surface deflections produced by a given contact pressure distribution can be computed in just  $O(N \log N)$  operations by using FFT, where  $N$  is the number of nodes in the surface grid used to solve the problem. When the same operation is performed by the conventional direct convolution method (Kubo et al. [10], Francis [11], and Ren and Lee [12]), its computational cost is  $O(N^2)$ . In the case of rough contact problems, where  $N \sim 10^5$  or even  $N \sim 10^6$  commonly arise (see [13] and references therein), the speed advantage of FFT becomes crucial. Rough contact analyses of layered solids have great practical significance, as one

of the main functions of tribological coatings is substrate protection against roughness-induced stresses. Furthermore, in many situations, the coating thickness is comparable to the spatial range of roughness-induced stress spikes ([4]).

However, even the most accurate prediction of roughness-induced contact stresses in a coated system does not answer all the equations. Even the most uniform coatings contain pre-existing microcracks, and new cracks can be initiated under severe contact conditions. Under repeated contact load, cracks can propagate through coating thickness, which can ultimately lead to coating failure ([14,15]). To understand how coatings behave under severe contact conditions and how their strength and toughness are affected by their microstructure, it is necessary to perform fracture mechanics analyses of layered solids under contact loading. Unfortunately, the highly efficient FFT-based methods discussed above are not directly applicable to such problems. For cracked solids, the elastic response functions are generally unavailable even in the wave number domain. However, knowledge of these response functions is essential for the method application (cf. [5,6]).

There is a different way of applying the FFT technique to elastic solids with imperfections. Moulinec and Suquet [16,17] described composite material as infinite elastic media containing periodic distributions of eigenstrains and used FFT for calculating the resulting stress fields. This is possible because elastic response functions of an infinite homogeneous medium are readily available in both the space and wave number of domains. Hermann et al. [18] applied a similar approach to stress analyses of solids containing periodic systems of inclusions or voids. They also indicated how the same method can be applied to cracked solids. However, the above approach is not directly applicable to contact mechanics problems because contacting solids are essentially non-periodic in the direction normal to their surfaces.

A number of workers have applied the finite element method to modeling contact-induced coating fracture (e.g., Tian and Saka [19], Eberhardt and Kim [20], and Souza et al. [21]). There is no conceptual difficulty for finite element method application to such problems. However, finite element method does require bulk discretization, and its numerical complexity is between  $O(N^2)$  and  $O(N^3)$ , where  $N$  is the number of elements. Consequently, compu-

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. 21, 2000; final revision, Mar. 13, 2001. Associate Editor: D. Kouris. Discussion on the paper should be addressed to the Editor, Prof. 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.

tation times can become exceedingly long in three-dimensional crack problems involving multiple cracks and/or complex crack shapes.

Lin and Keer [22,23] used the boundary element method to analyze a vertical three-dimensional crack lying in a multilayer solid. Kuo and Keer [24] extended the analysis to slant cracks. An important advantage of the boundary element method is that only the crack faces need to be discretized. However, computation of the corresponding influence coefficients, which involves evaluation of inverse Hankel transforms, turned out to be a costly operation even for relatively small numbers of elements. Furthermore, system assembly in the boundary element method is a rather nontrivial task, which is not easily automated. Consequently, computer code modification is often required to consider a different crack geometry.

The present work explores the possibility of applying the FFT technique to contact stress analyses of elastic solids with localized three-dimensional inhomogeneities, such as cracks. It is tantalizing to try to take advantage of the  $O(N \log N)$  complexity of FFT when dealing with typically complex geometries arising in such problems. A more immediate goal of this work is to develop a numerical tool suitable for analyzing complex crack patterns arising in coated systems subjected to roughness-induced contact stresses. To achieve these goals, a three-dimensional FFT algorithm is combined with the eigenstrain theory (similarly to the approach of Moulinec and Suquet) and the conjugate gradient (CG) method. To account for the essentially nonperiodic nature of contacting solids, a special correction term is used, which is constructed using in-plane two-dimensional FFT. The resulting numerical approach is described in this paper, followed by numerical examples. Application of the new approach to studying the effect of coating microstructure on crack propagation in thin coatings under contact fatigue conditions will be presented in a companion paper.

## 2 Numerical Techniques

The present analysis deals with layered elastic solids containing contact-induced three-dimensional cracks. The basic problem geometry is shown schematically in Fig. 1. The layered solid consists of a semi-infinite elastic substrate and an arbitrary number of perfectly bonded elastically dissimilar layers. The topmost layer of the solid can contain an arbitrary number of cracks. Cracks are indicated in Fig. 1 by bold black lines located below the  $x$ -axis. Although Fig. 1 only shows a cross section of such a cracked solid, the cracks in the figure have finite dimensions in all three directions, i.e., they are three-dimensional cracks. Planar crack areas lying in different planes can be joined, forming kinked three-dimensional cracks. The rest of the layered solid is assumed to remain intact. This special case is of a considerable practical significance, since roughness-induced coating fracture is likely to start from the surface. In principle, the method can be extended to cracks located in several layers and/or in the substrate, but such a

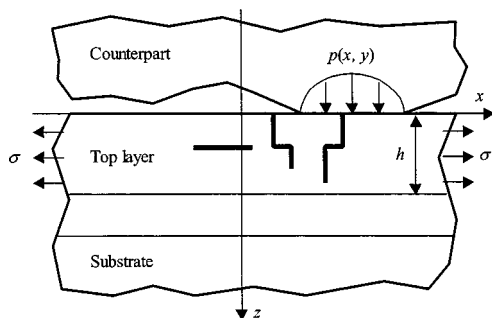


Fig. 1 Layered solid containing contact-induced cracks (shown as thick black lines). Counterpart roughness is exaggerated. The  $y$ -direction is normal to the picture plane.

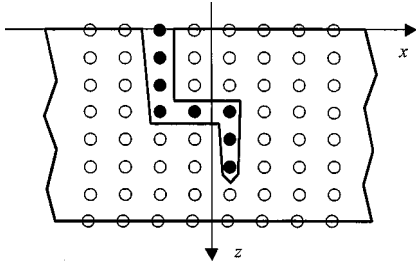
refinement is left to future work. Furthermore, it will be assumed for simplicity that all of the three-dimensional crack segments lie in either vertical or horizontal planes. The cracks are driven by subsurface stresses generated when the cracked solid comes into contact with a counterpart, which is shown in Fig. 1 as a jagged form located above the  $x$ -axis. The counterpart can also be layered, but it contains no cracks. One or both of the solids can have rough surfaces with known topographies. In addition to the contact stresses, uniform in-plane stresses, such as residual stresses, may be present in the cracked layer. The Cartesian coordinate system  $(x, y, z)$  used in the subsequent analysis is also shown in Fig. 1.

It is not obvious how to apply FFT to cracked elastic solids, especially in situations where cracks lie in different planes or have kinks. The basic approach explored in this paper is to replace the cracked layer by an intact elastic medium with an appropriate eigenstrain distribution, and then apply FFT to the latter. "Eigenstrain" is a generic term encompassing various kinds of inelastic strains in solids, such as thermal expansion, phase transformation strains, or plastic strains (Mura [25], Chapter 1). Cracks in elastic solids can also be modeled as eigenstrain distributions ([25], Chapter 5).

To illustrate this point, consider a two-dimensional crack normal to the  $z$ -direction and subjected to a tensile (Mode I) loading. In the loaded state, the crack faces are open, while the normal stress  $\sigma_{zz}$  vanishes along the crack line. The corresponding crack-opening displacement distribution will be denoted by  $\Delta u_z(x)$ . Next, imagine that the crack cavity is filled with the material of the cracked solid, and the continuity of the solid is restored. If the crack is filled completely, but without forcing any material in, the material inside the crack will be stress-free. Hence, the stress field existing in the solid will not be affected by the above procedure. However, exactly the same stress state can be achieved by assuming that the solid has been intact all the time and that the extra material has been produced by stress-free expansion of an infinitesimally thin strip within the solid. Therefore, the crack in its open state can be equivalently replaced by an eigenstrain distribution of the form  $g_{zz}(x) = \delta(z) \Delta u_z(x)$ , where  $\delta(z)$  is Dirac's delay function. The same reasoning can be applied to arbitrary system of cracks in both two-dimensional and three-dimensional cases, including mixed-mode and nonplanar cracks. The main advantage of modeling cracks by eigenstrain distributions is that the cracked solid geometry, which is often complex, is replaced by a much simpler geometry of the corresponding intact solid.

As the crack-opening displacement distribution is not known in advance in crack problems, the equivalent eigenstrain distribution is also initially unknown. It needs to be determined from the condition that the crack faces be free of traction stresses. If Green's functions in closed form are available for the solid in its intact state, the eigenstrain-induced stresses can be represented as an integral over the crack area, thus obtaining a boundary integral equation with respect to the crack-opening displacement. Discretization of this boundary integral equation leads to a boundary element method formulation suitable for three-dimensional crack analyses ([26–28]). However, Green's functions in closed form are unavailable for layered solids. Therefore, a different approach to solving crack-related eigenstrain problems is proposed here. The equivalent eigenstrain distribution is determined by iteration. During each iteration step, the stresses generated by the current approximation to the eigenstrain are computed with the aid of FFT. The stress computation procedure is described next.

To enable the use of FFT, the cracked portion of the topmost layer is discretized by using a uniformly spaced rectangular grid aligned with the layer boundaries (Fig. 2). The tensor components of all elastic fields are stored as discrete arrays of nodal values. In particular, each eigenstrain component  $g_{ij}(x, y, z)$  is represented by a discrete array of nodal values  $g_{ij}(l, m, n)$ . Here indices in parentheses refer to node positions in the three-dimensional grid, while subscript indices denote tensor components in the axes



**Fig. 2 Discretization of the cracked layer. Grid nodes are shown as circles. Nodes carrying eigenstrain are filled.**

shown in Fig. 1. In what follows, the node indices will often be omitted for brevity. As discussed above, we assume that all of the crack segments are either parallel or normal to the surface. In addition, the crack segments will be assumed to lie in planes containing grid nodes. Thus,  $g_{ij}=0$  for all nodes that do not lie within a crack. The above assumptions can be relaxed by using interpolation between grid nodes.

For each nodal array  $g_{ij}$ , the corresponding discrete Fourier transform is computed by applying a three-dimensional FFT algorithm:

$$\hat{g}_{ij} = \text{FFT}(g_{ij}), \quad i, j = x, y, z. \quad (1)$$

Based on the obtained discrete transformations  $\hat{g}_{ij}$ , the following spectral representation of the eigenstrain is considered:

$$\tilde{g}_{jk}(\mathbf{x}) = \sum_{l=0}^{M_x-1} \sum_{m=0}^{M_y-1} \sum_{n=0}^{M_z-1} \hat{g}_{jk}(l, m, n) \exp(i\xi(l, m, n) \cdot \mathbf{x}), \quad j, k = x, y, z. \quad (2)$$

Here  $\mathbf{x}$  is the coordinate vector,  $M_x$ ,  $M_y$ , and  $M_z$  are the numbers of grid nodes in the three directions,  $i$  is the imaginary unity, and  $\xi(l, m, n) = (\xi_x(l), \xi_y(m), \xi_z(n))$  is the wave vector. The components of  $\xi$  (i.e., the wave numbers) are given by

$$\xi_i(k) = 2\pi k' / (a_i M_i), \quad i = x, y, z,$$

where  $a_i$  is the grid spacing in the direction  $i$ ,  $k' = k$  for  $0 \leq M_i/2$ , and  $k' = M_i - k$  for  $M_i/2 < k < M_i$ .

Unlike the original eigenstrain distribution  $g_{ij}(\mathbf{x})$ , which vanishes outside the grid volume,  $\tilde{g}_{ij}(\mathbf{x})$  is infinite and periodic in all directions. Furthermore, while  $g_{ij}(\mathbf{x})$  is localized in the crack planes,  $\tilde{g}_{ij}(\mathbf{x})$  is a continuous oscillating function distributed over the grid volume. However,  $\tilde{g}_{ij}(\mathbf{x})$  vanishes at all grid nodes lying outside the cracks, has a relatively low amplitude between such nodes, and has relatively high and narrow peaks corresponding to the crack areas. Thus,  $\tilde{g}_{ij}(\mathbf{x})$  does approximate  $g_{ij}(\mathbf{x})$  within the grid volume. It can be said that in the present model, a crack has a small, but finite thickness (on the order of the grid spacing). When the grid spacings are small in comparison to the crack dimensions, the approximation of  $g_{ij}(\mathbf{x})$  by  $\tilde{g}_{ij}(\mathbf{x})$  will be accurate.

Next, imagine that the periodic eigenstrain  $\tilde{g}_{ij}(\mathbf{x})$  lies in an infinite elastic medium having the same elastic moduli as the cracked layer. Since the problem is linear, the terms in the right-hand side of Eq. (2) can be analyzed independently. The stresses and displacements corresponding to each spectral component  $\hat{g}_{jk} \exp(i\mathbf{x} \cdot \xi(l, m, n))$  will also be periodic, with the wave vector  $\xi$ . Their amplitudes will be given by the following expressions (Mura [29], and Mura [25], Chapter 1):

$$\hat{u}_j = -iK_{jk}^{-1} \hat{s}_{kl} \xi_l \quad (3a)$$

$$\hat{\sigma}_{ij} = C_{ijkl} e_{kl} = C_{ijkl} K_{km}^{-1} \hat{s}_{mn} \xi_l \xi_n - \hat{s}_{ij}, \quad (3b)$$

where  $\hat{s}_{ij} = C_{ijkl} \hat{g}_{kl}$ ,  $K_{ik} = C_{ijkl} \xi_l \xi_j$ ,  $e_{ij}$  are the elastic strains, and  $C_{ijkl}$  are the elastic constants (the stiffness matrix). The above formulas apply to any  $\xi \neq 0$ . For  $\xi = 0$ , it can be shown that  $\hat{u}_k$

$= 0$  and  $\hat{\sigma}_{ij} = -\hat{s}_{ij}$ . Using Eq. (3), the displacement and stress amplitudes are calculated for all terms in Eq. (2), thus obtaining the discrete Fourier transforms of these fields. Discrete nodal arrays corresponding to the periodic displacement and stress fields generated by the eigenstrain  $\tilde{g}_{ij}(\mathbf{x})$  are then obtained by applying the inverse three-dimensional FFT:

$$\tilde{u}_i = \text{FFT}^{-1}(\hat{u}_i), \quad i = x, y, z; \quad (4a)$$

$$\tilde{\sigma}_{ij} = \text{FFT}^{-1}(\hat{\sigma}_{ij}), \quad i, j = x, y, z. \quad (4b)$$

Algorithms for inverse FFT are very similar to, and as fast as those for direct FFT.

Generally, the eigenstrain-induced elastic fields obtained from Eq. (4) will satisfy neither the free surface conditions at  $z=0$ , nor the continuity conditions at  $z=h$ , where  $h$  is the layer thickness. To satisfy these conditions, and thus obtain a complete solution for the layer with eigenstrain, two additional solutions are superimposed on the periodic solution  $(\tilde{u}_i, \tilde{\sigma}_{ij})$ . One of them corresponds to a surface traction distribution  $\mathbf{t}^+$  acting on the surface of the elastic half-space  $z > 0$ , and the other to a traction distribution  $\mathbf{t}^-$  acting on the surface of the elastic half-space  $z < h$ . Both of these half-spaces are homogeneous and have the same elastic moduli as the layer. Note that the two additional solutions are not associated with any eigenstrain; they are conventional elasticity solutions produced by surface loads. The tractions  $\mathbf{t}^+$  and  $\mathbf{t}^-$  are initially unknown. They are determined with the aid of two-dimensional FFT applied in the  $x$  and  $y$ -directions. First, the discrete solutions  $(\tilde{u}_i, \tilde{\sigma}_{ij})$  is evaluated on the layer boundaries, thus obtaining the following two-dimensional nodal arrays:  $\tilde{t}_i^0 \equiv \tilde{\sigma}_{iz}(l, m, 0)$ ,  $\tilde{t}_i^h \equiv \tilde{\sigma}_{iz}(l, m, M_z - 1)$ , and  $\tilde{u}_i^h \equiv \tilde{u}_i(l, m, M_z - 1)$ . Their two-dimensional transforms, denoted by  $\hat{t}_i^0$ ,  $\hat{t}_i^h$ , and  $\hat{u}_i^h$ , respectively, are then obtained by using two-dimensional FFT. As the problem is linear, the spectral components can be analyzed independently. Using the matrix notation, the boundary conditions for each pair of wave numbers  $(\xi_x, \xi_y)$  can be expressed as follows:

$$\hat{\mathbf{t}}^+ + T^{0h} \hat{\mathbf{t}}^- + \hat{\mathbf{t}}^0 = 0; \quad (5a)$$

$$U^{h0} \hat{\mathbf{t}}^+ + U^{hh} \hat{\mathbf{t}}^- + \hat{\mathbf{u}}^h = U^{\text{sub}} (T^{h0} \hat{\mathbf{t}}^+ + \hat{\mathbf{t}}^- + \hat{\mathbf{t}}^h). \quad (5b)$$

Here the symbol  $T$  denotes traction responses and  $U$  denotes displacement responses. The matrices  $T^{0h}$  and  $U^{hh}$  pertain to the elastic half-space  $z < h$ , the matrices  $T^{h0}$  and  $U^{h0}$  to the half-space  $z < 0$ , and  $U^{\text{sub}}$  to the multilayered substrate to which the top layer is bonded. Note that all these matrixes are functions of the wave numbers  $(\xi_x, \xi_y)$ . Solving the system of matrix Eq. (5) for the vectors  $\hat{\mathbf{t}}^+$  and  $\hat{\mathbf{t}}^-$  yields

$$\hat{\mathbf{t}}^+ = -T^{0h} \hat{\mathbf{t}}^- - \hat{\mathbf{t}}^0, \quad (6a)$$

$$\hat{\mathbf{t}}^- = D^{-1} (U^{\text{sub}} \hat{\mathbf{t}}^h - \hat{\mathbf{u}}^h - B \hat{\mathbf{t}}^0), \quad (6b)$$

where

$$D = U^{hh} - U^{\text{sub}} + B T^{0h}, \quad (7a)$$

$$B = U^{\text{sub}} T^{h0} - U^{h0}. \quad (7b)$$

For the uniform term ( $\xi = 0$ ), one simply obtains  $\hat{\mathbf{t}}^+ = -\hat{\mathbf{t}}^0$  and  $\hat{\mathbf{t}}^- = 0$ . The stress amplitudes generated by the tractions  $\hat{\mathbf{t}}^+$  and  $\hat{\mathbf{t}}^-$  at a given depth  $z$  can be expressed as follows:

$$\hat{\boldsymbol{\sigma}}^+(z) = S^+(z) \hat{\mathbf{t}}^+, \quad \hat{\boldsymbol{\sigma}}^-(z) = S^-(z) \hat{\mathbf{t}}^-. \quad (8)$$

Here  $\hat{\boldsymbol{\sigma}}^+$  and  $\hat{\boldsymbol{\sigma}}^-$  are six-dimensional stress vectors, while  $S^+(z)$  and  $S^-(z)$  are the  $6 \times 3$  subsurface stress response matrices of the half-spaces  $z \geq 0$  and  $z < h$ , respectively.

For an elastically isotropic material, the half-space response matrices,  $U$ ,  $T$ , and  $S$  appearing in Eqs. (5)–(8) are easy to calculate by using periodic Papkovitch-Neuber potentials. They can also be obtained as a special case from the layer responses presented by Chen [30]. The multilayer substrate response  $U^{\text{sub}}$  is computed



by using the well-known propagator matrix technique ([31,32]). The principal recursive relation of this procedure can be expressed as follows:

$$U^{(n)} = U^{00} + (U^{0h} - U^{00}T^{0h})D^{-1}B, \quad (9)$$

where the matrices  $D$  and  $B$  are given by Eq. (7) with  $U^{\text{sub}} = U^{(n-1)}$ .

The additional stress fields generated by the traction distribution  $\mathbf{t}^+$  and  $\mathbf{t}^-$  are computed successively for all horizontal sections of the three-dimensional grid. For section  $n$ , the stress amplitudes corresponding to all pairs of in-plane wave numbers  $(\xi_x, \xi_y)$  are computed from Eq. (8) with  $z = na_z$ , where  $a_z$  is the grid spacing in the  $z$ -direction (see Fig. 2). The inverse two-dimensional FFT is then applied to the resulting stress transforms. This operation yields the two-dimensional stress arrays  $\sigma_{ij}^{n+}(l, m)$  and  $\sigma_{ij}^{n-}(l, m)$  corresponding to the tractions  $\mathbf{t}^+$  and  $\mathbf{t}^-$ , respectively. The total eigenstrain-induced stress  $\sigma_{ij}^e$  in the grid section  $n$  is obtained by adding these two-dimensional arrays to the periodic stress  $\tilde{\sigma}_{ij}$ :

$$\sigma_{ij}^e(l, m, n) = \tilde{\sigma}_{ij}(l, m, n) + \sigma_{ij}^{n+}(l, m) + \sigma_{ij}^{n-}(l, m), \quad 0 \leq l < M_x, \quad 0 \leq m < M_y. \quad (10)$$

The above procedure is performed successively for all grid sections ( $0 \leq n < M_z$ ). The resulting stress distribution  $\sigma_{ij}^e$  corresponds to the given eigenstrain  $g_{ij}$  (to the discretization error). At the same time, it satisfies both the free surface conditions at  $z = 0$  and the interface continuity conditions at  $z = h$ .

The stress field  $\sigma_{ij}^e$  is still periodic in the  $x$  and  $y$ -directions. In principle, additional terms could be used to reduce the associated periodicity error. However, such a refinement is left for future studies. Note that the present method is primarily intended for analyses of roughness-induced cracks in protective coatings. In such applications, multiple cracks randomly distributed over a large area are likely to arise. Hence, assuming in-plane periodicity of the problem is about as good an approximation as considering isolated crack systems.

To obtain the total stress field in the layer, the contact-induced stresses need to be computed. To a first approximation, the interaction between the cracks and the contact can be neglected. This leads to a conventional rough contact problem for a pair of intact layered solids with given surface topographies, which is solved by using the FFT and CG techniques. See [6,13] for algorithm details. In light of the above discussion, the periodicity correction procedure of Polonsky and Keer [6] need not be applied here. The surface deflection responses of layered solids, which are required for contact problem solution, are calculated with the aid of Eq. (9). Having computed the contact pressure distribution  $p(x, y)$ , and assuming proportional traction, the contact traction components  $t_i^c$  are obtained for each surface node as

$$t_x^c = -fp, \quad t_y^c = 0, \quad t_z^c = -p, \quad (11)$$

where  $f$  is the traction coefficient,  $p$  is the nodal pressure, and the frictional force is assumed to act in the negative  $x$ -direction. The subsurface stress field generated by these contact tractions is then computed using Eqs. (6)–(9) with  $\hat{t}_i^0 = -\hat{t}_i^c$ ,  $\hat{t}_i^h = 0$ , and  $\hat{u}_i^h = 0$ . The resulting contact stress field will be denoted by  $\sigma_{ij}^c$ . Finally, the total stress in the cracked layer is obtained as

$$\sigma_{ij} = \sigma_{ij}^e + \sigma_{ij}^c + \sigma_{ij}^u, \quad (12)$$

where  $\sigma_{ij}^u$  is the uniform stress field existing in the layer (such as a residual stress field).

To solve the crack problem, it is necessary to find such an eigenstrain distribution  $g_{ij}$  that the corresponding total stress  $\sigma_{ij}$  satisfies certain conditions on the crack faces. To obtain a concise and efficient discrete formulation for the crack problem, a status array describing the crack geometry is constructed. Each node in the grid is assigned three numbers:  $c_x$ ,  $c_y$ , and  $c_z$ . By definition,

$c_x = 1$  if the node lies within a crack normal to the  $x$ -direction, and  $c_x = 0$  otherwise. The numbers  $c_y$  and  $c_z$  are defined similarly. Then, the problem can be formulated as follows:

$$\sigma_{ii} = 0 \quad (c_i = 1, g_{ii} > 0, \text{ no summation}); \quad (13a)$$

$$\sigma_{ii} \leq 0 \quad (c_i = 1, g_{ii} = 0, \text{ no summation}); \quad (13b)$$

$$\sigma_{ij} = 0 \quad (c_i = 1, i \neq j); \quad (13c)$$

$$g_{ii} \geq 0 \quad (\text{no summation}); \quad (13d)$$

$$g_{ij} = 0 \quad (c_i = c_j = 0). \quad (13e)$$

The above equations and inequalities apply to every node in the grid. Equations (13a) and (13c) require that the three traction stresses vanish at the faces of an open crack. Equation (13b) means that a closed crack can support a compressive stress acting across its faces, but no tensile stress is allowed. Equation (13d) stipulates that crack faces cannot overlap. Thus, the present model is designed to treat the possibility of contact between crack faces in a consistent way. This feature is essential for realistic modeling of contact-induced cracks. On the other hand, Eq. (13c) implies that there is no friction between the faces of a closed crack. Although this assumption is only realistic for surface cracks in lubricated contact situations, it greatly simplifies the analysis. This assumption may be relaxed in the future versions of the method. Finally, Eq. (13e) implies that cracks are the only source of eigenstrain in the problem.

The system of equations and inequalities (13) is solved by iteration, until Eq. (13a) and (13c) are satisfied to the desired accuracy. Equations (13b), (13d), and (13e) are enforced at all times. The iteration scheme used here is based on the CG method and is very similar to the one used in our rough contact solver ([13]). Since a very detailed description of the scheme was included in ([13]), the details will be omitted here. The main differences between the two schemes are that in the present case, iteration is performed with respect to the eigenstrain instead of the contact pressure, the solution residual is based on the crack face tractions instead of the contact gap, and the load balancing procedure is not required. During each iteration step, the total stress  $\sigma_{ij}$  is recomputed for the current approximation to  $g_{ij}$  by using Eqs. (1)–(12). In the end of the solution procedure, the global stress field in the cracked layer is obtained, together with the eigenstrain distribution. The crack-opening displacement distributions for cracks in the layer can be approximated from the calculated eigenstrains as follows:

$$\Delta u_i = g_{ij}a_j \quad (i = x, y, z; \quad c_j = 1; \text{ no summation}). \quad (14)$$

Here  $a_j$  is the grid spacing in the direction  $j$ . The above formula is valid for all grid nodes except those at which two or more cracks intersect.

Numerical experimentation with the new computational method indicated that the convergence rate of the iteration procedure strongly depends on the problem geometry, particularly on the location of cracks with respect to the surface and the interface. For internal cracks, convergence is normally very rapid. On the other hand, surface-breaking cracks presented considerable challenge. The problem stems from the rapid variation of eigenstrains across the crack plane, and from the fact that the eigenstrain-induced stresses and the stress fields added to satisfy the boundary conditions are constructed from different fundamental solutions. However, this difficulty has been overcome by introducing long, narrow fictitious cracks lying in the free surface and running along the mouths of the surface-breaking cracks. The fictitious crack only needs to be three grid spacings wide to achieve acceptable convergence rates. Since the fictitious crack is so narrow, it does not significantly affect the contact-induced stress field. The solution accuracy in the near vicinity of the crack mouth can be further improved by extending both the grid and the surface-breaking crack beyond the surface. This modification, however, is not nearly as important as the previous one. Since rapid convergence

cannot be taken for granted in the present model, the iteration scheme was modified to restart the conjugate gradient algorithm after every ten iteration steps. This measure ensures that the conjugate direction does not become corrupted by numerical error even in the cases with slow convergence.

### 3 Results and Discussion

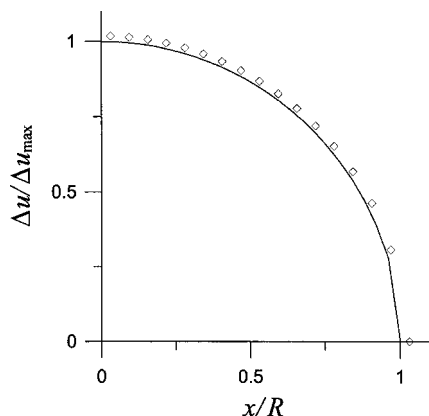
The numerical algorithms described above were implemented in a computer code using the C++ programming language and the Standard Template Library (STL). The FFT implementation was based on a function borrowed from Takuya Ooura's FFT package, which is freely distributed over the Internet. The numerical examples described below were run on a personal computer with a 450 MHz CPU and 256 MB RAM.

The new numerical method was first applied to subsurface stress computation for a pair of crack-free layered solids having rough surfaces and engaged in contact with friction. The results were compared to the output of a contact stress analysis program previously developed by the authors ([6]). The stress distributions computed by the two computer codes were identical to the round-off error. Thus, the procedure used to enforce the boundary condition (Eqs. (6)–(10)) was verified.

To check the basic eigenstrain analysis (Eq. (1)–(4)) and the iteration scheme, the new method was applied to a penny-shaped crack lying in an infinite elastic medium. This geometry was obtained by considering a counterpart perfectly bonded to the surface, and setting the elastic moduli of the layer, the substrate, and the counterpart to the same values. The crack was loaded by a uniform tensile stress. The problem was solved on a  $128 \times 128 \times 32$  grid with equal spacings  $a$  in the three directions. The crack was normal to the  $z$ -direction and was centered with respect to the grid. The crack radius was set to  $R = 16a$ , i.e. the grid length and width were four times greater than the crack diameter. Such an extended grid was used to reduce the periodicity error. The relative accuracy goal for the iteration scheme was set to  $2.5 \times 10^{-4}$ .

The iteration process converged quite rapidly: Only 15 iteration steps were required to compute the 2436 unknown eigenstrains arising in this example to the above accuracy level. The corresponding CPU time was 0.8 hrs, which is a remarkably short time for such a large-scale numerical model ( $N \approx 5 \times 10^5$ , about  $3 \times 10^6$  nodal stress values). Note that the present method calculates the global stress field in the cracked layer. That is, all  $3 \times 10^6$  nodal stresses were computed during each iteration step. The use of three-dimensional FFT in the present algorithm was essential for achieving such a remarkably high computation speed.

The obtained distribution of the Mode I crack-opening displacement  $\Delta u_z$  along a crack radius is shown in Fig. 3 together with the



**Fig. 3 Crack-opening displacement distribution along a crack radius for a penny-shaped crack: numerical solution (diamonds) and analytical solution (solid line)**

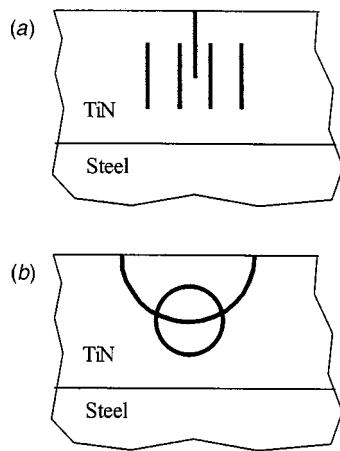
exact analytical solution. It is seen from Fig. 3 that the global crack-opening displacement behavior was reproduced reasonably accurately, despite the relatively coarse discretization ( $R/a = 16$ ). The average relative error of the numerical solution was about three percent. Such level of accuracy should suffice for most practical purposes. The main sources of numerical error in the present method are the discretization error and the periodicity error, both of which stem from the use of finite Fourier series (see Eq. (2)). The periodicity error will only arise for nonperiodic crack geometries, such as the penny-shaped crack. As discussed above, roughness-induced cracking in coatings can be modeled by periodic systems of cracks. The discretization error, however, can only be reduced by using finer grids, which would require more powerful computers. Thus, the present method is best suited for situations where a highly accurate analysis is not justified (e.g., due to uncertainty in the input parameters), but even an approximate solution is difficult to obtain because of the complex problem geometry. Roughness-induced cracking in tribological coatings certainly appears to fall into this category.

One of the main reasons for the numbers of nodes being so large in the present method is that the application of FFT necessitates the use of uniformly spaced grids. Hence, grid nodes cannot be condensed along the crack front, as is commonly done in finite element method models. On the other hand, the simplicity of a uniform rectangular grid is a major advantage, as it makes problem discretization a trivial step. This can be contrasted to the finite element method, where mesh generation is one of the most difficult steps in the analysis and typically requires the use of sophisticated software. The present method is especially advantageous for modeling propagating cracks, since successive crack geometries are analyzed on the same grid and no remeshing is required. Computation results for three-dimensional crack propagation in coatings with various microstructures will be reported in a companion paper.

It is also seen from Fig. 3 that the numerical solution significantly deviates from the exact one at the nodes lying next to the crack front. This is to be expected, since the crack front nodes are not treated in any special way in the present method. Thus, to extract a stress intensity factor from the near-tip elastic fields, the domain integral method (Moran and Shih [33]) would have to be used. Such a procedure has not yet been implemented. However, computation of the absolute value of the stress intensity factor is not very useful in the case of protective coatings, and more insight may be gained from comparative crack analyses. Such analyses can be used to shed light on the effect of various system parameters, including the coating microstructure, on the coating toughness and fatigue resistance. The present method is suitable for such analyses because the near-tip fields, although deviating from the exact solution, exhibit correct scaling with the global stress field and hence with the load and the crack geometry.

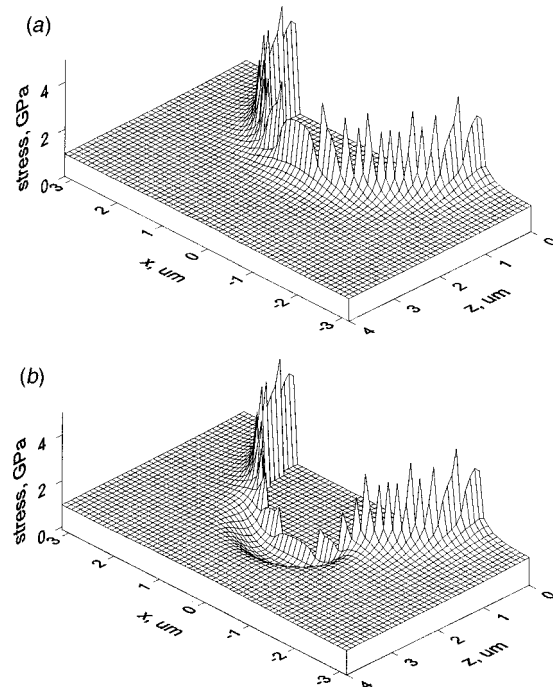
To illustrate this point, the method was applied to the following pair of two-dimensional problems: a Griffith crack lying in an infinite medium, and an edge crack normal to the free surface of a homogeneous half-plane. The Griffith crack was twice as long as the edge crack. The cracks were subjected to a uniform tensile stress of the same magnitude. The stress intensity factor ratio for these two cases is known to be 1.12, with accuracy better than one percent ([34]). Our computations yielded the following results: 1.115 for the ratio of the crack-opening displacement values just inside the crack, and 1.107 for the ratio of the stress values just outside the crack. The corresponding relative error values are about 0.5 percent and about 1 percent, respectively. Thus, the near-tip elastic fields computed by the present method can be used with confidence for comparative analyses of cracks.

Next, the new computational tool was used to study the shielding effect of multiple cracks distributed around the front of a larger crack in a thin protective coating. The problem geometry is shown in Fig. 4. The analyzed solid consisted of a thin coating layer ( $h = 4 \mu\text{m}$ ) on a homogeneous elastic substrate. The elastic

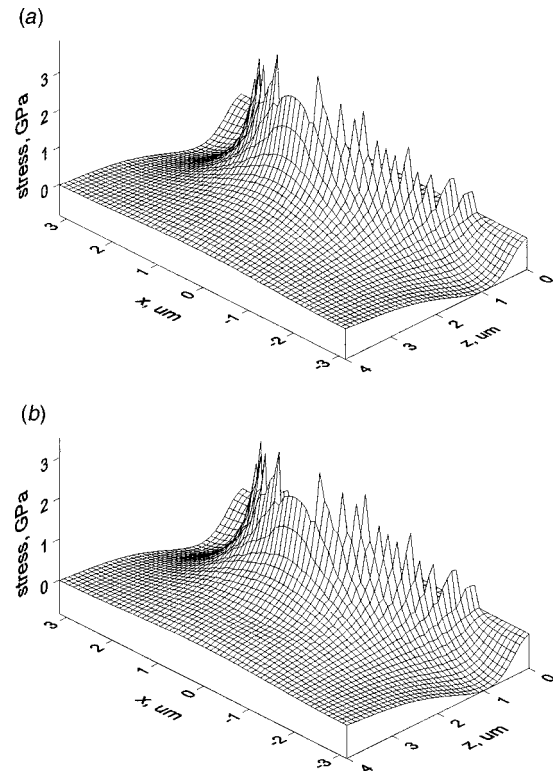


**Fig. 4 Multiple cracks in a thin coating: (x,z) view (a) and (y,z) view (b)**

moduli were chosen as follows:  $E=450$  GPa,  $\nu=0.2$  for the layer, and  $E=210$  GPa,  $\nu=0.28$  for the substrate, where  $E$  is Young's modulus and  $\nu$  is Poisson's ratio. This set of elastic moduli approximately corresponds to the TiN/steel system. The main crack in the coating was a semi-circular edge crack normal to the surface and having a radius of  $2\text{ }\mu\text{m}$ . On each side of the main crack tip, a pair of secondary cracks was located (Fig. 4). The four secondary cracks were circular in shape with  $R=1\text{ }\mu\text{m}$ , and were parallel to the main crack plane. The spacing between them was  $1\text{ }\mu\text{m}$ . Two types of loading were considered: Mode I loading by uniform tensile stress, and asperity contact loading. In the latter case, contact stresses were generated by a single asperity contact located on one side of the main crack. The contact area was a  $2.7\text{-}\mu\text{m}$ -diameter circle tangential to the crack mouth. The asperity contact pressure was almost uniform over the contact area due to plastic yielding of the counterpart. The yield pressure was set to



**Fig. 5 Distribution of the normal stress  $\sigma_{xx}$  in the main crack plane for tensile loading; single crack case (a) and multiple crack cases (b)**



**Fig. 6 Distribution of the shear stress  $\sigma_{xz}$  in the main crack plane for contact loading: single crack case (a) and multiple crack case (b)**

7 GPa and contact friction was neglected. For each of the two loading types, computations were performed both with and without secondary cracks. Thus, a total of four cases were considered.

All four problems were solved on the same  $64 \times 64 \times 64$  grid with equal spacings in all directions ( $a=0.1\text{ }\mu\text{m}$ ). Only a few lines in the input parameter file had to be modified when switching from one case to the next, which shows the convenience of the new method. The relative accuracy goal was again set to  $2.5 \times 10^{-4}$ . In each case, the accuracy goal was attained after no more than 35 iteration steps, and the maximum CPU time was 0.7 hrs.

The calculated stress distributions in the plane of the main crack are shown in Figs. 5 and 6. For uniform tensile loading (Fig. 5), the component  $\sigma_{xx}$ , which is the normal stress acting across the stress plane, is shown. For contact loading (Fig. 6), the crack faces are closed by the compressive contact stresses, so that the crack is only loaded in shear (Modes II and III). In this case, the component  $\sigma_{xz}$ , which is dominant in the central portion of the crack, is plotted. The stress plots appear jagged because the data are defined on a rectangular grid, while the crack front is semi-circular. Hence, the nodes nearest to the front and carrying the highest stresses lie at different distances from the front. This effect, however, does not affect the present comparative analysis, as the main crack has the same shape in all cases.

For tensile loading, stress concentration along the central portion of the main crack front is suppressed conspicuously when the secondary cracks are present (cf. Figs. 5(a) and 5(b)). For example, the stress acting at the crack front midpoint is reduced by a factor of almost five. These results are consistent with the classical concept of microcrack shielding, which is believed to be an important toughening mechanism in ceramic material ([35]).

In the case of contact loading, however, the shielding effect of secondary cracks is barely noticeable (cf. Figs. 6(a) and 6(b)). The stress at the crack front midpoint is only lowered by about 15



percent in the presence of secondary cracks. The present results indicate that the microcrack shielding effect is not nearly as strong for contact-induced cracks as it is for conventional cracks in structure, which are typically produced by tensile (Mode I) load. This result is rather unexpected. There appears to be no intuitive explanation for such a marked difference in the magnitude of crack tip shielding for the two loading types, as the crack geometry was the same in both cases. The above example demonstrates the potential value of comparative numerical analyses for complex crack patterns and/or nontrivial loading types.

#### 4 Conclusion

Cracks in layered elastic solids can be analyzed numerically by using the theory of periodic eigenstrains in combination with three-dimensional FFT and conjugate gradient techniques. The resulting numerical approach is fast and easy to use. Although the method is sufficiently accurate for most practical purposes, high accuracy is difficult to achieve with this approach. The method is best suited for comparative analyses of complex crack patterns, such as those produced in thin protective coatings by roughness-induced contact stresses. Numerical results obtained with the new method for multiple cracks in a thin coating indicate that for cracks driven by contact stresses, the microcrack shielding effect is much weaker than it is for the conventional Mode I loading.

#### Acknowledgment

The financial support of NSF is gratefully acknowledged.

#### References

- [1] Schulz, H., and Quinto, D. T., 1999, "Technological Development of PVD Hard Coatings for Industry," *Z. Metallkd.*, **90**, pp. 831–836.
- [2] Suchentrunk, R., Fuessler, H. J., Staudigl, G., Jonke, D., and Meyer, M., 1999, "Plasma Surface Engineering—Innovative Processes and Coating Systems for High-Quality Products," *Surf. Coat. Technol.*, **112**, pp. 351–357.
- [3] Ju, Y., and Farris, T. N., 1996, "Spectral Analysis of Two-Dimensional Contact Problems," *ASME J. Tribol.*, **118**, pp. 320–328.
- [4] Polonsky, I. A., Chang, T. P., Keer, L. M., and Sproul, W. D., 1997, "An analysis of the Effect of Hard Coatings on Near-Surface Rolling Contact Fatigue Initiation Induced by Surface Roughness," *Wear*, **208**, pp. 204–219.
- [5] Nogi, T., and Kato, T., 1997, "Influence on a Hard Surface Layer on the Limit of Elastic Contact—Part I: Analysis Using a Real Surface Model," *ASME J. Tribol.*, **119**, pp. 493–500.
- [6] Polonsky, I. A., and Keer, L. M., 2000, "A Fast and Accurate Method for Numerical Analysis of Elastic Layered Contacts," *ASME J. Tribol.*, **122**, pp. 30–35.
- [7] Chen, W. T., and Engel, P. A., 1972, "Impact and Contact Stress Analysis in Multilayer Media," *Int. J. Solids Struct.*, **8**, pp. 1257–1281.
- [8] Chiu, Y. P., and Hartnett, M. J., 1983, "A Numerical Solution for Layered Solid Contact Problems With Applications to Bearings," *ASME J. Lubr. Technol.*, **105**, pp. 585–590.
- [9] Cole, S. J., and Sayles, R. S., 1992, "A Numerical Model for the Contact of Layered Elastic Bodies With Real Rough Surfaces," *ASME J. Tribol.*, **114**, pp. 334–340.
- [10] Kubo, A., Okamoto, T., and Kurokawa, N., 1981, "Contact Stress Between Rollers With Surface Irregularity," *ASME J. Mech. Des.*, **103**, pp. 492–498.
- [11] Francis, H. A., 1983, "The Accuracy of Plane Strain Models for the Elastic Contact of Three-Dimensional Rough Surface," *Wear*, **85**, pp. 239–256.
- [12] Ren, N., and Lee, S. C., 1993, "Contact Simulation of Three-Dimensional Rough Surfaces Using Moving Grid Method," *ASME J. Tribol.*, **115**, pp. 597–601.
- [13] Polonsky, I. A., and Keer, L. M., 1999, "A Numerical Method for Solving Rough Contact Problems Based on the Multi-Level Multi-Summation and Conjugate Gradient Techniques," *Wear*, **231**, pp. 206–219.
- [14] Ziegele, H., Rebholz, C., Voevodin, A. A., Leyland, A., Rohde, S. L., and Matthews, A., 1997, "Studies of the Tribological and Mechanical Properties of Laminated CrC-SiC Coatings Produced by rf and dc Sputtering," *Tribol. Int.*, **30**, pp. 845–856.
- [15] Polonsky, I. A., Chang, T. P., Keer, L. M., and Sproul, W. D., 1998, "A Study of Rolling Contact Fatigue of Bearing Steel Coated With PVD TiN Films: Coating Response to Cyclic Contact Stress and Physical Mechanisms Underlying Coating Effect on the Fatigue Life," *Wear*, **215**, pp. 191–204.
- [16] Moulinec, H., and Suquet, P., 1994, "A Fast Numerical Method for Computing the Linear and Nonlinear Mechanical Properties of Composites," *C. R. Acad. Sci., Ser. II: Mec., Phys., Chim., Sci. Terre Univers*, **318**, No. 2, pp. 1417–1423.
- [17] Moulinec, H., and Suquet, P., 1998, "A Numerical Method for Computing the Overall Response of Nonlinear Composites with Complex Microstructures," *Comput. Methods Appl. Mech. Eng.*, **157**, pp. 69–94.
- [18] Herrmann, K. P., Muller, W. H., and Neumann, S., 1999, "Linear and Elastic-Plastic Fracture Mechanics Revisited by use of Fourier Transforms—Theory and Application," *Comput. Mater. Sci.*, **16**, pp. 186–196.
- [19] Tian, H., and Saka, N. J., 1992, "Finite-Element Analysis of Interface Cracking in Sliding Contacts," *Wear*, **155**, pp. 163–182.
- [20] Eberhardt, A. W., and Kim, B. S., 1998, "Crack Face Friction Effects on Mode II Stress Intensities for a Surface-Cracked Coating in Two-Dimensional Rolling Contact," *Tribol. Trans.*, **41**, pp. 35–42.
- [21] Souza, R. M., Mustoe, G. G. W., and Moore, J. J., 1999, "Finite-Element Modeling of the Stresses and Fracture During the Indentation of Hard Elastic Films on Elastic-Plastic Aluminum Substrates," *Thin Solid Films*, **356**, pp. 303–310.
- [22] Lin, W., and Keer, L. M., 1989, "Analysis of a Vertical Crack in a Multilayered Medium," *ASME J. Appl. Mech.*, **56**, pp. 63–69.
- [23] Lin, W., and Keer, L. M., 1989, "Three-Dimensional Analysis of Cracks in Layered Transversely Isotropic Media," *Proc. R. Soc. London, Ser. A*, **A424**, pp. 307–322.
- [24] Kuo, C. H., and Keer, L. M., 1995, "Three-Dimensional Analysis of Cracking in a Multilayered Composite," *ASME J. Appl. Mech.*, **62**, pp. 273–281.
- [25] Mura, T., 1982, *Micromechanics of Defects in Solids*, Kluwer, Dordrecht.
- [26] Murakami, Y., and Nemat-Nasser, S., 1983, "Growth and Stability of Interacting Surface Flaws of Arbitrary Shape," *Eng. Fract. Mech.*, **17**, pp. 193–210.
- [27] Lee, J. C., Farris, T. M., and Keer, L. M., 1987, "Stress Intensity Factors for Cracks of Arbitrary Shape Near an Interfacial Boundary," *Eng. Fract. Mech.*, **27**, pp. 27–41.
- [28] Hanson, M. T., Lin, W., and Keer, L. M., 1989, "Three-Dimensional Analysis of Cracking Through the Boundary of a Two-Phase Material," *ASME J. Appl. Mech.*, **56**, pp. 850–857.
- [29] Mura, T., 1964, "Periodic Distributions of Dislocations," *Proc. R. Soc. London, Ser. A*, **A280**, pp. 528–544.
- [30] Chen, W. T., 1971, "Computation of Stresses and Displacements in a Layered Elastic Medium," *Int. J. Eng. Sci.*, **9**, pp. 775–800.
- [31] Gilbert, F., and Backus, G., 1966, "Propagator Matrices in Elastic Wave and Vibration Problems," *Geophysics*, **31**, pp. 326–332.
- [32] Pan, E., 1991, "An Exact Solution for Transversely Isotropic, Simply Supported and Layered Rectangular Plates," *J. Elast.*, **25**, pp. 101–116.
- [33] Moran, B., and Shih, C. F., 1987, "Crack Tip and Associated Domain Integrals From Momentum and Energy-Balance," *Eng. Fract. Mech.*, **27**, pp. 615–642.
- [34] Murakami, Y., 1986, *Stress Intensity Factors Handbook*, 1st Ed., Pergamon, Oxford, UK.
- [35] Hutchinson, J. W., 1987, "Crack Tip Shielding by Micro-Cracking in Brittle Solids," *Acta Metall. Mater.*, **35**, pp. 1605–1619.



# On Plastic Collapse of Media With Random Yield Strength

A. P.-D. Ku<sup>1</sup>

R. P. Nordgren

Fellow ASME

Department of Civil Engineering,  
Rice University,  
Houston, TX 77005

*This paper concerns the plastic collapse of an elastic/perfectly plastic medium with randomly variable yield strength under a fixed load. The yield strength is represented by a Gaussian random field of known statistical properties. Using the theorems of limit analysis and the methods of reliability theory, algorithms are developed for the computation of upper and lower bounds on the probability of plastic collapse. By varying the magnitude of the fixed load, bounds on the probability distribution function for the collapse load can be computed. Results are given for uniform pressure applied to a rectangular region of the surface of an elastic/plastic half-space. For the corresponding plane problem, results for the classical Hill and Prandtl failure mechanisms are compared. Three-dimensional results are found to differ significantly from those of the plane problem. Comparison is made with results of a previous approximate method for three-dimensional problems. [DOI: 10.1115/1.1388011]*

## 1 Introduction

For deterministic problems, the theory of limit analysis provides a convenient method for determining upper and lower bounds on the plastic collapse load of a continuous medium or a structural system. For structural systems with randomly variable strengths, extensive research efforts have been devoted to combining reliability theory with plastic limit analysis; see, e.g., Wang et al. [1]. In this class of problems which includes trusses and frames, the yield limits of potential plastic hinges of structural members are treated as random variables. Then the upper and lower-bound theorems of limit analysis allow computation of bounds on reliability for the plastic collapse load using the methods of reliability theory.

A similar approach is possible for problems involving continuous media with yield strength represented as a random field. For example, continuous random fields have been applied in the probabilistic analysis of slope stability in geotechnical engineering by Vanmarcke [2], Li and Lumb [3], Yucemen and Al-Homoud [4], among others. These studies used the limit equilibrium method rather than the theorems of limit analysis to calculate the failure probability of the slope. Following a somewhat different approach, Nordgren [5] applied the kinematic theorem of plastic limit analysis to obtain an upper bound on the mean collapse load for problems of an elastic/perfectly plastic medium with random yield strength. To illustrate his method, Nordgren treated the plane problem of a wedge under pressure over a portion of one surface. In another probabilistic study, Ostoja-Starzewski and Ilies [6] used the slip-line theory of plasticity to obtain the probability density function (p.d.f.) of the plastic collapse load for the plane problem of a cylinder under internal pressure with spatially random yield strength. Both of these last two studies employed Monte Carlo simulation to obtain information on the probability of plastic collapse.

The present investigation provides general methods for analyzing the reliability of an elastic/perfectly plastic medium whose yield strength is regarded as a continuous random field. The proposed methods combine the upper and lower-bound theorems of

plastic limit analysis with the computational methods of reliability theory. For a particular realization of the random field of yield strength, application of each limit theorem involves dividing the spatial region of the problem into simple elements on which the constraints of the theorem are enforced while optimizing the collapse load. The collapse calculation based on the upper-bound theorem involves a nonlinear programming problem. The collapse load calculation based on the lower-bound theorem involves a linear programming problem. The calculated collapse load is used to evaluate the limit state function in the computational approach of modern reliability analysis. Bounds on the probability of failure under a given load can thus be calculated by the numerically efficient first-order reliability method (FORM) and by the Monte Carlo simulation (MCS) method. Repeated calculations of failure probability at different given loads enable bounds on the p.d.f. of the plastic collapse load to be determined.

As an application of the methodology for illustrative purposes, the classical indentation problem of an elastic/perfectly plastic half-space is treated. The material is assumed to admit the von Mises yield function with randomly variable yield strength modeled as a homogeneous Gaussian random field with known mean, variance, and covariance functions. The indentation load is considered to be constant pressure applied over a specified rectangular region of the surface of the half-space. It should be pointed out that the homogeneous Gaussian random field assumption is an approximation to the statistical structure of some engineering materials on the macroscopic level. Considerations of random field based on micromechanical analysis and its associated discretization for numerical analysis have been carried out by Jiang et al. [7] and Ostoja-Starzewski [8].

For the plane problem of the illustration, upper and lower bounds on the p.d.f. for the plastic collapse pressure are computed using the upper- and lower-bound theorems and FORM. The results are verified by MCS using importance sampling for representative cases. Results for the classical Hill and Prandtl failure mechanisms are compared. The p.d.f. for collapse pressure in the Hill mechanism is found to have a lower variance than that of the Prandtl mechanism although the mean failure load is approximately the same in both cases.

Next, the three-dimensional indentation problem is treated in a similar manner using the upper-bound theorem and the probabilistic results are compared with those of the plane indentation problem. In the presence of randomly varying yield strength, the indentation problem is essentially three-dimensional even if the geometry and loading suggests a two-dimensional approximation. Our numerical results show that two-dimensional models are inappropriate for this problem. Also, the present study offers a direct

<sup>1</sup>Now at EQE International, Inc., 16850 Diana Lane, Houston, TX 77058.

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, June 14, 2000; final revision, Feb. 6, 2001. Associate Editor: M.-J. Pindera. Discussion of the paper should be addressed to the 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.

way of calculating the three-dimensional reliability and assessing the validity of a previous approximate method ([2,4]). It is found that the reliability results given by the present study agree fairly well with this approximate method, thus confirming its applicability.

## 2 Theory of Plasticity and Limit Analysis

We begin with a brief review of the theorems and methods of plastic limit analysis for an elastic/perfectly plastic material with *deterministic* inhomogeneous material properties. The theorems are stated for the special case of surface tractions governed by a single load multiplier and in the absence of body force. Also, only zero displacements may be prescribed on a portion of the boundary. A fuller treatment of these theorems, including proofs, may be found, e.g., in Koiter [9]. In a later section we will consider the case of random material properties and methods for their characterization.

**2.1 Upper-Bound Method.** The upper-bound collapse theorem of plastic limit analysis states that an elastic/perfectly plastic body cannot support a given system of loads if any kinematically admissible collapse mechanism  $\dot{e}_{ij}$  (not necessarily continuous) exists for which the work rate of the given loads exceeds the rate of plastic energy dissipation in the body, i.e., the body will collapse if

$$\int_{S_\sigma} p \sigma_i^0 v_i ds > \int_V D(\dot{e}_{ij}) dV + \int_{S_D} D_S(\delta \dot{e}_{ij}) ds, \quad (1)$$

where  $V$  is the volume of the body,  $\sigma_i^0$  is the distribution of applied stress for  $p=1$  on a portion  $S_\sigma$  of the body surface,  $S$ ,  $p$  is the load multiplier,  $S_D$  is an internal surface of velocity discontinuity,  $v_i$  is the velocity field,  $\delta \dot{e}_{ij}$  is the strain rate jump across surface  $S_D$ , and  $D$  and  $D_S$  are the dissipation functions associated with the strain rate  $\dot{e}_{ij}$  in  $V$  and  $\delta \dot{e}_{ij}$  on  $S_D$ , respectively. To be admissible,  $\dot{e}_{ij}$  and  $\delta \dot{e}_{ij}$  must satisfy the strain rate-velocity relations

$$\dot{e}_{ij} = \frac{1}{2}(v_{i,j} + v_{j,i}), \quad \delta \dot{e}_{ij} = \frac{1}{2}(\delta v_{i,j} + \delta v_{j,i}), \quad (2)$$

where  $\delta v_{ij}$  denotes the jump in the discontinuous derivatives of  $v_i$  on  $S_D$ . Further, the velocity field must satisfy the zero displacement boundary conditions on the portion  $S_u$  of the body surface  $S(S_\sigma + S_u = S)$ . Also, restrictions on admissible  $\dot{e}_{ij}$  and  $\delta \dot{e}_{ij}$  are imposed by the plastic flow rule as will be discussed.

The von Mises yield function is employed in the present study. The flow rule associated with the von Mises yield criterion requires that the strain rate field be incompressible ( $\dot{e}_{kk}=0$ ) and that no separation occurs normal to  $S_D$ . Then the dissipation functions can be written as

$$D(\dot{e}_{ij}) = k \sqrt{2 \dot{e}_{ij} \dot{e}_{ij}}, \quad D_S(\delta \dot{e}_{ij}) = k |\delta v_t|, \quad (3)$$

where  $k$  is the yield strength in pure shear and  $\delta v_t$  is the tangential velocity jump across  $S_D$ . For an inhomogeneous material,  $k = k(\mathbf{x})$ .

The numerical method used in the present study for the upper bound theorem first divides the physical domain into a plastic region and a rigid region. The plastic region is then divided into an assemblage of triangular elements for plane problems or tetrahedral elements for three-dimensional problems. The velocity field is taken to be constant in each element. Since  $\dot{e}_{ij}=0$  in each element, the velocity field results in zero volumetric dissipation in  $V$ , i.e.,  $D(\dot{e}_{ij})=0$ . Velocity discontinuities are allowed at the inter-element boundaries and on boundaries between the plastic region and the rigid region. The element mesh used for the upper-bound method (as well as for the lower-bound method to be discussed later) will be called the *physical element mesh* hereafter.

The inequality (1), may be symbolically written as

$$p \dot{W}(\mathbf{v}, \mathbf{x}) > \dot{F}(\mathbf{v}, \mathbf{x}), \quad (4)$$

where  $\dot{W}(\mathbf{v}, \mathbf{x})$  is the external rate of energy performed by  $\sigma_i^0$  and  $\dot{F}(\mathbf{v}, \mathbf{x})$  is the rate of internal energy dissipation,  $\mathbf{v}$  represents nodal values of an admissible velocity field and  $\mathbf{x}$  represents the nodal coordinates of the mesh. The least upper bound on the plastic limit load multiplier is obtained by performing the following optimization problem:

$$p = \underset{\mathbf{v}, \mathbf{x}}{\text{minimize}} \left\{ \frac{\dot{F}(\mathbf{v}, \mathbf{x})}{\dot{W}(\mathbf{v}, \mathbf{x})} \right\} \quad (5)$$

subject to  $\mathbf{A}\mathbf{v}=\mathbf{0}$ , and  $v_i^*=1$ ,

where  $v_i^*$  is a particular velocity component which is taken as 1 for normalization. The equality constraint  $\mathbf{A}\mathbf{v}=\mathbf{0}$  arises from the admissibility condition of no normal velocity jump across  $S_D$  (associated with the von Mises yield criterion). In view of the element discretization,  $\dot{F}(\mathbf{v}, \mathbf{x})$  can be written as

$$\dot{F}(\mathbf{v}, \mathbf{x}) = \sum_{i=1}^{N_S} \left\{ \int_{B_i} k(\mathbf{x}) |\delta v_i^{(i)}| dB_i \right\}, \quad (6)$$

where  $N_S$  is the total number of velocity jumps on element faces, and  $B_i$  is the length of side  $i$  for a plane problem or the area of surface  $i$  in a three-dimensional problem. The lengthy expressions for  $\dot{W}$  and  $\mathbf{A}$  in (5) are given in the theses of Ku [10].

**2.2 Lower-Bound Method.** The lower-bound collapse theorem of plastic limit analysis states that if an equilibrium distribution of stress  $\sigma_{ij}$  can be found which satisfies the equilibrium equations and meets the stress boundary condition on  $S_\sigma$  and is everywhere below the yield limit, then the body will not collapse under the applied loads.

A numerical method similar to that proposed by Lysmer [11] is used here to determine the plastic lower bound. In this method the whole body is divided into an assemblage of physical elements as in the upper-bound method. A linear stress field with unknown coefficients is assumed for each physical element. Equilibrium within the element is satisfied by adjusting the coefficients. The external equilibrium between the applied load on  $S_\sigma$  and the internal force can be written as

$$p \mathbf{s}^0 = \mathbf{H}(\mathbf{x})^T \mathbf{z}, \quad (7)$$

where  $p \mathbf{s}^0$  is the applied load vector,  $\mathbf{z}$  is the collection of unknown coefficients from all elements, and  $\mathbf{H}$  is an index matrix relating  $p \mathbf{s}^0$  to  $\mathbf{z}$ . In addition to satisfying equilibrium within each element, the assumed stress field must satisfy equilibrium between each pair of elements. This condition of inter-element equilibrium can be written as

$$\mathbf{Q}(\mathbf{x}) \mathbf{z} = \mathbf{0}. \quad (8)$$

The lengthy expressions for the matrix  $\mathbf{H}$  and  $\mathbf{Q}$  are given by Ku [10].

For computational purposes in problems of plane strain, the nonlinear von Mises yield function is approximated by a linear lower bound.

$$(\sigma_x - \sigma_y)^2 + (2\sigma_{xy})^2 \leq (2k)^2 \quad (9)$$

The equality describes the interior of a circle in the  $\sigma_x - \sigma_y$  and  $2\sigma_{xy}$ -plane with radius  $2k$ . The circle is then replaced by a combination linear segments inside the circle. The linearized yield function may be symbolically written as

$$\mathbf{D}(\mathbf{x}) \mathbf{z} \leq \mathbf{b}, \quad (10)$$

where the expressions for  $\mathbf{D}$  and  $\mathbf{b}$  may be found in Ku [10]. For an homogeneous distribution of yield strength,  $k(\mathbf{x}) = k = \text{const.}$ , it is sufficient to ensure the no-yield condition throughout  $V$  by checking the linearized yield function at each vertex of all the triangular physical elements (since the stress field is linear within each physical element). For the inhomogeneous case  $k(\mathbf{x})$ , satis-

fraction of (10) requires checking at additional points depending on the form of  $k(\mathbf{x})$  as will be discussed in Section 4.3.

The problem of finding the lower-bound plastic limit load can be formulated as the following optimization problem:

$$p = \underset{\mathbf{x}, \mathbf{z}}{\text{maximize}} [\mathbf{H}(\mathbf{x})^T \mathbf{z}] \quad (11)$$

$$\text{subject to } \mathbf{Q}(\mathbf{x})\mathbf{z} = \mathbf{0}, \text{ and } \mathbf{D}(\mathbf{x})\mathbf{z} \leq \mathbf{b}.$$

In the special case where the mesh shape  $\mathbf{x}$  is specified, (11) is a standard linear programming problem in  $\mathbf{z}$  and efficient algorithms are available for solving such problems (see, e.g., [12]).

### 3 Reliability Problem Formulation

**3.1 Random Field Discretization.** In the present study the random yield strength,  $k(\mathbf{x})$ , is considered to be a Gaussian random field of the form

$$k(\mathbf{x}) = \bar{k} + f(\mathbf{x}), \quad (12)$$

where  $\bar{k}$  is the constant mean yield strength and  $f(\mathbf{x})$  is a zero-mean, homogeneous random field representing the fluctuation in the yield strength. For the Gaussian random field, the mean and the covariance function fully describe the probabilistic structure of  $f(\mathbf{x})$ . For plane problems the covariance function for  $f(\mathbf{x})$  is taken as exponentially decaying in the form

$$C(x_1, y_1; x_2, y_2) = \sigma^2 \exp \left\{ -\pi \left[ \left( \frac{x_1 - x_2}{\delta_x} \right)^2 + \left( \frac{y_1 - y_2}{\delta_y} \right)^2 \right] \right\}, \quad (13)$$

where  $\sigma^2$  is the constant variance of  $f(\mathbf{x})$ . Here,  $\delta_x$  and  $\delta_y$  may be interpreted as the correlation lengths of the random field in the  $x$  and  $y$ -directions, respectively. For three-dimensional problems an additional term can be inserted for the  $z$ -direction. In their study of a plane slope reliability problem, Li and Lumb [3] observed that the particular functional form of the covariance function does not significantly affect the reliability results, which depend mainly on the characteristic lengths and the variance of  $f(\mathbf{x})$ .

In order to discretize the random field, let  $\sigma y_i$  be the random nodal value of  $f$  at coordinate  $\mathbf{x}_i$ , where  $y_i$  is not to be confused with coordinate variable. Then the random function  $f(\mathbf{x})$  can be approximated as

$$f(\mathbf{x}) = \sum_{i=1}^{N_r} \sigma y_i h_i(\mathbf{x}), \quad (14)$$

where  $N_r$  is the total number of nodal points in the random field mesh and  $h_i(\mathbf{x})$  is the interpolating shape function for nodal point  $i$ , which satisfies  $h_i(\mathbf{x}_j) = \delta_{ij}$ . With the standard deviation  $\sigma$  included in the series,  $y_i$  is a zero-mean Gaussian random variable with unit variance. A convenient choice for  $h_i(\mathbf{x})$  is a piecewise linear function over a selected mesh. This particular linear shape function method will be called the SF-Linear method. Other forms of shape functions are possible. A more accurate representation of the shape function was given by Li and der Kiureghian [13] in the form of

$$\mathbf{h}(\mathbf{x}) = \mathbf{C}_{\mathbf{Y}\mathbf{Y}}^{-1} \mathbf{C}_{f(\mathbf{x})\mathbf{Y}}, \quad (15)$$

where  $\mathbf{h}(\mathbf{x})^T = [h_1(\mathbf{x}), \dots, h_{N_r}(\mathbf{x})]$  is a row vector of dimension  $N_r$ ;  $\mathbf{C}_{f(\mathbf{x})\mathbf{Y}}$  is a column vector of dimension  $N_r$  storing the covariance between  $f(\mathbf{x})$  and each random field nodal values; and  $\mathbf{C}_{\mathbf{Y}\mathbf{Y}}$  is the  $N_r \times N_r$  covariance matrix between each pair of nodal points. Both  $\mathbf{C}_{f(\mathbf{x})\mathbf{Y}}$  and  $\mathbf{C}_{\mathbf{Y}\mathbf{Y}}$  can be readily obtained by considering (13). This shape function is obtained from the optimal linear estimation (OLE) theory and this discretization method will be called the SF-OLE method.

**3.2 First-Order Reliability Method (FORM).** We review here a formulation of reliability theory given by Ditlevsen and Madsen [14]. The Gaussian random variables governing the reli-

ability of a system can be collected into a vector  $\mathbf{Y}$ , and a limit state function  $g(\mathbf{y})$  is introduced such that the system is safe if  $g(\mathbf{y}) > 0$  and fails if  $g(\mathbf{y}) \leq 0$ . The failure probability is given by the following integral:

$$p_f = \int_{g(\mathbf{y}) \leq 0} f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y}, \quad (16)$$

where  $f_{\mathbf{Y}}(\mathbf{y})$  is the joint probability density distribution function of random variables  $\mathbf{Y}$ . The reliability of the system is defined as  $1 - p_f$ . Let  $\mu_{\mathbf{Y}}$  and  $\mathbf{C}_{\mathbf{Y}\mathbf{Y}}$  denote the vector of the mean value and the covariance matrix of  $\mathbf{Y}$ , and let  $\mathbf{U}$  be a set of mutually independent, standard normal random variables. A linear transformation between  $\mathbf{Y}$  and  $\mathbf{U}$  may be written as

$$\mathbf{Y} = \mu_{\mathbf{Y}} + \mathbf{T}\mathbf{U}, \text{ or } \mathbf{u} = \mathbf{T}^{-1}(\mathbf{Y} - \mu_{\mathbf{Y}}), \text{ where } \mathbf{C}_{\mathbf{Y}\mathbf{Y}} = \mathbf{T}\mathbf{T}^T. \quad (17)$$

After the transformation from  $\mathbf{Y}$  to  $\mathbf{U}$ , the limit state function  $g(\mathbf{y})$  may be stated in the standard normal  $\mathbf{U}$  space as

$$G(\mathbf{u}) = g(\mu_{\mathbf{Y}} + \mathbf{T}\mathbf{u}). \quad (18)$$

The first-order reliability method evaluates the integral (16) approximately by first solving the following nonlinear optimization problem:

$$\underset{\mathbf{u}}{\text{minimize}} \{ \beta = |\mathbf{u}| \} \text{ subject to } G(\mathbf{u}) = 0, \quad (19)$$

where  $\beta$  is the reliability index and has the geometrical meaning of being the point on  $G(\mathbf{u}) = 0$  which is the closest to origin in  $\mathbf{u}$  space. The solution of (19) is termed the most probable failure point or design point and is denoted by  $\mathbf{u}^*$ . Then, the first-order approximation of failure probability is given by

$$p_f = \Phi(-\beta), \quad (20)$$

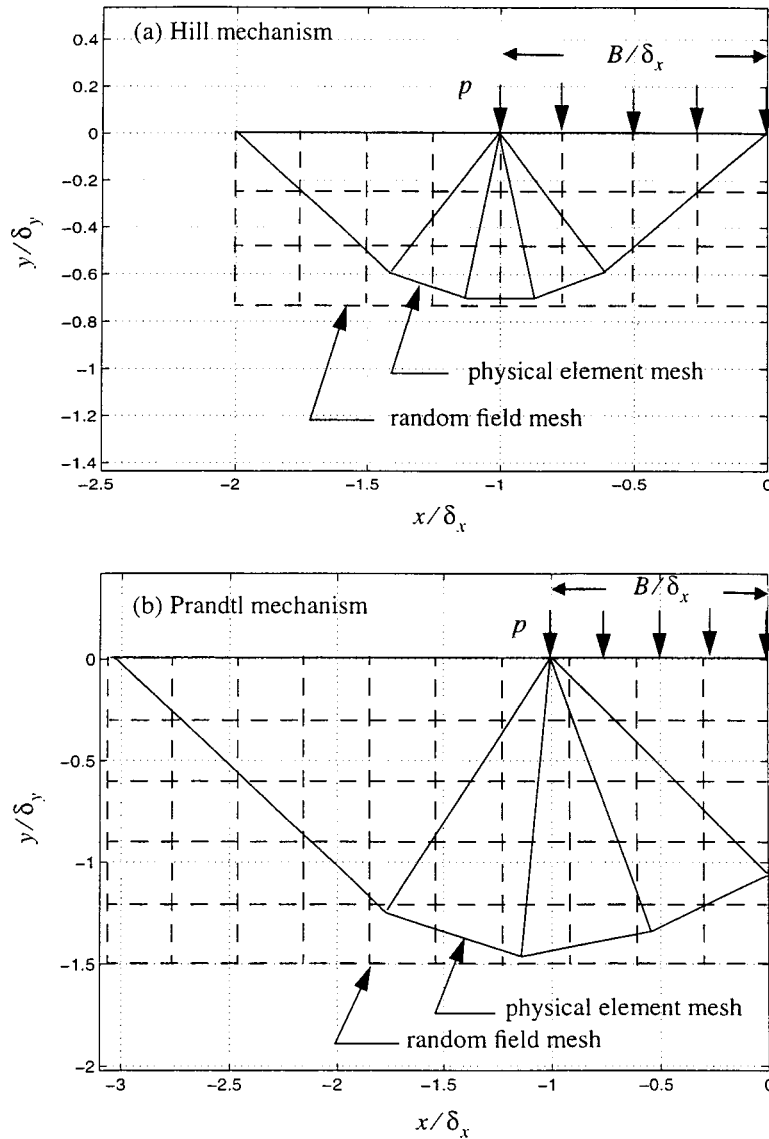
where  $\Phi(\cdot)$  is the cumulative Gaussian distribution function. There is also a second-order approximation to  $p_f$  which is not used in the present study.

### 4 Example Problem: Half-Space Under Normal Load

The principles state in the previous two sections will be combined and applied to a half-space under normal load as an example problem. In the reliability analysis the mesh coordinates in (5) and (11) can be kept fixed at the values obtained from a deterministic analysis to simplify the reliability computation. The validity of this simplification will be verified in a latter section where numerical results are presented. In addition, by applying this simplification, we obtain several useful characteristics of the limit state functions in the reliability method derived for the upper and lower-bound approaches. These derivations are presented in the present section.

#### 4.1 Two-Dimensional Upper-Bound Limit State Function.

Figures 1(a) and 1(b) show the mesh discretizations based on the slip line solutions of the well-known Hill and Prandtl mechanisms. The accuracy of these discretizations in the deterministic analysis can be improved by using more triangles to model the centered fan region. For the discretization meshes shown in Figs. 1(a) and 1(b), it can easily be shown ([10]) that the kinematic equality constraint  $\mathbf{A}\mathbf{v} = \mathbf{0}$  in (5) together with  $\mathbf{v}_i^* = 1$  constitutes a deterministic (square matrix) system of equations and  $\mathbf{v}$  is uniquely determined. However, for finer discretizations the kinematic equality constraints matrix is not square and optimization must be carried out on the velocity field. It will be illustrated in the following section by way of a numerical example that the simple discretizations as shown in Figs. 1(a) and 1(b) produce reasonably accurate results both in the deterministic and reliability analyses; thus velocity optimization will not be needed in the example problems considered here.



**Fig. 1 Element discretization for (a) Hill mechanism and (b) Prandtl mechanism**

In the upper-bound reliability method the limit state function for reliability analysis at a fixed load multiplier  $p$  may be written as

$$g(\mathbf{y}) = \frac{F(\mathbf{v}, \mathbf{y})}{\bar{W}(\mathbf{v})} - p. \quad (21)$$

After combining (6) and (14),  $\dot{F}(\mathbf{v}, \mathbf{y})$  can be expressed as

$$\dot{F}(\mathbf{v}, \mathbf{y}) = \sum_{i=1}^{N_s} \left\{ |\delta v_i^{(i)}| \int_{B_i} \sum_{j=1}^{N_r} \sigma y_j h_j(\mathbf{x}) dB_i \right\}. \quad (22)$$

It may be shown that the limit state function  $g(\mathbf{y})$  is linear in  $\mathbf{y}$ , i.e.,

$$g(\mathbf{y}) = (\mu_0 + \mathbf{d}^T \mathbf{y}) - p, \quad (23)$$

where  $\mu_0$  is the mean plastic strength. Using (17), the limit state function in standard normal  $\mathbf{U}$  space can be written as

$$G(\mathbf{u}) = R(\mathbf{u}) - p, \quad \text{where } R(\mathbf{u}) = \mu_0 + \mathbf{e}^T \mathbf{u}, \quad (24)$$

where  $\mathbf{e} = \mathbf{T}^T \mathbf{d}$ . The above limit state function is in the form of the difference between resistance  $R$  and the applied load multiplier  $p$ ,

thus  $R$  may be regarded as the plastic limit strength (limit load) of the half-space. It may be shown ([10]) that  $R$  is a Gaussian random variable with the following mean value and standard deviation:

$$\mu_R = \mu_0, \quad \sigma_R = |\mathbf{e}|. \quad (25)$$

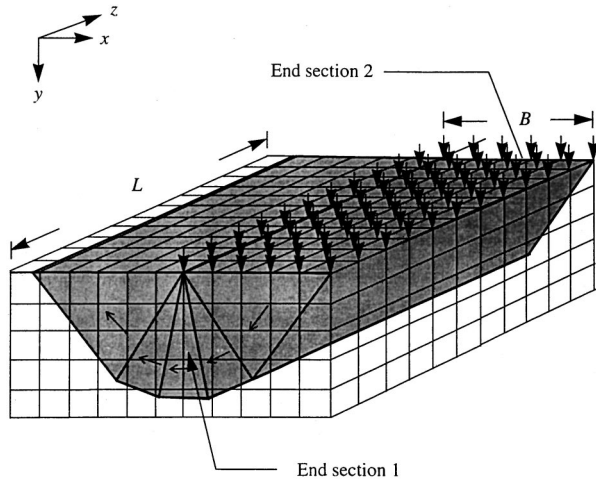
It was possible to obtain the analysis expression (24) for the distribution of  $R$  in the upper-bound method because the limit state function, (23) or (24), are linear and  $\mathbf{v}$  is known from deterministic analysis. In the lower-bound method, to be discussed next, the limit state function is nonlinear and the distribution of  $R$  has to be determined numerically by repeatedly changing the applied load multiplier  $p$  and re-solving the reliability problem.

#### 4.2 Three-Dimensional Upper-Bound Limit State Function.

When the indentation pressure acts over a limited length,  $L$ , a three-dimensional problem results, as illustrated in Fig. 2. Additional assumptions involved in considering such a three-dimensional failure mechanism will be discussed in later sections.

A limit state function for the three-dimensional problem similar to (23) can be expressed as





**Fig. 2 Failure mechanism of a three-dimensional half-space and random field discretization mesh**

$$g(\mathbf{y}) = (\mu'_0 + \mu_e + \mathbf{d}_0^T \mathbf{y} + \mathbf{d}_e^T \mathbf{y}) - p, \quad (26)$$

where  $\mu'_0$  is the mean plastic strength provided by the failure surfaces in the longitudinal direction, and  $\mu_e$  is the mean plastic strength contributed by the two vertical end sections. After transforming the variables into the standard normal space, the limit state function can be written as

$$G(\mathbf{u}) = (R_L + R_e) - p, \quad (27)$$

where  $R_L = \mu'_0 + \mathbf{b}_1^T \mathbf{u}$  is the random strength provided by the failure surfaces in the longitudinal direction,  $R_e = \mu_e + \mathbf{b}_2^T \mathbf{u}$  is the random strength provided from the two end sections, and  $\mathbf{b}_1 = \mathbf{d}_0^T \mathbf{T}$ ,  $\mathbf{b}_2 = \mathbf{d}_e^T \mathbf{T}$  in which  $\mathbf{T}$  is the transformation matrix between random variables  $\mathbf{y}$  and  $\mathbf{u}$ . The plastic limit load of the three-dimensional half-space,  $R$ , is the sum of  $R_L$  and  $R_e$ . The probability distribution of  $R$  is a normal distribution with the mean value and standard deviation given by

$$\mu_R = \mu'_0 + \mu_e, \quad (28)$$

$$\sigma_R^2 = |\mathbf{b}_1 + \mathbf{b}_2|^2 = |\mathbf{b}_1|^2 + |\mathbf{b}_2|^2 + 2\mathbf{b}_1 \cdot \mathbf{b}_2 \\ = \text{Var}[R_L] + \text{Var}[R_e] + 2\text{Cov}[R_L, R_e]. \quad (29)$$

**4.3 Lower-Bound Limit State Functions.** In applying the lower-bound reliability method, again we do not optimize on the physical element discretization mesh  $\mathbf{x}$  in (11), i.e.,  $\mathbf{H}(\mathbf{x})$ ,  $\mathbf{Q}(\mathbf{x})$ , and  $\mathbf{D}(\mathbf{x})$  are kept fixed in the reliability optimization problem (19). The validity of this simplification will be illustrated by way of numerical examples in the following section.

The limit state function  $g(\mathbf{y})$  is defined by (11) with the applied load  $p$  moved to the right-hand side, similar to (21). We note that the random variables only appear in the yield strength vector  $\mathbf{b}$ .

It has been mentioned previously that for homogeneous constant yield strength it is sufficient to check the yield conditions at the vertices of each physical element to ensure that no-yield condition is satisfied everywhere. For nonhomogeneous yield strength the yield check must be performed at more points. Since the stress field is assumed linear within each physical element, it is possible to ensure the no-yield condition when the SF-Linear method (discussed in Section 3.1) is employed to discretize the random field by checking at a small number of known points. However, if other random field discretization methods are used, e.g., the SF-OLE method, then relative minima of the yield function must be located in each physical element and the yield condition must be checked at these minima. This calculation would seriously deter-

riorate the numerical efficiency of the lower-bound reliability method. Thus, only the SF-Linear method is used in the present study.

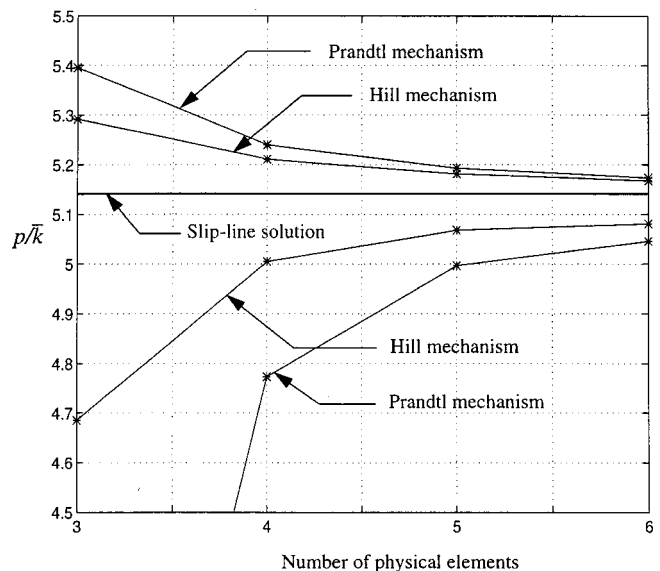
Determination of the gradient of the limit state function with respect to random variables is an important step in reliability optimization calculations using (19). An efficient method of calculating the gradient vector of  $g(\mathbf{y})$  with respect to component  $y_k$  exists for the proposed lower-bound reliability method. As stated above, since the function  $g(\mathbf{y})$  depends only on  $y_k$  through  $\mathbf{b}$ , we may write

$$\frac{\partial g}{\partial y_k} = \sum_{i=1}^{N_l} \frac{\partial g}{\partial b_i} \frac{\partial b_i}{\partial y_k}, \quad (30)$$

where  $N_l$  is the total number of linear inequality constraints. The term  $\partial b_i / \partial y_k$  can be readily obtained from (10), (11), (12), and (14). The term  $\partial g / \partial b_i$  is the sensitivity of the linear program (11) with respect to  $\mathbf{b}$ . This sensitivity measure is the dual solution of the linear programming problem. The dual solution is a by-product in solving a linear programming problem (see, e.g., [12]), thus only minimal numerical effort is needed to evaluate the gradient  $\partial g / \partial y_k$ . The behavior of the dual solution is such that it remains constant for a small perturbation of  $\mathbf{b}$  but varies when the perturbation is large enough. Thus the limit state function  $g(\mathbf{y})$  exhibits piecewise linear characteristics. However, numerical experience shows that the convergence behavior in solving (19) is similar to that of solving a smooth limit state function since the linear patterns are relatively small.

## 5 Numerical Results

Failure probabilities are evaluated numerically for two-dimensional and three-dimensional indentation problems for an elastic/perfectly plastic material with randomly varying yield strength using the methods developed in the preceding sections. The applied indentation pressure multiplier  $p$  is assumed to be a deterministic constant, although the proposed methods have no difficulty treating problems with  $p$  as a random variable. Before considering the random case, in order to verify proposed computational methods, we present deterministic numerical results for a material with constant yield strength  $k$ .



**Fig. 3 Deterministic upper and lower-bound results for a plane half-space indentation problem**

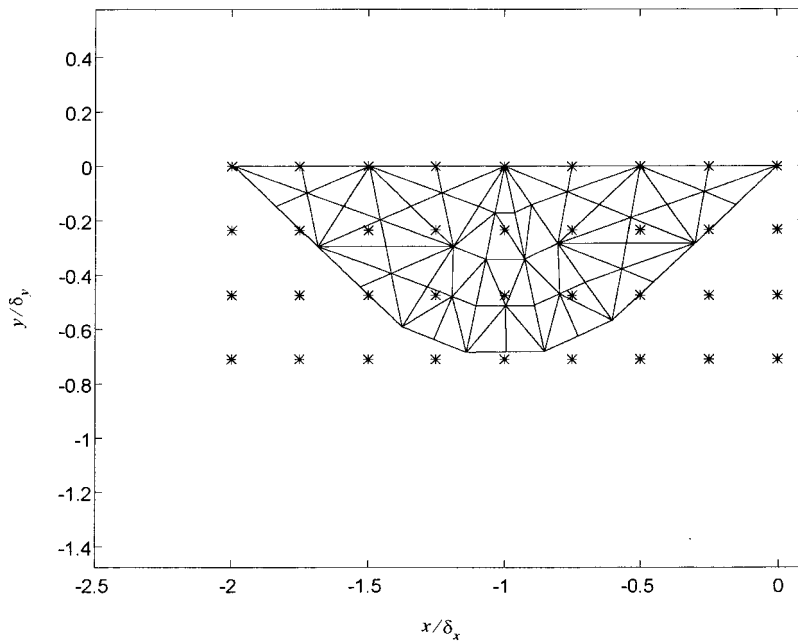


Fig. 4 Element discretization of 78 elements and 36 random field nodes

**5.1 Two-Dimensional Deterministic Half-Space Indentation Problem.** The well-known theoretical plastic collapse pressure for a half-space in plane strain without body force, according to the slip line theory, coincides for the Hill and Prandtl mechanisms and is given by

$$p^* = (2 + \pi)\bar{k}. \quad (31)$$

In the numerical computations the optimization on mesh coordinates has been performed as shown in Fig. 1 for Hill and Prandtl mechanisms with a five-element discretization. Numerical results for collapse pressure are shown in Fig. 3 for three, four, five, and six elements. The same optimized mesh coordinates are obtained for the upper and lower-bound methods using the same number of

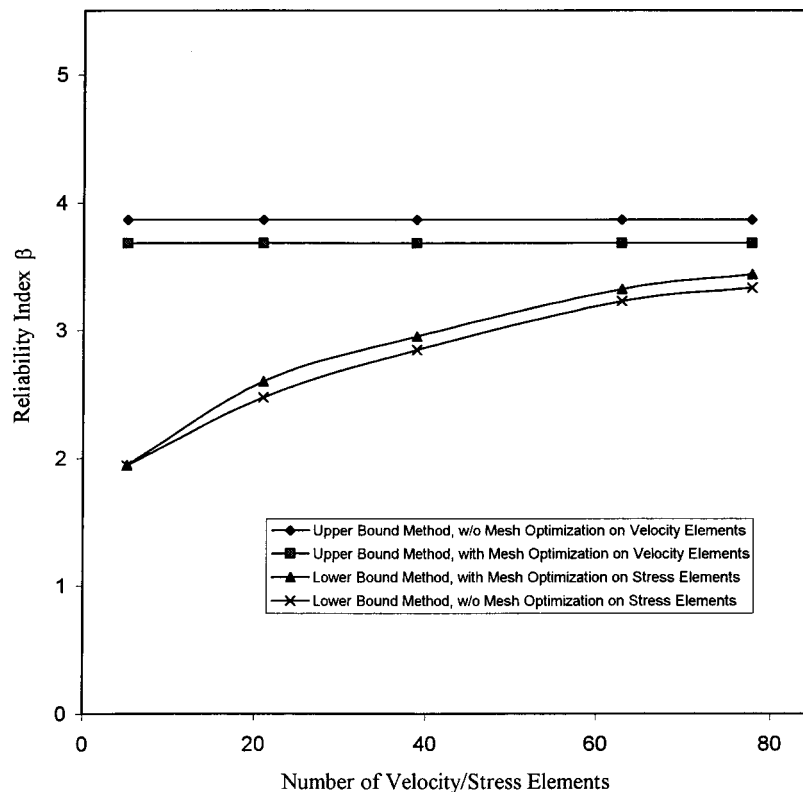


Fig. 5 Convergence of reliability index with increasing number of random field nodes

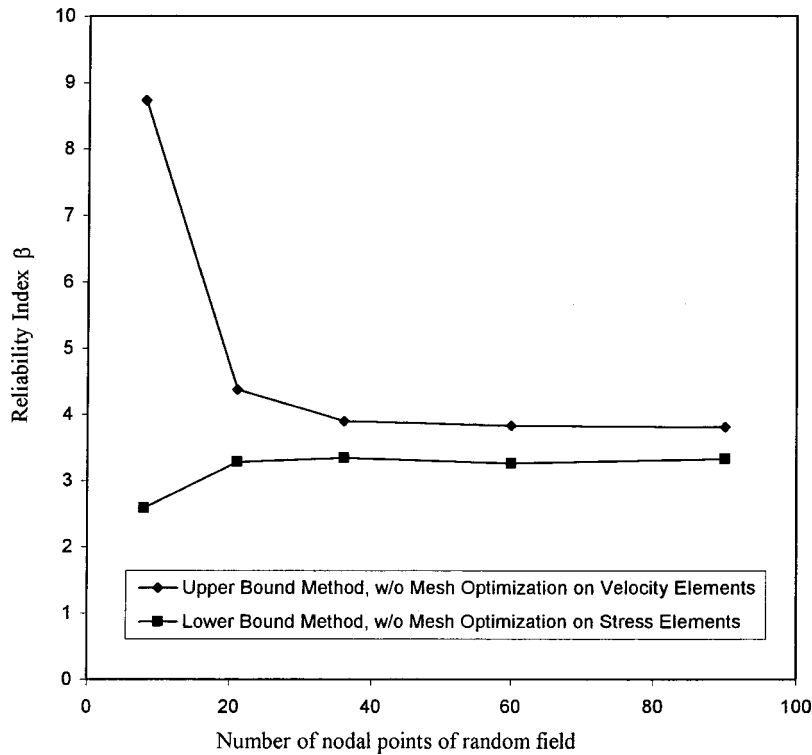


Fig. 6 Convergence of reliability index with increasing number of random field nodes

elements. The results of Fig. 3 show that five elements suffice for reasonably accurate (between 2.7 percent and 1.2 percent) calculation of bounds on the collapse pressure in the deterministic case.

**5.2 Reliability Analyses of a Two-Dimensional Half-Space Considering the Hill Mechanism.** A few convergence studies are conducted first to determine the appropriate sizes needed for the velocity and stress element, as well as for the random field mesh. A series of velocity and stress element discretizations were used beginning from the five-element mesh shown in Fig. 1 and ending with a fine element mesh with 78 elements shown in Fig. 4, where the average element linear size is approximately  $B/4$ . Three other meshes intermediate between  $B$  and  $B/4$  (with 21, 39, and 63 elements) are also used in the convergence study, where  $B$  is the length of the applied indentation load. These meshes are all based on the basic pattern of five-element mesh, with varying details of dividing within each of the basic five elements.

Also shown in Fig. 4 is the random field mesh with 36 ( $9 \times 4$ ) random field nodes. Four other meshes are used for the convergence study, with the discretizations of  $6(3 \times 2)$ ,  $21(7 \times 3)$ ,  $60(12 \times 5)$ , and  $90(15 \times 6)$  nodes. Only half of the geometry is considered due to the symmetry of the problem. However, in reliability analysis the material yield strength is randomly distributed and hence not symmetrical. It may be readily understood that failure over either half of the geometry is sufficiently critical than requiring failure on both side of the geometry; thus reliability analysis on only half of geometry is appropriate as will be done in the present study.

The following parameters are used for the convergence study:

$$p/\bar{k} = 0.85p^*/\bar{k}, \quad \sigma/\bar{k} = 0.115, \quad (32)$$

$$\delta_x/B = \delta_y/B = 0.4.$$

Figure 5 shows the convergence of the reliability index for the upper and lower bound methods using increasingly finer meshes for the velocity and stress elements. An interesting observation for the upper-bound method is that  $\beta$  is insensitive to using finer

meshes. This shows that within the five-element pattern the optimal solution for the upper-bound method requires only the five regions moving as rigid bodies even when the yield strength varies inhomogeneously. Convergence study for the stress element of the lower-bound method in Fig. 5 shows increasing accuracy of  $\beta$  as the stress elements are refined.

In Fig. 5, results are shown for cases where the mesh shape is treated as fixed or variable in the optimization process to find the lowest  $\beta$ . For the given parameter given in (32), the upper and lower bounds of  $\beta$  for the fixed mesh case are found to be  $\pm 7.4$  percent relative to the average (3.86 and 3.33), and  $\pm 3.4$  percent relative to average for the variable mesh case (3.68 and 3.44). It should be noted that significantly more computational efforts have to be spent for including the mesh shapes as variables in the optimization Eqs. (5) and (11).

Figure 6 shows the convergence study for refining the random field mesh while the velocity and stress element meshes are kept constant. The random field mesh of 36 nodes, shown in Fig. 4, is found to be appropriate for the parameters considered in (32).

Figure 7 shows the effects on the reliability indices of varying the correlation length  $\delta$ . For smaller  $\delta$  the realizations of random field will typically have more peaks and troughs than for larger  $\delta$ . With more peaks and troughs there is an averaging tendency upon integration to compute the internal energy dissipations (for the upper bound method), and the strength computed will be more concentrated around the average value given by (31). The associated reliability index  $\beta$  is higher for smaller  $\delta$ , because strength thus computed is less likely to fall below the given external load when simulations are performed. As  $\delta \rightarrow \infty$ , the random field becomes a single random variable and  $\beta$  approaches a constant value. Results in Fig. 7 for the lower bound methods follow the same characteristics.

For the case (32), the lower-bound FORM reliability analysis gives  $\beta = 3.33$  ( $p_f = 4.34 \times 10^{-4}$ ) for  $\delta_x = 0.4$ . The Monte Carlo simulation (MCS) using importance sampling ([15]) gives  $\beta = 3.36$  ( $p_f = 3.88 \times 10^{-4}$ ) for a coefficient of variation of four

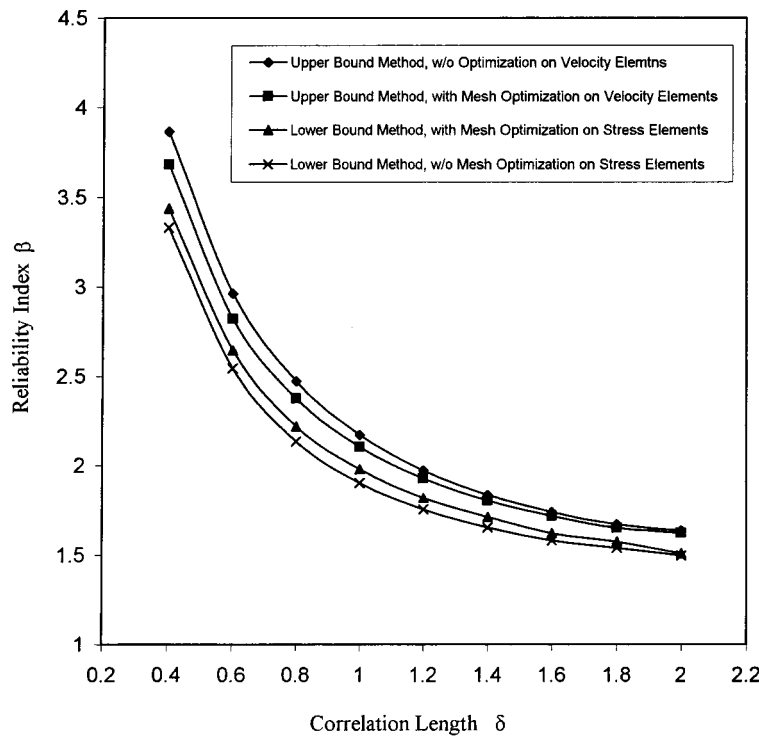


Fig. 7 Variation of reliability index with increasing correlation length

percent, which is considered sufficiently small. The sampling density function in the importance sampling simulation is chosen as a standard normal distribution centered at the design point  $\mathbf{u}^*$  (as calculated from FORM) in the standard normal space  $\mathbf{u}$ . A total of 3000 simulations is carried out in the MCS importance sampling calculation.

Based on the MCS result, it may be seen that the FORM analysis gives a very accurate approximation to the lower-bound reliability for this case. However, for very small correlation length,  $\delta_x = 0.4$ , the current stress element mesh may have to be further refined to obtain the same accuracy.

This confirmation of the FORM analysis is essential here since, as discussed in Section 4.3, the lower-bound limit state function is nonsmooth and exhibits piecewise linear characteristics. The computational efforts required by MCS using importance sampling is several hundred times that required for the FORM calculations. By varying the applied load a full cumulative probability distribution of  $p/\bar{k}$  has been obtained as shown in Fig. 8 for both the upper and lower-bound methods.

**5.3 Reliability Analyses of Two-Dimensional Half-Space Considering the Prandtl Mechanism.** An analysis similar to the foregoing is performed using the upper-bound method for the Prandtl mechanism. The physical element and random field element mesh are shown in Fig. 1(b). For the given problem parameters, (32), the upper-bound reliability index  $\beta$  is found to be 5.38, compared to  $\beta = 3.86$  when the Hill mechanism is considered. Thus, for small applied load,  $p < p^*$ , it is much more likely for the half-space to fail in the Hill mechanism rather than in the Prandtl mechanism. The plastic limit load for the two mechanisms are both Gaussian distributed. The two mechanisms have approximately the same mean values but the Prandtl mechanism has a smaller standard deviation. This is due to the larger extent of the plastic region involved in the Prandtl mechanism. Further, the smaller standard deviation explains the higher reliability index of the Prandtl mechanism.

#### 5.4 Reliability Analyses of Three-Dimensional Half-Space.

A three-dimensional indentation problem is illustrated in Fig. 2. The reliability of the three-dimensional problem is studied here using only the upper-bound method.<sup>2</sup> We assume that the plastic region of the three-dimensional half-space lies inside a cylindrical surface and is bounded by two vertical surfaces a distance  $L$  apart. The possibility of a plastic region with failure length less or

<sup>2</sup>The lower-bound method also could be applied to the three-dimensional problem but leads to intensive computations that were not pursued in this investigation.

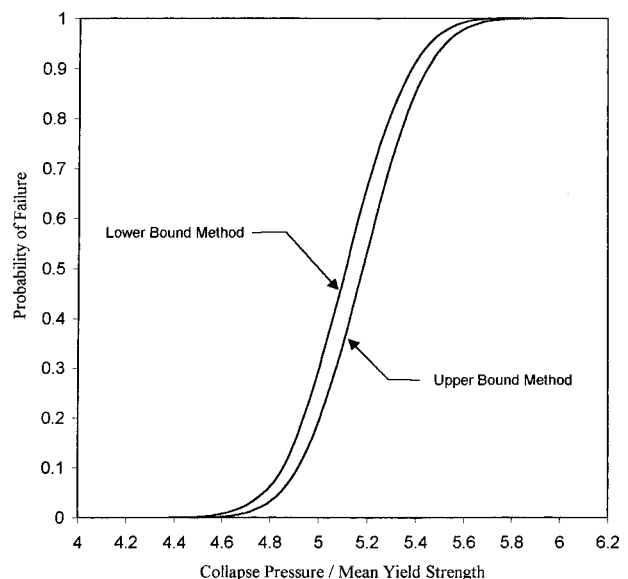
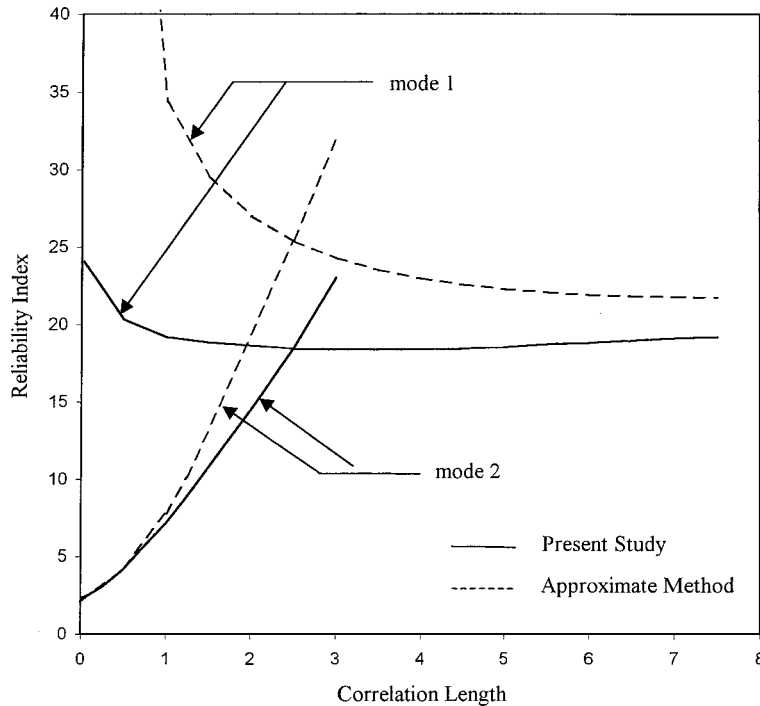


Fig. 8 Cumulative probability distribution of collapse pressure for the plane half-space problem with Hill's mechanism





**Fig. 9 Variation of reliability index with respect to increasing failure length for three-dimensional half-space under load  $p=0.85p^*$**

greater than  $L$  is not considered in this study. This point may call for further investigation depending on the nature of loading in actual applications.

Results of reliability calculations for the three-dimensional half-space using the method of Section 4.2 are shown in Fig. 9 with the failure length varying from 0 to  $7.5 \times \delta$  where  $\delta$  is the given correlation length of the random yield strength field. For simplicity of presentation the correlation lengths in the  $x$ ,  $y$ , and  $z$ -directions are assumed equal and denoted by  $\delta$ . Results are calculated for the case of  $B/\delta=2.5$ , where  $B$  is the width of the applied pressure region.

The assumed failure mechanism has two possible orientations (modes) for the failure cylinder. With reference to Fig. 2, in the mode 1 failure mechanism the failure cylinder rotates about an axis parallel to the  $z$ -direction. In the mode 2 failure mechanism the cylinder is turned 90 deg and rotates about the  $x$ -axis. The possible occurrence of both failure modes constitutes a series system reliability problem (see, e.g., [14]).

Results for the reliability index are shown in Fig. 9 for a range of  $L/\delta$  for the two modes of failure. For the mode 1 mechanism at low values of  $L/\delta$  the effect of the end sections dominates and leads to high reliability indices (low probability of failure). If the end sections are not considered, when  $L/\delta \rightarrow 0$  the reliability index will approach the two-dimensional upper-bound solution (in Section 5.2) for which  $\beta=3.86$ . However, with the consideration of end sections, the half-space is much stronger and thus much higher  $\beta$  values are observed in this regime.

For the mode 2 mechanism the roles of  $B/\delta$  and  $L/\delta$  are interchanged. After calculations based on the proposed method, the reliability index for mode 2 starts from  $\beta=2.35$ . From Fig. 9, it is seen that mode 2 is more likely to occur than mode 1 in the range of  $0 \leq L/\delta \leq 2.5 = B/\delta$ . For the range of  $L/\delta > 2.5$ , the mode 1 mechanism is more likely to occur than mode 2 mechanism.

As mentioned previously, the failure of the half-space is a series problem with two mechanisms combined. However, owing to the dominant nature of one on the other mechanism except near of  $L/\delta=2.5$ , the system reliability variation curve will be nearly the same as the more dominant component curve. At  $L/\delta=2.5$ , nu-

merical results show the two modes are weakly correlated (modal correlation coefficient 0.03). The system reliability index is virtually identical to the component index of  $\beta=18.5$  at  $L/\delta=2.5$  for our present example.

In summary, the system reliability index variation will be virtually the same  $\beta$  variation as the more dominant mode, except near  $L/\delta=2.5$  where we expect to see a very small smooth transition zone (not shown in Fig. 9).

For this same problem, the approximate method of spatial average and variance reduction ([2,4]) is employed to produce approximate solutions as follows: (1) the probability distribution of the limit load of a two-dimensional problem is obtained using a known method, (2)  $Var[R_L]$  (see Section 4.2) is estimated by the so-called "variance reduction function," (3)  $Var[R_e]$  and  $Cov[R_L, R_e]$  are neglected, (4) the probability distribution of three-dimensional limit load is assumed to be Gaussian with known mean value and the estimated variance from step (2). This approximate method has been applied to three-dimensional earth slope stability problems. The solutions for mode 1 and 2 obtained by this method are also plotted in Fig. 9. For very small  $L/\delta$  the approximate method produces unrealistically high  $\beta$  for mode 1. This is due to the fact that at this range of  $L/\delta$  the end effects are significant, thus neglecting  $Var[R_e]$  and  $Cov[R_L, R_e]$  result in large errors as has been discussed by Vanmarcke [2]. For the approximate mode 2 curve at  $L/\delta=0$ , the reliability index starts from  $\beta=2.10$  (a result of solving a two-dimensional problem by the proposed method). The general trends of the approximate solutions follows those of the solutions obtained by the present limit analysis-based methods.

## 6 Conclusion

Methods are developed in this study to assess the reliability of an elastic/perfect plastic medium against plastic collapse. The methods are illustrated by considering a half-space under normal load. By combining the plastic limit theorems with the methods of modern reliability theory, upper and lower bounds on the reliability index may be computed efficiently for plane problems. The

probabilistic plastic collapse based on the classical Prandtl and Hill mechanisms are compared. The three-dimensional reliability problem of half-space under normal load over a rectangle is solved using the upper-bound method. The present results are compared with those of a previous approximate method and fair agreement is found for the critical failure mode. The methods developed here can be applied to other problems and other yield conditions. A study of slope stability under the Mohr-Coulomb yield criterion is in progress.

## References

- [1] Wang, W., Ramirez, M., and Corotis, R., 1994, "Reliability Analysis of Rigid-Plastic Structures by the Static Approach," *Struct. Safety*, **15**, pp. 209–235.
- [2] Vanmarcke, E. H., 1980, "Probabilistic Stability Analysis of Earth Slopes," *Engineering Geology*, **16**, pp. 29–50.
- [3] Li, K. S., and Lumb, P., 1987, "Probabilistic Design of Slopes," *Can. Geotech. J.* **24**, pp. 520–535.
- [4] Yucemen, M. S., and Al-Homoud, A. S., 1990, "Probabilistic Three-Dimensional Stability Analysis of Slopes," *Struct. Safety*, **9**, pp. 1–20.
- [5] Nordgren, R. P., 1992, "Limit Analysis of a Stochastically Inhomogeneous Plastic Medium With Application to Plane Contact Problems," *ASME J. Appl. Mech.*, **59**, pp. 477–484.
- [6] Ostoja-Starzewski, M., and Ilies, H., 1996, "The Cauchy and Characteristic Boundary Value Problems of Random Rigid-Perfectly Plastic Media," *Int. J. Solids Struct.*, **33**, No. 8, pp. 1119–1136.
- [7] Jiang, M., Ostoja-Starzewski, M., and Jasiuk, I., 2000, "Bounding of Effective Elastoplastic Response of Random Composites," *PLASTICITY'2000—Eighth Intl. Symposium on Plasticity*, Whistler, BC, Canada, Neat Press, MD.
- [8] Ostoja-Starzewski, M., 1999, "Microstructural Disorder, Mesoscale Finite Elements, and Macroscopic Response," *Proc. R. Soc. London, Ser. A*, **A455**, pp. 3189–3199.
- [9] Koiter, W. T., 1960, "General Theorems for Elastic-Plastic Solids," *Progress in Solid Mechanics*, Vol. 1, I. N. Sneddon and R. Hill, eds., North-Holland, Amsterdam, pp. 167–221.
- [10] Ku, P.-D., 1997, "Probabilistic Analysis of the Plastic Collapse of Random Media," Ph.D dissertation, Rice University, Houston, TX.
- [11] Lysmer, J., 1970, "Limit Analysis of Plane Problems in Solid Mechanics," *J. Soil Mech. Found. Div., Am. Soc. Civ. Eng.*, **96**, No. 4, pp. 1311–1334.
- [12] Luenberger, D. G., 1984, *Linear and Nonlinear Programming*, Addison-Wesley, Reading, MA.
- [13] Li, C.-C., and Der Kiureghian, A., 1993, "Optimal Discretization of Random Fields," *J. Eng. Mech.*, **119**, No. 6, pp. 1136–1154.
- [14] Ditlevsen, O., and Madsen, H. O., 1996, *Structural Reliability Methods*, John Wiley and Sons, West Sussex, UK.
- [15] Melchers, R. E., 1999, *Structural Reliability Analysis and Prediction*, John Wiley and Sons Ltd., West Sussex, UK.

# Mode II Edge Delamination of Compressed Thin Films

**D. S. Balint**

Assoc. Mem. ASME

**J. W. Hutchinson**

Fellow ASME

Division of Engineering and Applied Sciences,  
Harvard University,  
Cambridge, MA 02138

*Ceramic coatings deposited on metal substrates generally develop significant compressive stresses when cooled from the temperature at which they are processed as a result of thermal expansion mismatch. One of the main failure modes for these coatings is edge delamination. For an ideally brittle interface, the edge delamination of a compressed thin film involves mode II interface cracking. The crack faces are in contact with normal stress acting across the faces behind the advancing tip. Frictional shielding of the crack tip has been shown to increase the apparent fracture toughness. Roughness effects associated with the separating faces can also contribute to the apparent toughness. A model of mode II steady-state edge delamination that incorporates combined friction and roughness effects between the delaminated film and substrate is proposed and analyzed. This model is used to assess whether frictional shielding and surface roughness effects are sufficient to explain the large apparent mode II fracture toughness values observed in experiments.*

[DOI: 10.1115/1.1388012]

## 1 Introduction

Many thin film manufacturing techniques create residual stresses in the film that can lead to failure. Ceramic coatings deposited on metal substrates generally develop significant compressive stresses as a result of thermal expansion mismatch when cooled from the temperature at which they are processed. Systems of this type are of interest as thermal barrier and wear coatings. Two of the primary failure modes for films in compression are edge delamination and buckle delamination (c.f. Fig. 1). The elastic energy per unit area stored in the film which is available upon edge delamination is

$$G_0 = \frac{(1 - \nu^2)\sigma^2 h}{2E} \quad (1)$$

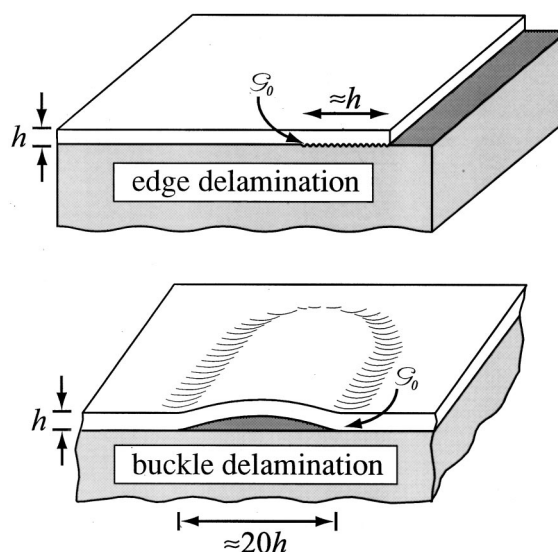
where  $E$  and  $\nu$  are the Young's modulus and Poisson's ratio of the film,  $h$  is the film thickness, and  $\sigma$  is the uniform compressive biaxial prestress in the film. To a first approximation,  $G_0$  is also the energy available to drive the interface delamination crack for buckle delaminations. The typical flaw size (in the form of a debonded region) needed to initiate a buckle delamination is about  $20h$ . Delaminations that emanate from a film edge and terminate in the interior of the substrate surface only require debond flaws as small as 1 or  $2h$  for initiation ([1]). This would suggest that edge delaminations would be more commonly observed than buckle delaminations. In fact, the opposite is true. Many systems seem to fail primarily by buckle delaminations initiated away from the edges of the film.

There are many reasons why edge delaminations are less common than might be expected. Edge delamination is a mode II cracking phenomenon when the film is in compression, and it is now well known that mode II tends to be associated with the highest interface toughness. By contrast, buckle delamination is mixed mode but approaches mode II as the delamination spreads and arrests ([2]). In addition, there are extrinsic effects accompanying mode II edge delamination which contribute to the apparent toughness. Frictional sliding is one such effect ([3]). Another, which is the primary focus of this paper, is the effect that surface

roughness has on the delaminated film as it slides over the substrate. The roughness forces open the film-substrate interface as the delamination crack faces displace, resulting in normal stresses at the interface that are larger than those produced in the absence of roughness. This leads to greater frictional dissipation and shielding of the crack tip. At the same time, however, it can also wedge open the crack tip producing a mode I stress intensity component. Most failed interfaces have a characteristic roughness. The present study suggests that frictional sliding and wedging due to roughness should generally be considered in combination in problems such as this. Specifically, it will be shown that nano-scale roughness has a significant influence on the effective mode II toughness of films whose thickness is in the micron range.

## 2 Formulation of the Model

**2.1 Modeling the Delaminated Interface.** We study steady-state edge delamination of a thin film of thickness  $h$  with shear modulus  $\mu = E/(2(1 + \nu))$  and Poisson's ratio  $\nu$  that is in uniform residual compression and is bonded to a very thick sub-



**Fig. 1 A schematic showing an edge delamination and a buckle delamination and the minimum flaw sizes necessary to achieve steady-state**

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 7, 2000; final revision, March 3, 2001. Associate Editor: H. Gao. Discussion on the paper should be addressed to the 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.

strate having the same elastic properties. Stringfellow and Freund [3] showed that the effect of elastic mismatch between film and substrate is secondary to the role of friction. The same is expected for the roughness effect studied here, and thus an investigation of the role of dissimilar materials is postponed. The delamination is assumed to have propagated away from the edge a distance that is much larger than the film thickness, such that  $h$  is the only relevant length scale. Friction and roughness effects are significant mainly within a few film thicknesses of the interfacial crack tip, so the film-substrate system can be modeled as an elastic half-space with a thin film perfectly bonded along half of its length and fully delaminated along the other half. Under plane strain conditions the problem reduces to a two-dimensional one where the film is in a state of uniform compression far ahead of the interfacial crack tip and is stress free far behind the crack tip.

Using the superposition scheme shown in Fig. 2, the problem of interest (c.f. Fig. 2(a)) can be decomposed into the problem depicted in Fig. 2(b) and the reduced problem shown in Fig. 2(c). Note that there is no displacement of the film relative to the substrate for the problem shown in Fig. 2(b). This can be understood by imagining a film that is under uniform compression and then is debonded along half of its length with a compressive stress applied remotely to hold it in place. Since the displacement is zero, the stress intensity is also zero. Thus the displacement and stress intensity for the problem of interest are identical to those for the reduced problem. It is the reduced problem that is solved in this paper.

Roughness on a scale that is small relative to the film thickness is assumed to be present at the interface between the delaminated film and the substrate. Specifically, the results which emerge from the present study suggest that roughness on the order of one hundredth the film thickness or even somewhat smaller has the largest effect on the apparent energy release rate. The roughness is assumed to be random on the delaminated interface such that once sliding across the interface has occurred on the order of one roughness half-wavelength  $l$ , the two surfaces become uncorrelated and are thereafter propped open a distance  $R$ , the amplitude of the roughness. This is depicted in Fig. 3, although the influence of the two-dimensionality of the roughness is not portrayed.

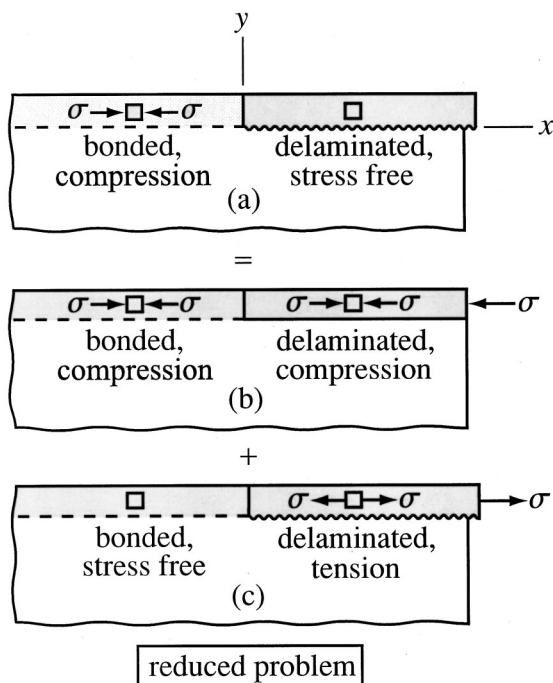


Fig. 2 A schematic of the superposition scheme

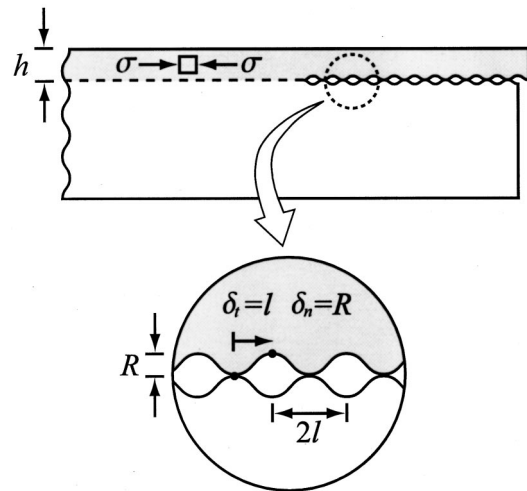


Fig. 3 The coupling of normal and tangential displacements caused by surface roughness

To model the roughness, the relative normal displacement of the two surfaces,  $\delta_n$ , is assumed to be related to the relative tangential displacement of the two surfaces,  $\delta_t$ , by

$$\delta_n(s) = R(1 - e^{-\delta_t(s)/l}), \quad (2)$$

which is plotted in Fig. 4. We retain the assumption of Coulomb friction at the interface such that on a scale that is large relative to  $l$ , but small compared to  $h$

$$\sigma_{xy}(x) = -\mu_f \sigma_{yy}(x) \quad x > 0, \quad (3)$$

when sliding occurs. Thus, the rough interface is replaced by a planar crack where the two components of crack-face displacement are constrained by (2) and the two components of traction are constrained by (3).

Conditions must be imposed to ensure that the solution is consistent with the sign of the friction condition: (i) the normal stress behind the delamination crack tip must be compressive and (ii) the tangential displacement must be a monotonically increasing function of distance behind the tip. The latter of these conditions ensures that the sliding is in one direction under the steady-state propagation.

**2.2 Integral Equation Formulation.** As already noted, the solution to the problem in Fig. 2(b) makes no contribution to the stress intensity factors. The reduced problem shown in Fig. 2(c) is formulated and solved thereby providing the stress intensity fac-

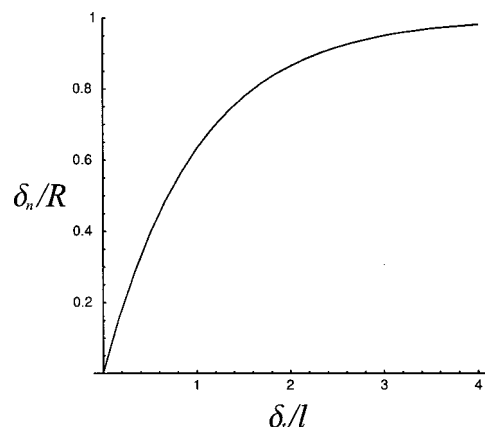


Fig. 4 A plot of the displacement coupling relationship



tors for the problem in Fig. 2(a). The interfacial crack between the film and the substrate is equivalent to a continuous distribution of elastic edge dislocations ([4]). A single elastic edge dislocation a distance  $h$  below the free surface of the half-space (i.e., on the  $x$ -axis of the coordinate system depicted in Figure 2(a)) creates a stress field in the surrounding material that is given by the Airy stress function:

$$\Phi = \frac{\mu}{2\pi(1-\nu)} \left\{ \frac{2b_y x(h-y) + b_x[x^2 + (2h-y)y]}{x^2 + (2h-y)^2} h + \frac{b_y x - b_x y}{2} \log \left[ \frac{x^2 + y^2}{x^2 + (2h-y)^2} \right] \right\}, \quad (4)$$

where  $(b_x, b_y)$  is the Burger's vector for the dislocation (a factor of two missing from the expression for  $\Phi$  given by Stringfellow and Freund has been incorporated here) ([3,5]). The stress field for a single edge dislocation located at a position  $s$  on the  $x$ -axis can be derived from Eq. (4). With

$$\beta_x(s) = \frac{d\delta_t(s)}{ds} \quad (5)$$

$$\beta_y(s) = \frac{d\delta_n(s)}{ds} \quad (6)$$

on the  $x$ -axis a distance  $x$  behind the crack tip, the stresses are given in terms of the dislocation distribution  $(\beta_x(s), \beta_y(s))ds$  by

$$\sigma_{xy}(x) = \frac{\mu}{2\pi(1-\nu)} \left[ \int_0^\infty \frac{g_{11}(\xi)}{\xi} \beta_x(s) ds + \int_0^\infty g_{12}(\xi) \beta_y(s) ds \right], \quad (7)$$

$$\sigma_{yy}(x) = \frac{\mu}{2\pi(1-\nu)} \left[ \int_0^\infty g_{21}(\xi) \beta_x(s) ds + \int_0^\infty \frac{g_{22}(\xi)}{\xi} \beta_y(s) ds \right], \quad (8)$$

where  $\xi = x - s$  and

$$g_{11}(\xi) = \frac{64h^2 + 16h^4\xi^2 + 16h^2\xi^4}{(4h^2 + \xi^2)^3}, \quad (9)$$

$$g_{12}(\xi) = \frac{-32h^5 + 24h^3\xi^2}{(4h^2 + \xi^2)^3}, \quad (10)$$

$$g_{21}(\xi) = \frac{32h^5 - 24h^3\xi^2}{(4h^2 + \xi^2)^3}, \quad (11)$$

$$g_{22}(\xi) = \frac{64h^6 - 48h^4\xi^2}{(4h^2 + \xi^2)^3}. \quad (12)$$

The single governing integral equation expressed with  $\beta_x(s)$  as the unknown is obtained by imposing Eq. (3):

$$\begin{aligned} & \int_0^\infty \frac{g_{11}(\xi)}{\xi} \beta_x(s) ds + \int_0^\infty g_{12}(\xi) \beta_y(s) ds \\ &= -\mu_f \left[ \int_0^\infty g_{21}(\xi) \beta_x(s) ds + \int_0^\infty \frac{g_{22}(\xi)}{\xi} \beta_y(s) ds \right], \end{aligned} \quad (13)$$

where from Eqs. (2) and (6),

$$\beta_y(s) = \frac{R}{l} \beta_x(s) \exp \left[ -\frac{1}{l} \int_0^s \beta_x(\eta) d\eta \right]. \quad (14)$$

It is known from linear elastic fracture mechanics that  $\beta_x(s)$  has a  $s^{-1/2}$  singularity at the crack tip. Far behind the tip the film is in a state of plane strain extension. These conditions are described as

$$\beta_x(s) \propto \frac{1}{\sqrt{s}} \text{ as } s \rightarrow 0, \quad (15)$$

$$\beta_x(s) \rightarrow \frac{\sigma(1-\nu^2)}{E} \text{ as } s \rightarrow \infty. \quad (16)$$

The stress intensity factors are given by

$$K_I = \lim_{s \rightarrow 0} \frac{\sqrt{2\pi s}}{4} \frac{E}{1-\nu^2} \beta_y(s), \quad (17)$$

$$K_{II} = \lim_{s \rightarrow 0} \frac{\sqrt{2\pi s}}{4} \frac{E}{1-\nu^2} \beta_x(s). \quad (18)$$

Note that by Eq. (14) that as  $s \rightarrow 0$ ,

$$\beta_y(s) \rightarrow \frac{R}{l} \beta_x(s), \quad (19)$$

therefore the mode mix is fixed according to

$$K_I = \frac{R}{l} K_{II}. \quad (20)$$

The mode II stress intensity factor for a steady-state edge delamination with zero friction and no roughness is

$$K_{II}^0 = \sigma \sqrt{\frac{h}{2}}, \quad (21)$$

and the energy release rate is given by Eq. (1). Normalizing by these values and applying

$$\mathcal{G} = \frac{1-\nu^2}{E} (K_I^2 + K_{II}^2), \quad (22)$$

yields the following relationships for the energy release rate:

$$\frac{\mathcal{G}}{\mathcal{G}_0} = \left( \frac{K_{II}}{K_{II}^0} \right)^2 \left[ 1 + \left( \frac{R}{l} \right)^2 \right]. \quad (23)$$

The dimensionless equation shown in the Appendix reveals that the solution is determined by three dimensionless parameters, i.e.,

$$\frac{K_{II}}{K_{II}^0} = f \left( \bar{\sigma} = \frac{\sigma h(1-\nu^2)}{El}, \bar{R} = \frac{R}{l}, \mu_f \right). \quad (24)$$

Further details of the formulation, along with aspects of the numerical solution scheme, are presented in the Appendix.

### 3 Results

The combined effect of friction and roughness on the mode II stress intensity factor is shown in Fig. 5. There, plots of  $K_{II}/K_{II}^0$  as a function of  $\bar{\sigma}$  are presented for various roughness levels and two values of the coefficient of friction. When friction is present without any roughness, the normalized stress intensity factor is independent of the film stress. Its reduction below the frictionless limit ( $K_{II}/K_{II}^0 = 1$ ) is relatively small and is precisely in accord with the results of Stringfellow and Freund [3]. Combined friction and roughness lead to a dependence of the stress intensity factor on  $\bar{\sigma}$  such that the full effect of the roughness is attained when  $\bar{\sigma} \approx 2$ . For  $\bar{\sigma} \geq 2$ , the effect of roughness on  $K_{II}$  is pronounced. Values of  $R/l$  as large as unity cannot be ruled out and, indeed, are to be expected when one or both of the materials are polycrystalline.

The effect of friction and roughness on the normalized energy release rate from Eq. (23) is plotted in Fig. 6 for the same two values of  $\mu_f$ . The competition between the reduction in  $K_{II}$  due to combined friction and wedging and the increase in  $K_I$  due to wedging (c.f. Eq. (14)) is evident. At sufficiently small  $\bar{\sigma}$ ,  $\mathcal{G}/\mathcal{G}_0$  is increased above the zero roughness limit due to the dominant influence of wedging on  $K_I$ . For  $\mu_f = .5$ ,  $\mathcal{G}/\mathcal{G}_0$  exceeds the zero

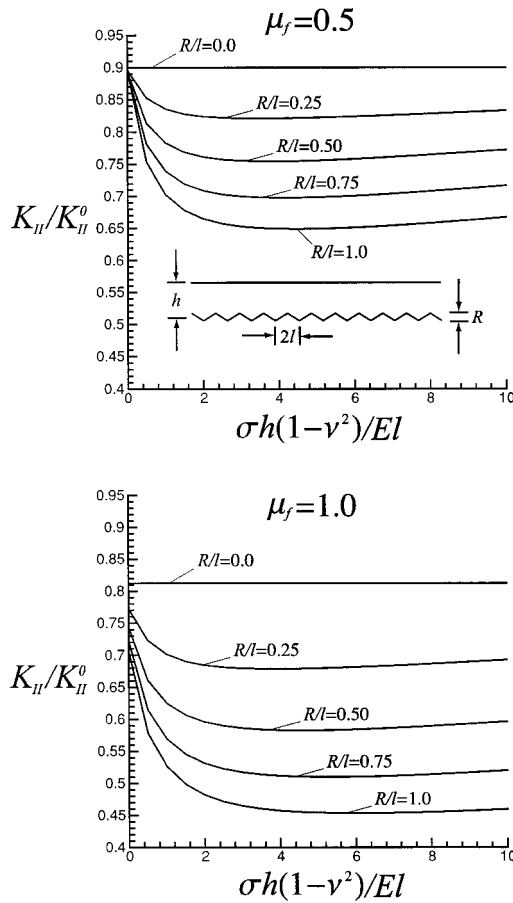


Fig. 5 Normalized mode II stress intensity factor

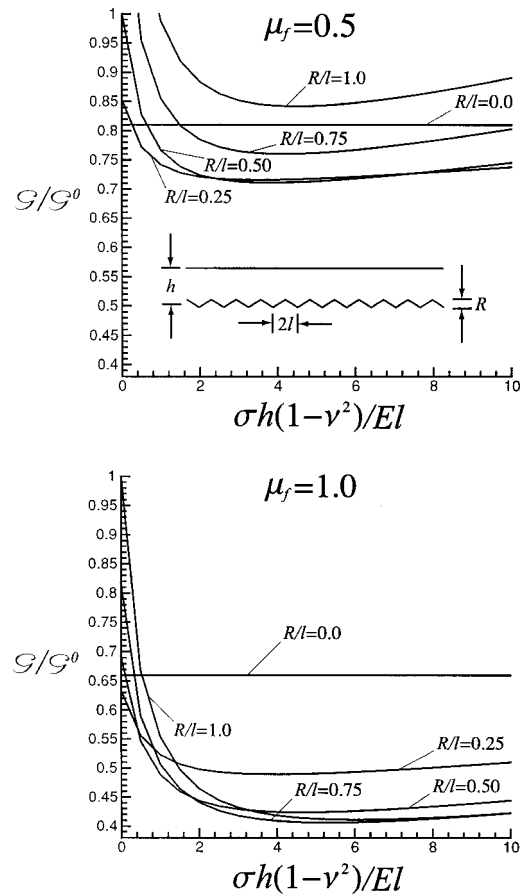


Fig. 6 Normalized energy release rate

roughness limit for all  $\bar{\sigma}$  if  $R/l=1$ . However, it is apparent that there is a significant range of  $R/l$  and  $\bar{\sigma}$  such that the combined effect of friction and roughness reduces  $G/G_0$  by approximately twice the effect of friction alone. For  $\mu_f=.5$ , the maximum crack-tip shielding corresponds to  $G/G_0 \approx .7$  for a roughness level  $R/l \approx .5$  and  $\bar{\sigma} > 2$ . For  $\mu_f=1$ , the corresponding value is  $G/G_0 \approx .4$ .

Examples of the normal stress distribution acting on the interface behind the crack tip are given in Fig. 7. Roughness increases the normal stress within a distance of about one film thickness from the tip. It is this increase which provides greater frictional dissipation and thereby diminishes the mode II stress intensity factor. The normal stress becomes very slightly negative along part of the interface at  $x/h > 5$ , but its magnitude is so small that there is no need to extend the formulation to account for a segment of the interface that is open. This condition has been checked for the full range of parameters governing the solution. In addition, the monotonicity condition for the tangential crack displacement, which is necessary for consistency of the imposed friction condition, was satisfied.

The reduction in the energy release rate (c.f. Fig. 6) gives further insight into the combined effect of friction and interface roughness. The simplest possible condition for crack advance based on crack-tip stress intensity would be the mode-independent criterion

$$\mathcal{G} = \Gamma_0 \quad (25)$$

where  $\Gamma_0$  is considered as the separation energy for the interface. Then, the apparent mode II toughness for steady-state propagation would be  $\Gamma \equiv \mathcal{G}_0$  since  $\mathcal{G}_0$  is the overall energy release rate. Using the results of Fig. 6, one can plot the normalized apparent mode II toughness,  $\Gamma/\Gamma_0$ , as is done in Fig. 8 for  $R/l=.5$ .

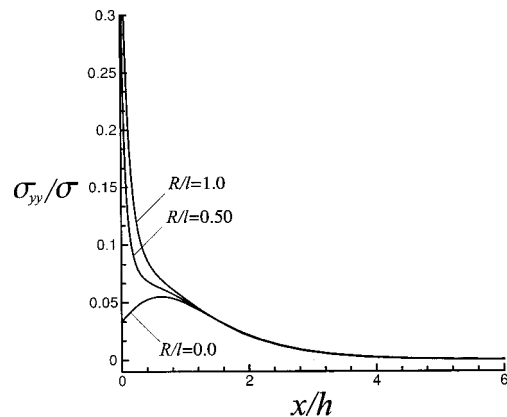
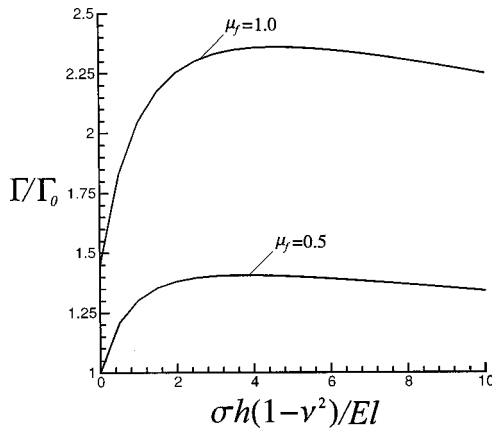


Fig. 7 Normal stress at the interface for  $\mu_f=1.0$  and  $\bar{\sigma}=5.0$

#### 4 Concluding Remarks

Results of this study indicate that the mechanism of combined frictional sliding and roughness-wedging has an effect on the apparent mode II fracture toughness that can be as much as twice the effect of friction alone. Toughness values as large as 2.5 times the separation energy of the interface were predicted. The stress intensity at the crack tip was found to decrease significantly with an increase in the amplitude of the roughness. It was also found that the presence of roughness at the interface induces a mode I stress intensity that is at most equal to the mode II stress intensity for the



**Fig. 8** Normalized apparent mode II fracture toughness for  $\bar{R} = 0.5$

range of roughness parameter values considered. Thus, what is generally considered a pure mode II phenomenon is in fact mixed mode at the tip when the interface is rough.

In addition to the roughness parameter, the analysis revealed that the normalized energy release rate depends on one important dimensionless stress parameter,  $\bar{\sigma} = \sigma h(1 - \nu^2)/El$ . Consider a film whose thickness is about one micron and supports a stress to modulus level of about .01. A roughness half-wavelength,  $l$ , on the order of ten nanometers would produce a value of  $\bar{\sigma}$  large enough to significantly affect the apparent interface toughness. Such a wavelength is fully consistent with the underlying assumptions of the model and the analysis.

Previous models that considered only frictional sliding showed that material mismatch can further increase the apparent mode II fracture toughness when the film is more compliant than the substrate ([3]). It is expected that the same trend would be observed when roughness effects are considered in combination with frictional sliding for the bimaterial problem, yielding apparent mode II fracture toughness values even larger than those predicted by this study. For an interface with an array of contacting asperities, a more realistic representation of the coefficient of friction can be expressed as the sum of a constant term and a term that is proportional to the dilatancy of the interface ([6]). For the roughness model assumed in this paper, the dilatancy term is positive. Thus incorporating a more realistic friction model would result in additional frictional dissipation and a corresponding increase in the apparent mode II toughness. The results of this model in conjunction with these additional mechanisms that enhance the toughness shows that it is indeed possible to predict mode II toughness values that begin to become comparable to those found in experiments by modeling the combined effect of frictional sliding and roughness-wedging at the interface. Furthermore, the strong coupling between frictional sliding and roughness-wedging shows that both effects should be considered in combination when modeling the delamination of compressed thin films.

## Acknowledgment

This work was supported in part by a grant entitled "Prime Reliant Coatings" from the Office of Naval Research, by the National Science Foundation (CMS-96-34632), and by the Division of Engineering and Applied Sciences. Daniel S. Balint acknowledges the support of the National Defense Science and Engineering Graduate Fellowship Program.

## Appendix

To facilitate numerical solution, the domain of integration is transformed from  $s, x \in [0, \infty)$  to  $t, z \in [-1, 1]$  by setting

$$s = h \frac{1+t}{1-t}$$

and

$$x = h \frac{1+z}{1-z}.$$

When the displacement coupling relationship given by Eq. (2) is used, Eq. (13) can be expressed in terms of dimensionless variables as

$$\begin{aligned} & \int_{-1}^1 \beta(t) \left\{ \frac{k_{11}(z, t)(1-z)}{(z-t)(1-t)} + 2\mu_f \frac{k_{21}(z, t)}{(1-t)^2} \right\} dt \\ &= -\bar{R} \int_{-1}^1 \beta(t) \exp[-\bar{\sigma} \delta(t)] \left\{ 2 \frac{k_{12}(z, t)}{(1-t)^2} \right. \\ & \quad \left. + \mu_f \frac{k_{22}(z, t)(1-z)}{(z-t)(1-t)} \right\} dt, \end{aligned} \quad (26)$$

for the unknown

$$\beta(t) = \frac{E}{\sigma(1-\nu^2)} \beta_x \left( h \frac{1+t}{1-t} \right), \quad (27)$$

where

$$\bar{R} = \frac{R}{l}, \quad (28)$$

$$\bar{\sigma} = \frac{\sigma h(1-\nu^2)}{El}, \quad (29)$$

$$\delta(t) = \frac{E}{\sigma(1-\nu^2)} \int_{-1}^t \beta_x \left( h \frac{1+\tau}{1-\tau} \right) \frac{2}{(1-\tau)^2} d\tau, \quad (30)$$

$$k_{11}(z, t) = g_{11} \left( h \frac{2(z-t)}{(1-z)(1-t)} \right), \quad (31)$$

$$k_{12}(z, t) = g_{21} \left( h \frac{2(z-t)}{(1-z)(1-t)} \right) h, \quad (32)$$

$$k_{21}(z, t) = g_{12} \left( h \frac{2(z-t)}{(1-z)(1-t)} \right) h, \quad (33)$$

$$k_{22}(z, t) = g_{22} \left( h \frac{2(z-t)}{(1-z)(1-t)} \right). \quad (34)$$

The dimensionless end conditions are expressed as

$$\beta(t) \propto \frac{1}{\sqrt{1+t}} \text{ as } t \rightarrow -1, \quad (35)$$

$$\beta(t) \rightarrow 1 \text{ as } t \rightarrow 1. \quad (36)$$

Any real continuous function defined on the interval  $[-1, 1]$  can be approximated by a finite linear combination of Chebyshev polynomials of the first kind. Thus the dislocation density can be expressed as

$$\beta(t) = \frac{1}{\sqrt{1+t}} \left[ \sqrt{2} + (1-t) \sum_{n=1}^N a_n T_{n-1}(t) \right], \quad (37)$$

where the integer  $N$  is adjusted to achieve the desired accuracy and the expansion coefficients  $a_n$  are unknown. Both boundary conditions are satisfied by this form. Substituting Eq. (37) into Eq. (26) yields an integral equation involving the  $N$  unknown expansion coefficients that must be satisfied for all  $z \in [-1, 1]$ . Satisfying the integral equation at  $N$  values of  $z$  produces an  $N \times N$  linear system  $\mathbf{Ax} = \mathbf{b}$  where the components of the  $N \times 1$  vector  $\mathbf{x}$  are the expansion coefficients. This linear system can be solved for the

unknown coefficients which can then be used to construct the dislocation density, stress intensity factors, and energy release rate.

The components of the  $N \times N$  matrix  $\mathbf{A}$  and the  $N \times 1$  vector  $\mathbf{b}$  are singular integrals, the integrands of which are expressible in a form that permits numerical evaluation using Gauss-Chebyshev sums for singular integrals derived by Erdogan and Gupta [7]. Convergence must be achieved for the individual Gauss-Chebyshev sums used to evaluate the integrals by choosing the number of terms in each sum,  $M$ , sufficiently large and for the Chebyshev expansion by choosing  $N$  sufficiently large. Numerical error in the linear system increases with  $N$  and for some  $N$  surpasses the accuracy gained. Thus an optimal value of  $N$  exists and was determined.

The Gauss-Chebyshev sums can only be evaluated at the zeros of the  $(M-1)$ th Chebyshev polynomial of the second kind. If  $M/2N$  is a positive integer, the zeros of the  $N$ th Chebyshev polynomial of the first kind are a subset of the zeros of the  $(M-1)$ th Chebyshev polynomial of the second kind. This fact was utilized to evaluate the Gauss-Chebyshev sums at the zeros of the  $N$ th Chebyshev polynomial of the first kind, which was the only

set of  $N$  values of  $z$  for which good convergence was found. Convergence to six decimal places was achieved for  $N=12$  and  $M=48,000$  for the case where  $\mu_f=0$  and  $\bar{R}=0$  which has a known analytical solution. These values of  $N$  and  $M$  were used for all subsequent calculations involving nonzero  $\mu_f$  and  $\bar{R}$ .

## References

- [1] Yu, H., He, M. Y., and Hutchinson, J. W., 2001, "Edge Effects in Thin Films," *Acta Mater.*, in print.
- [2] Hutchinson, J. W., and Suo, Z., 1992, "Mixed Mode Cracking in Layered Materials," *Adv. Appl. Mech.*, **29**, pp. 63–191.
- [3] Stringfellow, R. G., and Freund, L. B., 1993, "The Effect of Interfacial Friction on the Buckle-Driven Spontaneous Delamination of a Compressed Thin Film," *Int. J. Solids Struct.*, **30**, pp. 1379–1395.
- [4] Rice, J. R., 1968, "Mathematical Analysis in the Mechanics of Fracture," *Fracture*, **2**, pp. 191–311.
- [5] Dundurs, J., 1969, "Elastic Interaction of Dislocations With Inhomogeneities," *Mathematical Theory of Dislocations*, ASME, New York, pp. 70–115.
- [6] Scott, R. F., 1963, *Principles of Soil Mechanics*, Addison-Wesley, Reading, MA.
- [7] Erdogan, F., and Gupta, G. D., 1972, "On the Numerical Solution of Singular Integral Equations," *Quart. Appl. Math.*, pp. 525–534.



# A Complex Potential-Variational Method for Stress Analysis of Unsymmetric Laminates With an Elliptical Cutout

E. Madenci

A. Barut

Department of Aerospace and Mechanical  
Engineering,  
The University of Arizona,  
Tucson, AZ 85721

M. P. Nemeth

Mechanics and Durability Branch,  
NASA Langley Research Center,  
Hampton, VA 23681-2199

*A combined complex potential-variational solution method is developed for the analysis of unsymmetrically laminated plates with finite planform geometry, subjected to arbitrary edge loads, and with an inclined elliptical cutout. This method uses complex potentials and their Laurent series expansions to reduce the potential energy of a plate to a contour integral that is evaluated numerically by the trapezoidal rule. A variational statement of equilibrium is applied to the potential energy to obtain a linear system of equations in terms of the unknown coefficients of the Laurent series, whose solutions yield the stress and displacement fields for a given problem. This approach represents a computationally efficient alternative to boundary collocation procedures that are typically used to solve problems based on complex potential theory. Comparisons are made with corresponding results obtained from finite element analysis for a square unsymmetrically laminated plate with a central inclined elliptical cutout and subjected to biaxial tension. The results confirm the validity of the solution method. [DOI: 10.1115/1.1379528]*

## Introduction

Unsymmetrically laminated composite structures are continuing to receive attention in the search for ways to enhance structural performance by exploiting material anisotropy. A fundamental structural element of this class of structures, which is of great practical importance, is the flat laminated plate with one or more cutouts. Structural elements of this type often appear as components of aircraft wing spars and ribs. Behavior trends that show the effects of unsymmetrically laminated construction on this class of structural elements are essentially unknown at the present time. Thus, it is beneficial to have an efficient special-purpose analysis method that can be used to conduct extensive parametric studies in a timely manner and at relatively low costs. Although the finite element method and the boundary element method are capable of determining the response of general plate configurations, like an unsymmetrically laminated plate of general planform shape and with a cutout, they often require extensive computational effort to obtain accurate results. One example that typically manifests this computational difficulty is a plate with a very narrow cutout that is used to simulate a crack.

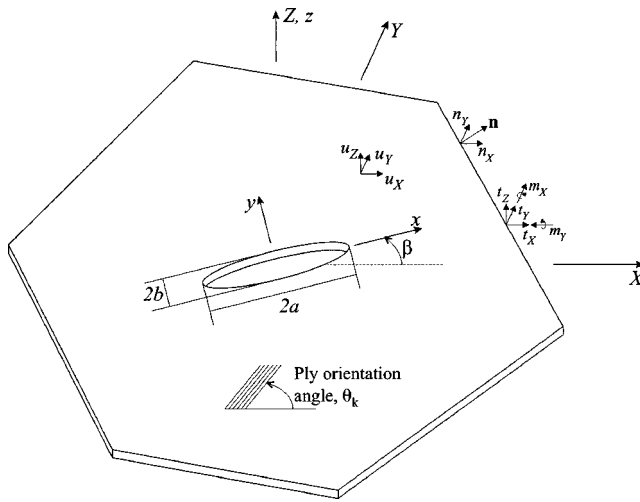
An analytical method that appears to be suitable for inexpensive parametric studies of the structural response of flat unsymmetrical laminated plates with one or more cutouts of general shape and with a general planform shape involves the use of complex potential theory. In the past, complex potential theory has been used successfully for the stress analysis of isotropic plates and symmetrically laminated composite plates with cutouts. Laminated plates of this class exhibit, at most, anisotropy in the form of coupling between in-plane extension or contraction and shear, and between pure bending and twisting deformations. For this class of plates, the in-plane and bending responses are completely uncoupled, can be solved independently, and require only two com-

plex potentials in the analysis. As a result, complex potential theory has been applied to the in-plane and out-of-plane response problems separately. For example, Prasad and Stuart [1] solved the bending problem for a symmetrically laminated plate with an elliptical cutout. This study used a one-term solution to determine the stresses around a cutout in an infinite plate subjected to far-field bending moments. Later, Stuart and Prasad [2] solved the bending problem for a symmetrically laminated plate with finite length and width, with an elliptical cutout, and subjected to edge moments by using their previous formulation and combining a Laurent series approximation for the stresses and displacements with a boundary collocation method. In that study, the boundary collocation technique used a least-squares minimization procedure. Owen and Klang [3] and Britt [4] applied a similar procedure to the in-plane stress analysis of symmetrically laminated composite rectangular plates with finite length and width and subjected to compression, shear, and a combination of these loads.

Complex potential theory has also been combined with other methods to solve problems. For example, Chen [5] formulated a solution for a plate with finite dimensions and a cutout and subjected to bending and twisting moments by combining Laurent series approximations for the complex potentials with an energy method. In this approach, the energy method is used to eliminate the need for a boundary collocation procedure to determine the unknown constants in the Laurent series. Owens [6], Jones [7], Jones and Klang [8], and Britt [9] solved the buckling problem for symmetrically laminated rectangular composite plates with finite length and with a cutout by combining complex potential theory with the Rayleigh-Ritz method. In these studies, the prebuckling stress field was obtained by using Laurent series approximations for the complex potentials and boundary collocation procedure. Then, the two-dimensional stability functional, defined over a doubly connected region, was integrated numerically to form the linear and geometric stiffness matrices of the buckling problem. Qi [10] applied a similar approach to an analysis for the dynamic stability of plates with an elliptical cutout.

Becker [11] was the first to analyze unsymmetrically laminated composite plates, infinite in extent, with a crack-shaped cutout, and subjected to far-field in-plane loading by using complex po-

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, Jan. 11, 2001. Associate Editor: V. R. Kinra. Discussion on the paper should be addressed to the Editor, Prof. 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.



**Fig. 1 Planform geometry, coordinate systems, and loading conditions for laminated plate with an inclined elliptical cutout**

tential theory. In his analysis, Becker introduced four independent complex potentials satisfying the governing equations identically, and used a one-term expansion for each of the four complex potentials. Later, Becker [12] examined the behavior of similar plates subjected to far-field transverse shear loads.

The goal of the present study is to develop a method that is well suited for parametric studies and that accurately predicts the stress and displacement fields in initially flat, unsymmetrically laminated plates, with finite planform dimensions and an elliptical cutout. Toward this goal, the present study combines Becker's complex potential analysis with Chen's variational approach. With this approach, the total potential energy of the plate is expressed solely as a contour integral, and the need for a boundary collocation procedure is eliminated. This approach enables computationally efficient solution of boundary value problems for rectangular and general polygonal plate geometries by simple numerical integration of the potential-energy contour integral. Moreover, a mapping function is used to transform the elliptical cutout boundary to a unit circle in order to further simplify the numerical analysis.

In the remainder of the present paper, details of the analysis method are presented and results from this approach are discussed. First, the boundary value problem is defined. Next, the analysis details and numerical solution procedure are described. Then, results for a square laminate with an inclined elliptical cutout are presented.

## Problem Definition

The general boundary-value problem considered in the present study consists of a thin polygonal plate with an elliptical cutout (Fig. 1). The elliptical cutout, which can be located anywhere inside the exterior plate boundary, has a semi-major axis and a semi-minor axis of length  $a$  and  $b$ , respectively. The special cases of a circular cutout and a line-shaped crack are given by  $a=b$  and  $a=0$  or  $b=0$ , respectively. Two coordinate systems whose origins coincide with the center of the cutout are shown in Fig. 1. The global structural coordinates are given by  $(X, Y, Z)$  and the principal coordinates of the elliptical cutout are given by  $(x, y, z)$ . The orientation of the cutout axes is defined with respect to the global structural coordinate frame as defined by the angle  $\beta$ .

The unsymmetrically laminated plate is assumed to be flat prior to loading, and the exterior edges of the plate are subjected to both external resultant forces and bending moments. The external resultant forces include components  $t_X$ ,  $t_Y$ , and  $t_Z$ , and the bending moments include components  $m_X$  and  $m_Y$ . The external resultant forces and moments are defined with respect to the  $(X, Y, Z)$  global structural coordinates, and their positive-valued directions

are shown in Fig. 1. The global displacement components in the  $X$ ,  $Y$ , and  $Z$ -directions are denoted by  $u_X$ ,  $u_Y$ , and  $u_Z$ , respectively. The unit vector that is normal to the edges of the plate is denoted by  $\mathbf{n}$ , and its components in the  $X$  and  $Y$ -directions are  $n_X$  and  $n_Y$ , respectively. The plate is made of  $K$  specially orthotropic layers, and each layer has an orientation angle,  $\theta_k$ , that is defined with respect to the global  $X$ -axis (Fig. 1). Moreover, each layer has thickness  $t_k$ , elastic moduli  $E_L$  and  $E_T$ , shear modulus  $G_{LT}$ , and Poisson's ratio  $\nu_{LT}$ , where  $L$  and  $T$  are the longitudinal (fiber) and transverse principal material directions of a given ply, respectively.

In the derivation of the equations that govern the plate response, it is convenient to formulate the analysis in terms of the principal coordinates of the elliptical cutout,  $(x, y, z)$ , because of the use of mapping functions. In this case, the displacement components in the  $x$ ,  $y$ , and  $z$ -directions are  $u_x$ ,  $u_y$ , and  $u_z$ , respectively, and  $(t_x, t_y, t_z)$  and  $(m_x, m_y)$  respectively denote the external resultant forces and moments in the  $(x, y, z)$  coordinates. The displacement and the traction components that are defined in the cutout coordinate system are related to those defined with respect to the global structural coordinates by the transformation matrix  $\mathbf{T}$ ,

$$\mathbf{T} = \begin{bmatrix} \cos \beta & \sin \beta \\ -\sin \beta & \cos \beta \end{bmatrix} \quad (1)$$

## Solution Procedure

The goal of this section is to develop an analytical method for determining the displacements and stresses in unsymmetrically laminated composite plates, with finite dimensions and an elliptical cutout, that is amenable to numerical solutions. Toward this goal, the principle of stationary potential energy is used in conjunction with complex potentials that satisfy the three plate equilibrium equations identically. With this approach, the potential energy of a plate is reduced to a contour integral. The complex potentials are expressed as truncated Laurent series that contain unknown complex constants. The boundary conditions are satisfied in an average sense by requiring the first variation of the potential-energy contour integral to vanish, thus resulting in a system of equations to solve for the unknown constants. The details of this solution procedure are as follows and are based on the classical theory of laminated plates, which neglects transverse-shear deformation.

The total potential energy of a laminated plate that is subjected to edge loads is expressed as

$$\pi = \mathcal{U} + \mathcal{V} \quad (2)$$

in which  $\mathcal{U}$  and  $\mathcal{V}$  represent the strain energy in the laminate and the potential energy due to the external edge loads, respectively. The strain energy is expressed as

$$\mathcal{U} = \frac{1}{2} \int_A \mathbf{s}^T \mathbf{e} dA \quad (3)$$

where  $A$  represents the doubly connected planform area (domain) of the laminate shown in Fig. 1. The vectors  $\mathbf{s}$  and  $\mathbf{e}$  represent the stress resultants and strain components as

$$\mathbf{s}^T = \{N_{xx}, N_{yy}, N_{xy}, M_{xx}, M_{yy}, M_{xy}\} \quad (4a)$$

$$\mathbf{e}^T = \{u_{x,x}, u_{y,y}, u_{x,y} + u_{y,x}, -u_{z,xx}, -u_{z,yy}, -2u_{z,xy}\} \quad (4b)$$

where a subscript after a comma denotes differentiation and the functions  $u_x$ ,  $u_y$ , and  $u_z$  represent the in-plane and out-of-plane displacements, respectively, of the plate midplane (Fig. 1). Based on classical laminate theory, the in-plane and bending stress resultants are defined by

$$\{N_{xx}, N_{yy}, N_{xy}\} = \sum_{k=1}^K \{\sigma_{xx}^{(k)}, \sigma_{yy}^{(k)}, \sigma_{xy}^{(k)}\} (z_k - z_{k-1}) \quad (5a)$$

$$\{M_{xx}, M_{yy}, M_{xy}\} = \frac{1}{2} \sum_{k=1}^K \{\sigma_{xx}^{(k)}, \sigma_{yy}^{(k)}, \sigma_{xy}^{(k)}\} (z_k^2 - z_{k-1}^2) \quad (5b)$$

in which  $z_k$  and  $z_{k-1}$  define the distance to the top and bottom surfaces of the  $k$ th layer with respect to the midplane of the laminate.

The stress resultant and strain vectors  $\mathbf{s}$  and  $\mathbf{e}$  are related to each other by

$$\mathbf{s} = \mathbf{C} \mathbf{e} \quad (6)$$

where  $\mathbf{C}$ , defining the laminate material properties, is composed of the extensional, membrane-bending coupling, and bending stiffness matrices,  $\mathbf{A}$ ,  $\mathbf{B}$ , and  $\mathbf{D}$ ,

$$\mathbf{C} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B} & \mathbf{D} \end{bmatrix}. \quad (7)$$

The expressions for the  $\mathbf{A}$ ,  $\mathbf{B}$ , and  $\mathbf{D}$  submatrices can be found in textbooks (e.g., [13]). The desired form of the strain energy is obtained by substituting for  $\mathbf{s}$  and  $\mathbf{e}$  from Eq. (4) and then applying Gauss' theorem. This action yields

$$\begin{aligned} \mathcal{U} = & \frac{1}{2} \int_l \{ (N_{xx} u_x + N_{xy} u_y) n_x + (N_{yy} u_y + N_{xy} u_x) n_y - (M_{xx} u_{z,x} \\ & + M_{xy} u_{z,y}) n_x + (M_{yy} u_{z,y} + M_{xy} u_{z,x}) n_y + (M_{xx,x} + M_{xy,y}) u_z n_x \\ & + (M_{yy,y} + M_{xy,x}) u_z n_y \} dl - \frac{1}{2} \int_A \{ (N_{xx,x} + N_{xy,y}) u_x + (N_{xy,x} \\ & + N_{yy,y}) u_y + (M_{xx,xx} + 2M_{xy,xy} + M_{yy,yy}) u_z \} dA \end{aligned} \quad (8)$$

in which  $l$  represents the boundary of the doubly connected plan-form area  $A$ . The potential energy,  $\mathcal{V}$ , arising from the external resultant forces and moments acting along the exterior boundary is expressed as

$$\mathcal{V} = - \int_l (t_x u_x + t_y u_y + t_z u_z - m_x u_{z,x} - m_y u_{z,y}) dl. \quad (9)$$

In the present study, only the external plate boundary is subjected to external loads. Enforcing the first variation of the total potential energy,  $\delta\pi$ , to vanish results in the well-known equilibrium equations of classical plate theory

$$N_{xx,x} + N_{xy,y} = 0 \quad (10a)$$

$$N_{xy,x} + N_{yy,y} = 0 \quad (10b)$$

$$M_{xx,xx} + 2M_{xy,xy} + M_{yy,yy} = 0. \quad (10c)$$

The next step in the solution procedure is to derive a solution that satisfies these equations identically and reduces the potential energy to a contour integral. As suggested by Becker [11], the solution to these coupled equilibrium equations is established by introducing four linearly independent complex potentials,  $\varphi_k(z_k)$  with  $k = 1, 2, 3, 4$ . These potentials are functions of the modified complex variables,  $z_k$ , defined by

$$z_k = x + \mu_k y \quad \text{with } k = 1, 2, 3, 4 \quad (11)$$

where  $\mu_1, \mu_2, \mu_3$ , and  $\mu_4$  are complex constants that must be determined. First, the stress and moment resultants are expressed as linear combinations of these potentials in the form

$$\begin{aligned} (N_{xx}, N_{yy}, N_{xy}, M_{xx}, M_{yy}, M_{xy}) \\ = 2 \operatorname{Re} \left[ \sum_{k=1}^4 (a_k, b_k, c_k, d_k, e_k, f_k) \varphi_k \right] \quad \text{with } k = 1, 2, 3, 4 \end{aligned} \quad (12)$$

where  $a_k, b_k, c_k, d_k, e_k$ , and  $f_k$  are unknown complex constants. Then, the in-plane displacement components are expressed as linear combinations of complex potentials,  $\Phi_k(z_k)$ , with  $k = 1, 2, 3, 4$ , in the form

$$(u_x, u_y) = 2 \operatorname{Re} \left[ \sum_{k=1}^4 (p_k, q_k) \Phi_k \right] \quad (13)$$

where

$$\Phi_k(z_k) = \int \varphi_k(z_k) dz_k. \quad (14)$$

Similarly, the out-of-plane deflection is expressed as

$$u_z = 2 \operatorname{Re} \left[ \sum_{k=1}^4 F_k(z_k) \right] \quad (15)$$

where

$$F_k(z_k) = \int \Phi_k(z_k) dz_k. \quad (16)$$

By substituting from Eqs. (13) and (15) for the displacement components in Eq. (4b) and using the chain rule for partial differentiation, the strain vector  $\mathbf{e}$  is expressed in terms of the unknown complex constants as

$$\mathbf{e}^T = 2 \operatorname{Re} \left[ \sum_{k=1}^4 \{ p_k, \mu_k q_k, p_k \mu_k + q_k, -1, -\mu_k^2, -2\mu_k \} \varphi_k \right]. \quad (17)$$

Then, the unknown coefficients  $a_k, b_k, c_k, d_k, e_k$ , and  $f_k$  are expressed in terms of the unknown constants  $p_k, q_k$ , and  $\mu_k$  by substituting Eqs. (12) and (17) into the constitutive relations, Eq. (6), as

$$\begin{aligned} \begin{Bmatrix} a_k \\ b_k \\ c_k \\ d_k \\ e_k \\ f_k \end{Bmatrix} &= \begin{bmatrix} A_{11} & A_{12} & A_{16} & B_{11} & B_{12} & B_{16} \\ A_{12} & A_{22} & A_{26} & B_{12} & B_{22} & B_{26} \\ A_{16} & A_{26} & A_{66} & B_{16} & B_{26} & B_{66} \\ B_{11} & B_{12} & B_{16} & D_{11} & D_{12} & D_{16} \\ B_{12} & B_{22} & B_{26} & D_{12} & D_{22} & D_{26} \\ B_{16} & B_{26} & B_{66} & D_{16} & D_{26} & D_{66} \end{bmatrix} \\ &\times \begin{Bmatrix} p_k \\ \mu_k q_k \\ p_k \mu_k + q_k \\ -1 \\ -\mu_k^2 \\ -2\mu_k \end{Bmatrix} \quad \text{with } k = 1, 2, 3, 4. \end{aligned} \quad (18)$$

Equation (18) is then substituted into Eq. (12) to obtain the stress and moment resultants in terms of the remaining unknowns,  $p_k, q_k$ , and  $\mu_k$ . Substituting these expressions for the stress and moment resultants into the equilibrium equations given by Eq. (10) results in a system of algebraic equations given by

$$\begin{aligned} (A_{11} + \mu_k A_{16}) p_k + (A_{12} + \mu_k A_{26}) q_k \mu_k + (A_{16} + \mu_k A_{66}) (p_k \mu_k + q_k) \\ - (B_{11} + \mu_k B_{16}) - (B_{12} + \mu_k B_{26}) \mu_k^2 - (B_{16} + \mu_k B_{26}) 2\mu_k = 0 \end{aligned} \quad (19a)$$

$$\begin{aligned} (A_{16} + \mu_k A_{12}) p_k + (A_{26} + \mu_k A_{22}) q_k \mu_k + (A_{66} + \mu_k A_{26}) (p_k \mu_k + q_k) \\ - (B_{16} + \mu_k B_{12}) - (B_{26} + \mu_k B_{22}) \mu_k^2 - (B_{66} + \mu_k B_{26}) 2\mu_k = 0 \end{aligned} \quad (19b)$$

$$\begin{aligned} (B_{11} + 2\mu_k B_{16} + \mu_k^2 B_{12}) p_k + (B_{12} + 2\mu_k B_{26} + \mu_k^2 B_{22}) q_k \mu_k + (B_{16} \\ + 2\mu_k B_{66} + \mu_k^2 B_{26}) (p_k \mu_k + q_k) - (D_{11} + 2\mu_k D_{16} + \mu_k^2 D_{12}) \\ - (D_{12} + 2\mu_k D_{26} + \mu_k^2 D_{22}) \mu_k^2 - (D_{16} + 2\mu_k D_{66} \\ + \mu_k^2 D_{26}) 2\mu_k = 0. \end{aligned} \quad (19c)$$

These equations are linear in  $p_k$  and  $q_k$  but nonlinear in  $\mu_k$ . Thus, considering the first two of these equations permits the solution for  $p_k$  and  $q_k$  in terms of  $\mu_k$  as

$$p_k = \frac{\sum_{i=0}^5 R_i \mu_k^i}{\sum_{i=0}^5 T_i \mu_k^i}, \quad (20a)$$

$$q_k = \frac{\sum_{i=0}^5 S_i \mu_k^i}{\sum_{i=0}^5 T_i \mu_k^i} \quad (20b)$$

The explicit expressions for  $R_i$ ,  $S_i$ , and  $T_i$  are given in the Appendix. Substituting these expressions for  $p_k$  and  $q_k$  into Eq. (19c) yields the characteristic (polynomial) equation for the generalized complex variables  $z_k$  in terms of  $\mu_k$  as

$$\sum_{i=0}^8 G_i \mu_k^i = 0 \quad (21)$$

The coefficients of this polynomial are real-valued constants leading to distinct roots, and are given in the Appendix. The complex parameters  $\mu_k$  ( $k=1, 2, 3, 4$ ) and their complex conjugates are the roots of this characteristic equation, i.e.,  $\mu_5 = \bar{\mu}_1$ ,  $\mu_6 = \bar{\mu}_2$ ,  $\mu_7 = \bar{\mu}_3$ , and  $\mu_8 = \bar{\mu}_4$ . Determination of these roots leads to the expressions for the stress and moment resultants that satisfy the in-plane and bending equilibrium equations for arbitrary complex potentials, thus rendering the area integrals in Eq. (8) to vanish from the expression for the total potential energy,  $\pi$ . Thus, the boundary value problem has been transformed into a form that involves only a boundary-contour integral and four unknown complex potential functions,  $\Phi_k(z_k)$  with  $k=1, 2, 3, 4$ . The auxiliary unknown complex potential functions,  $\varphi_k(z_k)$  and  $F_k(z_k)$ , are obtained by their differentiation and integration, respectively.

These analytic functions, related to  $\Phi_k(z_k)$ , are assumed in the series form in terms of mapping functions  $\xi_k$  as

$$\Phi_k(\xi_k) = \sum_{n=1}^N [\alpha_{nk} \xi_k^n + \beta_{nk} \xi_k^{-n}]. \quad (22)$$

This form of  $\Phi_k(\xi_k)$  results in the potentials for  $F_k(\xi_k)$  and  $\varphi_k(\xi_k)$  as

$$\begin{aligned} F_k(\xi_k) &= \alpha_{1k} \left( \frac{1}{2} r_k \xi_k^2 - s_k \ln \xi_k \right) + \beta_{1k} \left( \frac{1}{2} s_k \xi_k^{-2} - r_k \ln \xi_k \right) \\ &+ \sum_{n=2}^N \left[ \alpha_{nk} \left( \frac{1}{n+1} r_k \xi_k^{n+1} - \frac{1}{n-1} s_k \xi_k^{n-1} \right) \right. \\ &\left. + \beta_{nk} \left( \frac{1}{1-n} r_k \xi_k^{-n+1} + \frac{1}{n+1} s_k \xi_k^{-n-1} \right) \right] \end{aligned} \quad (23)$$

and

$$\varphi_k(\xi_k) = \sum_{n=1}^N \frac{n}{\omega'_k(\xi_k)} [\alpha_{nk} \xi_k^{n-1} - \beta_{nk} \xi_k^{-n-1}] \quad (24)$$

where a prime denotes differentiation with respect to  $\xi_k$ . The functions  $\omega_k(\xi_k)$  are the inverse mapping of  $\xi_k$ .

The mapping functions,  $\xi_k$  ( $k=1, 2, 3, 4$ ), map an elliptical cutout in the  $(x, y)$ -plane onto unit circles in the four complex planes that correspond to the complex variables,  $z_k$ . In this analysis, the mapping functions for an elliptical cutout, introduced by Lekhnitskii [14], are employed and are given in the form

$$\xi_k = \frac{z_k \pm \sqrt{z_k^2 - a^2 - \mu_k^2 b^2}}{a - i \mu_k b}, \quad k=1,2,3,4 \quad (25)$$

in which  $a$  and  $b$  are the major and minor axes of an ellipse and  $i = \sqrt{-1}$ . The sign of the square root term is chosen so that  $|\xi_k| \geq 1$ ; that is, the domain of the plate is mapped onto the exterior of the unit circles. Inverting the mapping function provides  $\omega_k(\xi_k)$  as

$$z_k = \omega_k(\xi_k) = r_k \xi_k - \frac{s_k}{\xi_k} \quad (26)$$

in which

$$r_k = \frac{1}{2} (a - i \mu_k b), \quad (27a)$$

$$s_k = \frac{1}{2} (a + i \mu_k b). \quad (27b)$$

The unknown complex coefficients  $\alpha_{nk}$  and  $\beta_{nk}$  in Eqs. (22)–(24) are determined by requiring the first variation of the total potential to vanish, i.e.,  $\delta\pi = 0$ . This step yields a system of linear algebraic equations involving boundary integrals that must be solved numerically to obtain  $\alpha_{nk}$  and  $\beta_{nk}$  for plates with finite planform geometry.

## Numerical Procedure

In this section, the matrix form of the potential-energy contour integral is derived to facilitate numerical solutions. In order to apply the matrix operations, the following quantities are defined:

$$\varphi_{nk}^* = \frac{n \xi_k^{n-1}}{\omega'_k(\xi_k)} \quad (28)$$

$$\Phi_{nk}^* = \xi_k^n \quad (29)$$

$$F_{nk}^* = \begin{cases} \frac{r_k}{n+1} \xi_k^{n+1} - \frac{s_k}{n-1} \xi_k^{n-1}, & |n| \geq 2 \\ \frac{r_k \xi_k^2}{2} - s_k \ln \xi_k, & n = 1 \\ r_k \ln \xi_k + \frac{s_k \xi_k^{-2}}{2}, & n = -1 \end{cases} \quad (30)$$

such that

$$\Phi_k = \sum_{n=1}^N (\alpha_{nk} \Phi_{nk}^* + \beta_{nk} \Phi_{-nk}^*) \quad (31a)$$

$$\varphi_k = \sum_{n=1}^N (a_{nk} \varphi_{nk}^* + \beta_{nk} \varphi_{-nk}^*) \quad (31b)$$

$$F_k = \sum_{n=1}^N (a_{nk} F_{nk}^* + \beta_{nk} F_{-nk}^*). \quad (31c)$$

Considering first the expression for  $N_{xx}$  in Eq. (12) and substituting from Eq. (31b), the following term can be rewritten as

$$\begin{aligned} \sum_{k=1}^4 a_k \varphi_k &= \{^{(1)} \mathbf{N}_{xx}^* T \boldsymbol{\alpha}_1 + ^{(-1)} \mathbf{N}_{xx}^* T \boldsymbol{\beta}_1 + \dots + ^{(N)} \mathbf{N}_{xx}^* T \boldsymbol{\alpha}_N \\ &+ ^{(-N)} \mathbf{N}_{xx}^* T \boldsymbol{\beta}_N \} \end{aligned} \quad (32)$$

where

$$^{(n)} \mathbf{N}_{xx}^* T = \{a_1 \varphi_{n1}^*, a_2 \varphi_{n2}^*, a_3 \varphi_{n3}^*, a_4 \varphi_{n4}^*\} \quad (33)$$

$$\boldsymbol{\alpha}_n^T = \{\alpha_{n1}, \alpha_{n2}, \alpha_{n3}, \alpha_{n4}\} \quad (34a)$$

$$\boldsymbol{\beta}_n^T = \{\beta_{n1}, \beta_{n2}, \beta_{n3}, \beta_{n4}\} \quad (34b)$$



and the superscript  $T$  denotes matrix transposition. The same term can be rewritten as

$$\sum_{k=1}^4 a_k \varphi_k = \mathbf{N}_{xx}^{*T} \mathbf{q} \quad (35)$$

in which the vectors  $\mathbf{N}_{xx}^*$  and  $\mathbf{q}$  are formed as

$$\mathbf{N}_{xx}^{*T} = \{ {}^{(1)}\mathbf{N}_{xx}^{*T}, {}^{(-1)}\mathbf{N}_{xx}^{*T}, {}^{(2)}\mathbf{N}_{xx}^{*T}, {}^{(-2)}\mathbf{N}_{xx}^{*T}, \dots, {}^{(N)}\mathbf{N}_{xx}^{*T}, {}^{(-N)}\mathbf{N}_{xx}^{*T} \} \quad (36)$$

$$\mathbf{q}^T = \{ \alpha_1^T, \beta_1^T, \alpha_2^T, \beta_2^T, \dots, \alpha_N^T, \beta_N^T \}. \quad (37)$$

With these equations, the stress resultant  $N_{xx}$  takes its final form as

$$N_{xx} = 2 \operatorname{Re} \{ \mathbf{N}_{xx}^{*T} \mathbf{q} \} \quad (38a)$$

or

$$N_{xx} = \mathbf{N}_{xx}^{*T} \mathbf{q} + \bar{\mathbf{N}}_{xx}^{*T} \bar{\mathbf{q}} \quad (38b)$$

or

$$N_{xx} = \hat{\mathbf{N}}_{xx}^T \hat{\mathbf{q}} \quad (39)$$

where the vectors  $\hat{\mathbf{N}}_{xx}$  and  $\hat{\mathbf{q}}$  are defined as

$$\hat{\mathbf{N}}_{xx}^T = \{ \mathbf{N}_{xx}^{*T}, \bar{\mathbf{N}}_{xx}^{*T} \} \quad (40)$$

and

$$\hat{\mathbf{q}}^T = \{ \mathbf{q}^T, \bar{\mathbf{q}}^T \} \quad (41)$$

in which the overbars denote complex conjugation. The remaining stress resultants and displacement components can be obtained by replacing the symbol  $N_{xx}$  with the appropriate one in Eq. (39) and by using one of the following corresponding expressions:

$${}^{(n)}\mathbf{N}_{yy}^{*T} = \{ b_1 \varphi_{n1}^*, b_2 \varphi_{n2}^*, b_3 \varphi_{n3}^*, b_4 \varphi_{n4}^* \} \quad (42a)$$

$${}^{(n)}\mathbf{N}_{xy}^{*T} = \{ c_1 \varphi_{n1}^*, c_2 \varphi_{n2}^*, c_3 \varphi_{n3}^*, c_4 \varphi_{n4}^* \} \quad (42b)$$

$${}^{(n)}\mathbf{M}_{xx}^{*T} = \{ d_1 \varphi_{n1}^*, d_2 \varphi_{n2}^*, d_3 \varphi_{n3}^*, d_4 \varphi_{n4}^* \} \quad (42c)$$

$${}^{(n)}\mathbf{M}_{yy}^{*T} = \{ e_1 \varphi_{n1}^*, e_2 \varphi_{n2}^*, e_3 \varphi_{n3}^*, e_4 \varphi_{n4}^* \} \quad (42d)$$

$${}^{(n)}\mathbf{M}_{xy}^{*T} = \{ f_1 \varphi_{n1}^*, f_2 \varphi_{n2}^*, f_3 \varphi_{n3}^*, f_4 \varphi_{n4}^* \} \quad (42e)$$

$${}^{(n)}\mathbf{u}_x^{*T} = \{ p_1 \Phi_{n1}^*, p_2 \Phi_{n2}^*, p_3 \Phi_{n3}^*, p_4 \Phi_{n4}^* \} \quad (42f)$$

$${}^{(n)}\mathbf{u}_y^{*T} = \{ q_1 \Phi_{n1}^*, q_2 \Phi_{n2}^*, q_3 \Phi_{n3}^*, q_4 \Phi_{n4}^* \} \quad (42g)$$

$${}^{(n)}\mathbf{u}_z^{*T} = \{ F_{n1}^*, F_{n2}^*, F_{n3}^*, F_{n4}^* \}. \quad (42h)$$

The terms  $M_{xx,x}$ ,  $M_{xy,x}$ ,  $M_{xy,y}$ ,  $M_{yy,y}$ ,  $u_{z,x}$ , and  $u_{z,y}$  that appear in the potential-energy contour integral are obtained by replacing the symbol  $N_{xx}$  with the appropriate one in Eq. (39) and by using one of the following corresponding expressions:

$${}^{(n)}\mathbf{M}_{xx,x}^{*T} = \left\{ \frac{d_1}{\omega_1'} \varphi_{n1}'^*, \frac{d_2}{\omega_2'} \varphi_{n2}'^*, \frac{d_3}{\omega_3'} \varphi_{n3}'^*, \frac{d_4}{\omega_4'} \varphi_{n4}'^* \right\} \quad (43a)$$

$${}^{(n)}\mathbf{M}_{xy,x}^{*T} = \left\{ \frac{f_1}{\omega_1'} \varphi_{n1}'^*, \frac{f_2}{\omega_2'} \varphi_{n2}'^*, \frac{f_3}{\omega_3'} \varphi_{n3}'^*, \frac{f_4}{\omega_4'} \varphi_{n4}'^* \right\} \quad (43b)$$

$${}^{(n)}\mathbf{M}_{xy,y}^{*T} = \left\{ \frac{f_1 \mu_1}{\omega_1'} \varphi_{n1}'^*, \frac{f_2 \mu_2}{\omega_2'} \varphi_{n2}'^*, \frac{f_3 \mu_3}{\omega_3'} \varphi_{n3}'^*, \frac{f_4 \mu_4}{\omega_4'} \varphi_{n4}'^* \right\} \quad (43c)$$

$${}^{(n)}\mathbf{M}_{yy,y}^{*T} = \left\{ \frac{e_1 \mu_1}{\omega_1'} \varphi_{n1}'^*, \frac{e_2 \mu_2}{\omega_2'} \varphi_{n2}'^*, \frac{e_3 \mu_3}{\omega_3'} \varphi_{n3}'^*, \frac{e_4 \mu_4}{\omega_4'} \varphi_{n4}'^* \right\} \quad (43d)$$

$${}^{(n)}\mathbf{u}_{z,x}^{*T} = \{ \Phi_{n1}^*, \Phi_{n2}^*, \Phi_{n3}^*, \Phi_{n4}^* \} \quad (43e)$$

$${}^{(n)}\mathbf{u}_{z,y}^{*T} = \{ \mu_1 \Phi_{n1}^*, \mu_2 \Phi_{n2}^*, \mu_3 \Phi_{n3}^*, \mu_4 \Phi_{n4}^* \} \quad (43f)$$

in which prime denotes differentiation with respect to its argument. By using Eq. (39) and the analogous expressions for the other stress resultants and displacement components, the potential-energy contour integral is expressed as

$$\pi = \hat{\mathbf{q}}^T \left\{ \frac{1}{2} \oint_l \hat{\mathbf{S}} \hat{\mathbf{Q}}^T dl \right\} \hat{\mathbf{q}} - \hat{\mathbf{q}}^T \oint_l \hat{\mathbf{Q}} \mathbf{p} dl \quad (44)$$

where  $\hat{\mathbf{S}}$  and  $\hat{\mathbf{Q}}$  are matrices and  $\mathbf{p}$  is a vector given by

$$\hat{\mathbf{S}} = [ (\hat{\mathbf{N}}_{xx} n_x + \hat{\mathbf{N}}_{xy} n_y) (\hat{\mathbf{N}}_{xy} n_x + \hat{\mathbf{N}}_{yy} n_y) (\hat{\mathbf{M}}_{xx} n_x + \hat{\mathbf{M}}_{xy} n_y) + \hat{\mathbf{M}}_{xy} n_x + \hat{\mathbf{M}}_{yy} n_y) (-\hat{\mathbf{M}}_{xx} n_x - \hat{\mathbf{M}}_{xy} n_y) (-\hat{\mathbf{M}}_{xy} n_x - \hat{\mathbf{M}}_{yy} n_y) ] \quad (45a)$$

$$\hat{\mathbf{Q}} = [ \hat{\mathbf{u}}_x \quad \hat{\mathbf{u}}_y \quad \hat{\mathbf{u}}_z \quad -\hat{\mathbf{u}}_{z,x} \quad -\hat{\mathbf{u}}_{z,y} ] \quad (45b)$$

$$\mathbf{p}^T = \{ t_{xx}, t_{yy}, t_{zz}, m_{xx}, m_{yy} \}. \quad (45c)$$

Requiring the first variation of the total potential to vanish provides the equilibrium equations in the form

$$\mathbf{K} \hat{\mathbf{q}} = \mathbf{F} \quad (46)$$

in which  $\hat{\mathbf{q}}$  involves only the unknown coefficients  $\alpha_{nk}$  and  $\beta_{nk}$  and their complex conjugates. The known coefficient matrix  $\mathbf{K}$  and forcing vector  $\mathbf{F}$  in Eq. (46) are given by

$$\mathbf{K} = \frac{1}{2} \oint_l (\hat{\mathbf{S}} \hat{\mathbf{Q}}^T + \hat{\mathbf{Q}} \hat{\mathbf{S}}^T) dl \quad (47a)$$

$$\mathbf{F} = \oint_l \hat{\mathbf{Q}} \mathbf{p} dl. \quad (47b)$$

Also, a constraint equation

$$\operatorname{Re} \left[ \sum_{k=1}^4 \frac{\mu_k p_k - q_k}{a - i \mu_k b} \alpha_{1k} \right] = 0 \quad (48)$$

is used to eliminate the rigid-body motion of the plate from Eq. (46) by requiring the constant term in the rotation field, given by

$$\Omega = u_{x,y} - u_{y,x} \quad (49)$$

to vanish. In addition, the two constraint equations

$$\operatorname{Im} \left[ \sum_{k=1}^4 s_k \alpha_{1k} \right] = 0, \quad (50a)$$

$$\operatorname{Im} \left[ \sum_{k=1}^4 r_k \beta_{1k} \right] = 0 \quad (50b)$$

are used to ensure the single-valuedness of the logarithmic terms in Eq. (23) that result in single-valuedness of the out-of-plane displacement field.

These three constraint equations are directly invoked into the governing equations by using Lagrange multipliers. The remaining step needed to solve Eq. (46) is to compute the contour integrals, which include the exterior plate boundary and the interior cutout boundary, in Eqs. (47). In the present study, the contour integrals are evaluated by using a trapezoidal rule of the form

$$\begin{aligned} \int_0^L f[x(l), y(l)] dl &= \frac{L}{2} \int_{-1}^1 f[x(\eta), y(\eta)] d\eta \\ &\cong \frac{L}{2} \sum_{k=0}^K H_k f[x(\eta_k), y(\eta_k)] \end{aligned} \quad (51)$$

where

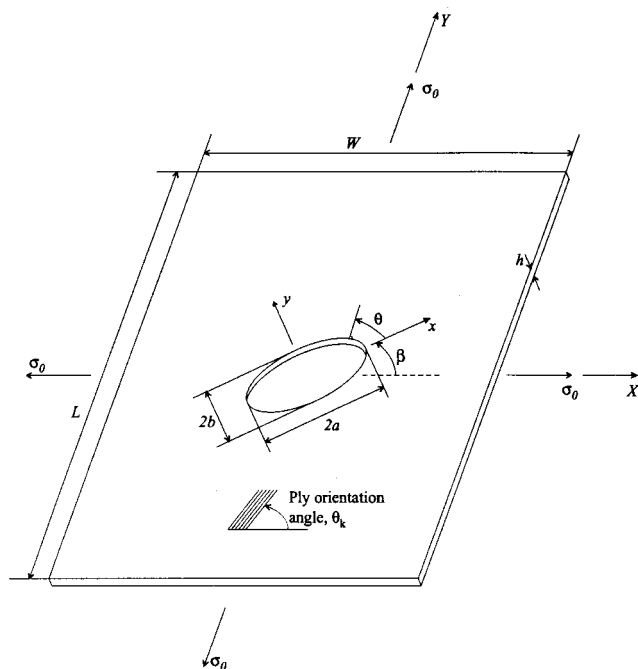


Fig. 2 Rectangular plate with an inclined elliptical cutout and subjected to biaxial tension

$$L = \int dl, \quad \eta = \frac{2l}{L} - 1$$

$$H_k = \begin{cases} \frac{1}{2} & \text{for } k=0 \text{ and } K \\ 1 & \text{for } 2 < k < K \end{cases}, \quad \eta_k = \frac{2k}{K} - 1, \quad (52)$$

with  $K$  denoting the number of intervals on the contour  $l$ . In this equation,  $x(\eta)$  and  $y(\eta)$  are the parametric equations of the contour  $l$  in which  $-1 < \eta < 1$ . The function  $f$  is arbitrary in that it represents symbolically the integrand of any integral used in the construction of Eq. (47).

## Numerical Results

Results are presented in this section for a square  $[\pm 45 \text{ deg}]$  unsymmetrically laminated plate ( $W=L=4$  inches) that is subjected to a biaxial load of  $\sigma_0=1$  lb/in. and that has a relatively small centrally located elliptical cutout (Fig. 2). The elliptical cutout is inclined at an angle of  $\beta=45 \text{ deg}$  and has a semi-major axis,  $a=0.5$  inch, and semi-minor axis,  $b=0.2$  inch. These dimensions correspond to a cutout aspect ratio  $b/a=0.4$ , a semi-major axis-to-plate-width ratio  $a/W=0.125$ , and a plate width-to-thickness ratio  $L/h=400$ . Each layer of the two-layer plate is 0.005 inch thick and is specially orthotropic with material properties  $E_L=19 \times 10^6$  psi,  $E_T=1.9 \times 10^6$  psi,  $G_{LT}=0.9 \times 10^6$  psi, and  $\nu_{LT}=0.3$ . The total plate thickness is given by  $h=0.01$  inch.

The roots of the characteristic equation, Eq. (21), for this problem were determined to be  $\mu_1=-0.54867+0.83604i$ ,  $\mu_2=3.38364i$ ,  $\mu_3=0.29554i$ ,  $\mu_4=0.54867+0.83604i$ ,  $\mu_5=\bar{\mu}_1$ ,  $\mu_6=\bar{\mu}_2$ ,  $\mu_7=\bar{\mu}_3$ , and  $\mu_8=\bar{\mu}_4$ . The deformed shape of the laminate is shown in Fig. 3 and is essentially a saddle-like shape because of the bending-stretching coupling. Contour plots of the nondimensional in-plane shear stress resultant,  $N_{xy}/\sigma_0$ , and the nondimensional twisting stress resultant,  $M_{xy}/(\sigma_0 h)$ , defined in terms of the global coordinate system (Fig. 1) are shown in Figs. 4 and 5, respectively. These results indicate a central point of inversion symmetry (polar symmetry) at the origin and concentra-

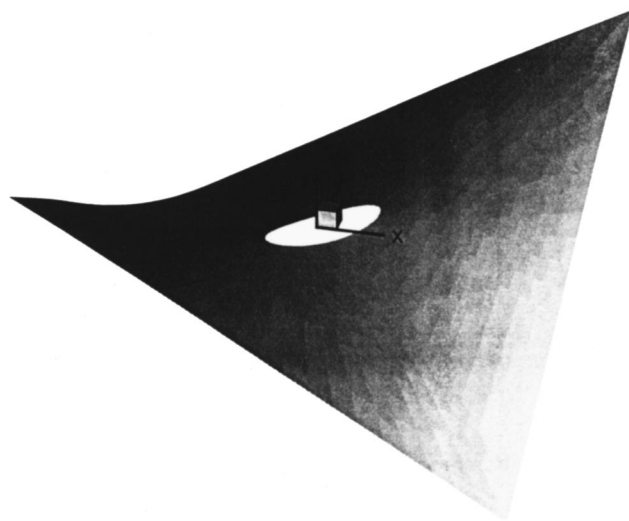


Fig. 3 Deformed geometry of a square  $[\pm 45 \text{ deg}]$  laminated plate with an elliptical cutout inclined at  $45 \text{ deg}$  and subjected to biaxial tension

tions in the stress resultants at  $\theta=0 \text{ deg}$  and  $180 \text{ deg}$ . The maximum values of  $N_{xy}/\sigma_0$  and  $M_{xy}/(\sigma_0 h)$  are  $-4.96$  and  $.74$ , respectively.

The fidelity of the present method for this particular set of problem parameters is illustrated in Fig. 6 by a plot of the out-of-plane displacement at the cutout edge versus the angle  $\theta$ . In particular, five curves are shown in the figure that correspond to values of  $N=1, 3, 5, 10$ , and  $15$ , in which  $N$  is the number of terms in the Laurent series defined by Eq. (31). The values of  $N=1, 3, 5, 10$ , and  $15$ , correspond to the 16, 48, 80, 160, and 260 equations that are used to obtain solutions to Eq. (46). An additional curve is shown that corresponds to results that were ob-

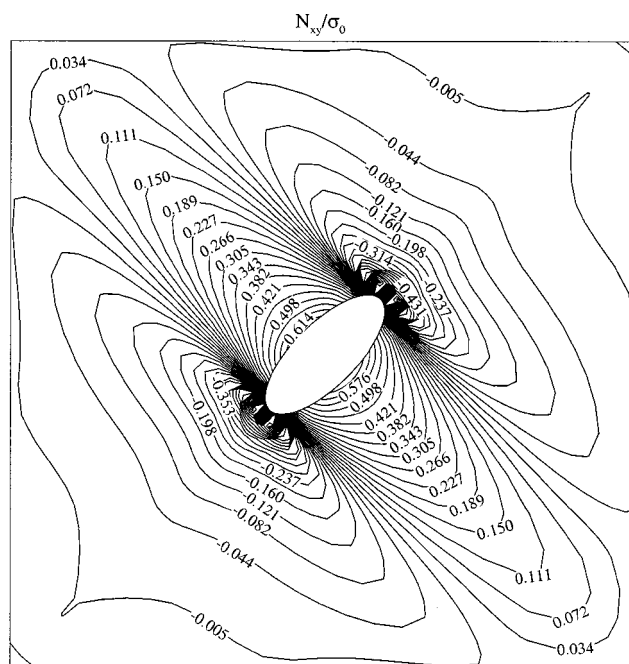
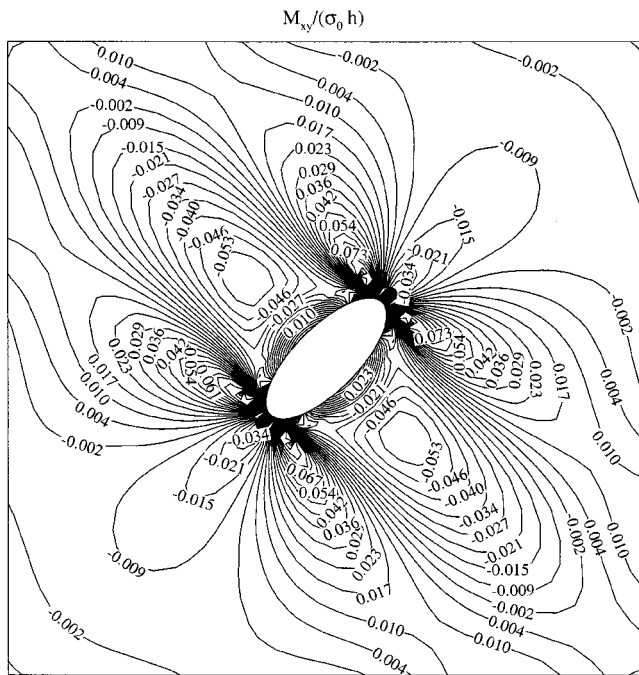


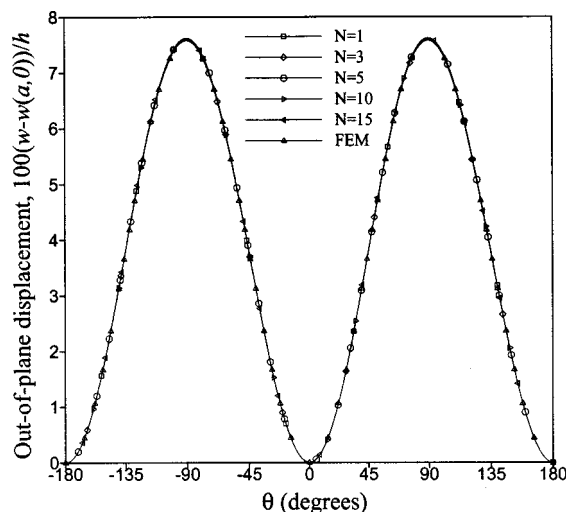
Fig. 4 Nondimensional in-plane shear stress resultant distribution in a square  $[\pm 45 \text{ deg}]$  laminated plate with an elliptical cutout inclined at  $45 \text{ deg}$  and subjected to biaxial tension



**Fig. 5 Nondimensional twisting stress resultant distribution in a square  $[\pm 45 \text{ deg}]$  laminated plate with an elliptical cutout inclined at 45 deg and subjected to biaxial tension**

tained by the finite element analysis. The finite element model for one quarter of the plate had 3150 elements with 1656 nodes resulting in 9720 degrees-of-freedom. Comparison of the curves shown in Fig. 6, which are coincident, indicates remarkable agreement, even for  $N=1$ .

Results are also presented in Table 1 that give the stress resultants  $N_{xx}$  and  $M_{xx}$  at the edge of the cutout and at  $\theta=90 \text{ deg}$ . Similarly, results are presented in Table 2 that give  $N_{yy}$  and  $M_{yy}$  at the edge of the cutout and at  $\theta=0 \text{ deg}$ . In both of these tables, the stress resultants are referred to the local  $(x, y)$  coordinate system (Fig. 2) and results are given that were obtained from a finite element analysis and from the present analysis with  $N=1, 3, 5, 10$ , and 15 in Eq. (31). The finite element solutions were obtained by using a constant-strain, shear deformable shell element de-



**Fig. 6 Nondimensional out-of-plane displacement around the edge of an elliptical cutout inclined at 45 deg and subjected to biaxial tension**

**Table 1 Comparison of stress resultants at the cutout edge given by  $\theta=90 \text{ deg}$  for a square  $[\pm 45 \text{ deg}]$  laminated plate with an elliptical cutout inclined at 45 deg and subjected to biaxial tension**

	$N_{xx}$ (lb/in.)	$M_{xx}$ (lb-in./in.)
$N=1$	1.3875	0.0117
$N=3$	1.3330	0.0106
$N=5$	1.3050	0.0100
$N=10$	1.3060	0.0100
$N=15$	1.3058	0.0100
FEM	1.3276	0.0099

scribed by Barut et al. [15]. The results in Table 1 indicate differences of less than five percent for  $N_{xx}$  and 18 percent for  $M_{xx}$ , for  $N=1$ . For  $N \geq 5$ , all differences are less than 1.5 percent for  $N_{yy}$  and 2 percent for  $M_{yy}$ . This good agreement with the shear-deformable finite element results at these two plate locations and the fact that the present method is based on classical plate theory suggest that transverse shear deformations are negligible, as expected, for a very thin plate ( $L/h=400$ ).

Results are presented in Figs. 7 and 8 that show the nondimensional out-of-plane displacement  $w/h$  and the nondimensional bending moment resultant  $M_{\theta\theta}/(\sigma_0 h)$ , respectively, along the edge of the cutout. For these results, all the problem parameters are identical to those described in the beginning of this section except for the cutout inclination angle  $\beta$ . Seven curves are shown in Figs. 7 and 8 that correspond to values of  $\beta = -45 \text{ deg}, -30 \text{ deg}, -15 \text{ deg}, 0 \text{ deg}, 15 \text{ deg}, 30 \text{ deg},$  and  $45 \text{ deg}$ . Location around the edge of the cutout is indicated by values of the angle  $\theta$  (Fig. 2).

The results in Fig. 7 for the out-of-plane displacement field show a family of oscillatory curves that exhibit a phase shift and growth in amplitude as the cutout inclination angle  $\beta$  deviates from 0 deg. The maximum amplitude of the displacement along the cutout is predicted for  $\beta = \pm 45 \text{ deg}$ . Each of these curves exhibits two upward peaks and two downward peaks that shift farther away from  $\theta=45 \text{ deg}$  as the magnitude of the cutout inclination angle increases.

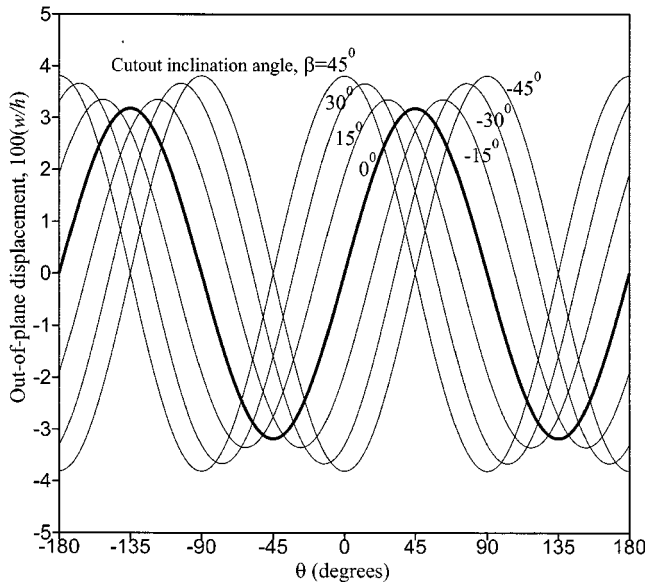
The results in Fig. 8 for the bending stress resultant also show a family of oscillatory curves that exhibit a growth in amplitude as the cutout inclination angle  $\beta$  deviates from 0 deg. Consistent with the results for the out-of-plane displacements, the results in Fig. 8 also show the maximum amplitude for  $\beta = \pm 45 \text{ deg}$ . Furthermore, the results in Fig. 8 also show that the curves for  $\beta \neq 0 \text{ deg}$  have two positive and two negative peak values that generally have different magnitudes. In contrast, the curve for  $\beta = 0 \text{ deg}$  has two positive and two negative peak values that have the same magnitude.

## Conclusions

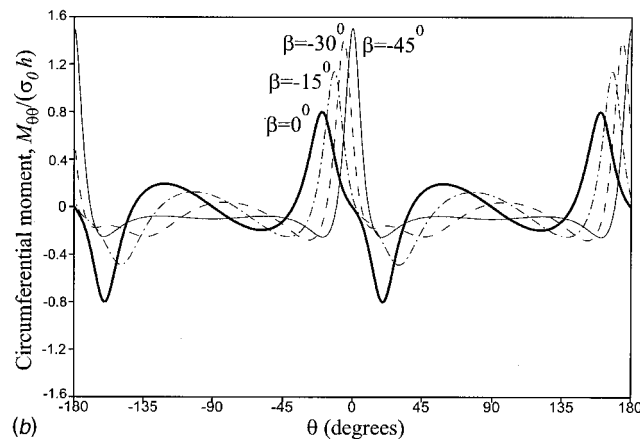
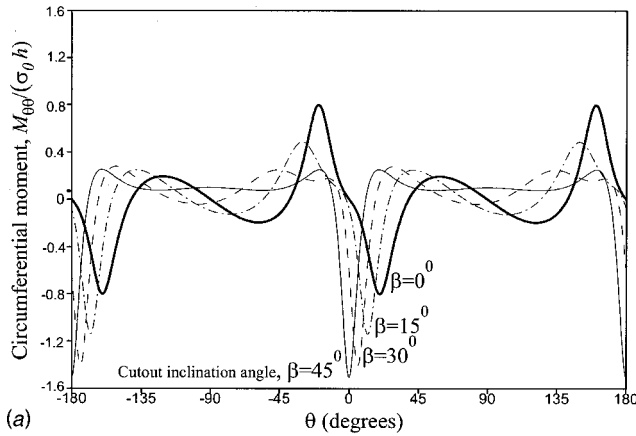
A solution method for the analysis of unsymmetrically laminated plates with finite planform geometry and with an elliptical cutout has been presented. This method uses complex potentials

**Table 2 Comparison of stress resultants at the cutout edge given by  $\theta=0 \text{ deg}$  for a square  $[\pm 45 \text{ deg}]$  laminated plate with an elliptical cutout inclined at 45 deg and subjected to biaxial tension**

	$N_{yy}$ (lb/in.)	$M_{yy}$ (lb-in./in.)
$N=1$	10.0342	-0.0144
$N=3$	10.2424	-0.0148
$N=5$	10.3622	-0.0150
$N=10$	10.3723	-0.0150
$N=15$	10.3808	-0.0150
FEM	10.2400	-0.0150



**Fig. 7 Effect of elliptical cutout inclination on the nondimensional out-of-plane displacement around the edge of the cutout for a square  $[\pm 45^\circ]$  laminated plate subjected to biaxial tension**



**Fig. 8 Effect of elliptical cutout inclination on the nondimensional bending stress resultant around the edge of the cutout for a square  $[\pm 45^\circ]$  laminated plate subjected to biaxial tension**

and their Laurent series expansions to reduce the potential energy of a plate to a contour integral that is evaluated numerically by the trapezoidal rule. In addition, the variational statement of equilibrium is applied to the potential energy to obtain a linear system of equations in terms of the unknown coefficients of the Laurent series. Solution of these equations yields the stress and displacement fields for a given problem. This approach represents a computationally efficient alternative to boundary collocation procedures that are typically used to solve problems of the type presented herein.

Results have been presented for a square two-layer unsymmetrically laminated plate that is subjected to biaxial tension on the edges and that has a relatively small, inclined central elliptical cutout. The results were compared to corresponding results obtained from finite element analysis and show excellent agreement. The results also show that the amount of cutout inclination affects greatly the out-of-plane displacements and bending stress resultant at the cutout boundary.

## Appendix

In Eq. (19), the constant coefficients  $R_k$ ,  $S_k$ , and  $T_k$  are defined as

$$R_0 = A_{66}B_{11} - A_{16}B_{16}$$

$$R_1 = 2A_{66}B_{16} + 2A_{26}B_{11} - A_{16}B_{12} - 2A_{16}B_{66} - A_{12}B_{16}$$

$$R_2 = A_{22}B_{11} + 5A_{26}B_{16} - 3A_{16}B_{26} - A_{12}B_{12} - 2A_{12}B_{66}$$

$$R_3 = -2A_{66}B_{26} + A_{26}B_{12} + 2A_{26}B_{66} + 3A_{22}B_{16} - A_{16}B_{22} - 3A_{12}B_{26}$$

$$R_4 = -A_{26}B_{26} + A_{22}B_{12} + 2A_{22}B_{66} - A_{12}B_{22} - A_{66}B_{22}$$

$$R_5 = A_{22}B_{26} - A_{26}B_{22}$$

$$S_0 = A_{11}B_{16} - A_{16}B_{11}$$

$$S_1 = A_{11}B_{12} + 2A_{11}B_{66} - A_{16}B_{16} - A_{12}B_{11} - A_{66}B_{11}$$

$$S_2 = 3A_{11}B_{26} + A_{16}B_{12} + 2A_{16}B_{66} - 2A_{66}B_{16} - 3A_{12}B_{16} - A_{26}B_{11}$$

$$S_3 = A_{11}B_{22} + 5A_{16}B_{26} - A_{12}B_{12} - 2A_{12}B_{66} - 3A_{26}B_{16}$$

$$S_4 = 2A_{16}B_{22} + 2A_{66}B_{26} - A_{12}B_{26} - A_{26}B_{12} - 2A_{26}B_{66}$$

$$S_5 = A_{66}B_{22} - A_{26}B_{26}$$

$$T_0 = A_{11}A_{66} - A_{16}^2$$

$$T_1 = 2A_{11}A_{26} - 2A_{12}A_{16}$$

$$T_2 = 2A_{16}A_{26} + A_{11}A_{22} - A_{12}^2 - 2A_{12}A_{66}$$

$$T_3 = 2A_{16}A_{22} - 2A_{12}A_{26}$$

$$T_4 = A_{22}A_{66} - A_{26}^2$$

The coefficients  $G_k$  of the characteristic equation given in Eq. (20) are defined in the form

$$G_0 = S_5B_{22} + R_5B_{26} - T_4D_{22}$$

$$G_1 = R_5B_{12} + S_4B_{22} + R_4B_{26} + 3S_5B_{26} + 2R_5B_{66} - T_3D_{22} - 4S_4D_{26}$$

$$G_2 = R_4B_{12} + S_5B_{12} + 3R_5B_{16} + S_3B_{22} + R_3B_{26} + 3S_4B_{26} + 2R_4B_{66} + 2S_5B_{66} - 2T_4D_{12} - T_2D_{22} - 4T_3D_{26} - 4T_4D_{66}$$

$$G_3 = R_5B_{11} + R_3B_{12} + S_4B_{12} + 3R_4B_{16} + S_5B_{16} + S_2B_{22} + R_2B_{26} + 3S_3B_{26} + 2R_3B_{66} + 2S_4B_{66} - 2T_3D_{12} - 4T_4D_{16} - T_1D_{22} - 4T_2D_{26} - 4T_3D_{66}$$

$$G_4 = R_4B_{11} + R_2B_{12} + S_3B_{12} + 3R_3B_{16} + S_4B_{16} + S_1B_{22} + R_1B_{26} + 3S_2B_{26} + 2R_2B_{66} + 2S_3B_{66} - T_4D_{11} - 2T_2D_{12} - 4T_3D_{16} - T_0D_{22} - 4T_1D_{26} - 4T_2D_{66}$$



$$\begin{aligned}
G_5 &= R_3 B_{11} + R_1 B_{12} + S_2 B_{12} + 3R_2 B_{16} + S_3 B_{16} + S_0 B_{22} + R_0 B_{26} \\
&\quad + 3S_1 B_{26} + 2R_1 B_{66} + 2S_2 B_{66} - T_3 D_{11} - 2T_1 D_{12} - 4T_2 D_{16} \\
&\quad - 4T_0 D_{26} - 4T_1 D_{66} \\
G_6 &= R_2 B_{11} + R_0 B_{12} + S_1 B_{12} + 3R_1 B_{16} + S_2 B_{16} + 3S_0 B_{26} + 2R_0 B_{66} \\
&\quad + 2S_1 B_{66} - T_2 D_{11} - 2T_0 D_{12} - 4T_1 D_{16} - 4T_0 D_{66} \\
G_7 &= R_1 B_{11} + S_0 B_{12} + 3R_0 B_{16} + S_1 B_{16} + 2S_0 B_{66} - T_1 D_{11} - 4T_0 D_{16} \\
G_8 &= R_0 B_{11} + S_0 B_{16} - T_0 D_{11} .
\end{aligned}$$

## References

- [1] Prasad, C. B., and Stuart, M. S., 1990, "Moment Distributions Around Holes in Symmetric Composite Laminates Subjected to Bending Moments," *AIAA J.*, **28**, pp. 877–882.
- [2] Stuart, M. J., and Prasad, C. B., 1990, "Analysis and Experiments for Composite Laminates With Holes and Subjected to 4-Point Bending," *Proceedings, 31st AIAA/ASME/ASCE/AHS Structures, Structural Dynamics, and Materials Conference*, Part 2, Vol. 4, AIAA, New York, pp. 748–758.
- [3] Owen, V. L., and Klang, E. C., 1990, "Shear Buckling of Specially Orthotropic Plates With Centrally Located Cutouts," *Proceedings, 8th DOD/NASA/FAA Conference on Fibrous Composites in Structural Design*, Norfolk, VA, WL/FIBA, Wright-Patterson AFB, OH.
- [4] Britt, V. O., 1991, "Analysis of Stresses in Finite Anisotropic Panels with Centrally Located Cutouts," *Proceedings, 9th DOD/NASA/FAA Conference on Fibrous Composites in Structural Design*, Lake Tahoe, NV, WL/FIBA, Wright-Patterson AFB, OH.
- [5] Chen, H. C., 1997, "Bending of Laminated Composite Plates With Cutouts," *Proc. 38th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, Part 3, Vol. 3, Kissimmee, FL, AIAA, New York, pp. 1884–1892.
- [6] Owens, V. L., 1990, "Shear Buckling of Anisotropic Plates With Centrally Located Cutouts," M.S. thesis, North Carolina State University, Raleigh, NC.
- [7] Jones, K. M., 1992, "Buckling Analysis of Fully Anisotropic Plates Containing Cutouts and Elastically Restrained Edges," M.S. thesis, North Carolina State University, Raleigh, NC.
- [8] Jones, K. M., and Klang, E. C., 1992, "Buckling Analysis of Fully Anisotropic Plates Containing Cutouts and Elastically Restrained Edges," *Proc. 33rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, Dallas, TX, AIAA, New York, pp. 190–200.
- [9] Britt, V. O., 1994, "Shear and Compression Buckling Analysis for Anisotropic Panels with Elliptical Cutouts," *AIAA J.*, **32**, pp. 2293–2299.
- [10] Qi, Z., 1998, "Dynamical Stability of Finite Anisotropic Panels with Elliptical Cutouts," *J. Shanghai Univ.*, **2**, pp. 35–39.
- [11] Becker, W., 1991, "A complex Potential Method for Plate Problems With Bending Extension Coupling," *Arch. Appl. Mech.*, **61**, pp. 318–326.
- [12] Becker, W., 1992, "Closed-Form Analytical Solutions for a Griffith Crack in a Non-Symmetric Laminate Plate," *Composite Struc.*, **21**, pp. 49–55.
- [13] Jones, R. M., 1999, *Mechanics of Composite Materials*, 2nd Ed., Taylor and Francis, Philadelphia, PA.
- [14] Lekhnitskii, S. G., 1968, *Anisotropic Plates*, Gordon and Breach, New York.
- [15] Barut, A., Madenci, E., and Tessler, A., 1997, "Nonlinear Analysis of Laminates Through a Mindlin-Type Shear Deformable Shallow Shell Element," *Comput. Methods Appl. Mech. Eng.*, **143**, pp. 155–173.

**Q. Yang**

Associate Professor  
e-mail: yangq@mail.tsinghua.edu.cn

**W. Y. Zhou**

Professor  
e-mail: zhouwy@mail.tsinghua.edu.cn

Department of Hydraulic Engineering,  
Tsinghua University,  
100084 Beijing, P. R. China

**G. Swoboda**

Professor,  
Faculty of Civil Engineering and Architecture,  
University of Innsbruck,  
Technikerstr. 13  
A-6020 Innsbruck, Austria  
e-mail: gunter.swoboda@uibk.ac.at

# Asymptotic Solutions of Penny-Shaped Inhomogeneities in Global Eshelby's Tensor

*In this paper, a three-dimensional penny-shaped isotropic inhomogeneity surrounded by unbounded isotropic matrix in a uniform stress field is studied based on Eshelby's equivalent inclusion method. The solution including the deduced equivalent eigenstrain and its asymptotic expressions is presented in tensorial form. The so-called energy-based equivalent inclusion method is introduced to remove the singularities of the size and eigenstrain of the Eshelby's equivalent inclusion of the penny-shaped inhomogeneity, and yield the same energy disturbance. The size of the energy-based equivalent inclusion can be used as a generic damage measurement. [DOI: 10.1115/1.1380676]*

## 1 Introduction

In this paper, a three-dimensional penny-shaped isotropic inhomogeneity surrounded by unbounded isotropic matrix in a uniform stress field is studied based on Eshelby's equivalent inclusion method ([1]). The penny-shaped inhomogeneity is treated as the limit of a spheroidal shape inhomogeneity. An ellipsoidal shape inhomogeneity, as shown in Fig. 1, in a uniform stress field is a classic problem ([1,2]) briefed in the Appendix. In particular, in plane elasticity, the penny-shaped inhomogeneities are studied by Hurtado et al. [3], and asymptotic solutions for lamellar inhomogeneities of arbitrary shape have been obtained by Homocntcov-schi and Dasalu [4].

The Eshelby's tensor is usually defined by each component in a certain local coordinate system, see, e.g., Eq. (A7) in Appendix A, and so is the solution based on Eshelby's equivalent inclusion method. Even in the well-chosen local coordinate system, the solutions are very complicated; see, e.g., the work of Hurtado et al. [3], Zhao and Weng [5], Shafiro and Kachanov [6], let alone in the global coordinate system. It is hard to see the influence of Eshelby's tensor as an entity from such solutions. If we are mainly concerned about the macroscopic responses of inhomogeneities rather than their local stress fields in detail, e.g., damage modeling, such solutions are generally too complicated to be used.

A typical microstructure-based damage variable is the crack tensor; see, e.g., Kachanov [7], Swoboda and Yang [8,9], and Yang et al. [10]. For a solid of size  $V^0$  weakened by  $n$  cracks, the crack tensor takes the form

$$\Omega = \frac{1}{V^0} \sum_{k=1}^n a_k^3 \mathbf{n}^k \mathbf{n}^k \quad (1)$$

where  $a_k$  and  $\mathbf{n}^k$  are the radius and normal vector of the  $k$ th crack. Equation (1) shows that only few characteristic geometric quantities are incorporated into the definition of the damage tensor. Based on such a definition, any crack of arbitrary shape can be

replaced with an equivalent penny-shaped crack. Evidently, the crack tensor is just the generalized form of the Budiansky and O'Connell's damage parameter ([11])

$$\omega = \frac{1}{V^0} \sum_{k=1}^n a_k^3 \quad (2)$$

which is associated with the energy disturbance caused by the presence of the cracks; see, e.g., Eq. (85). In this paper, we focus on the penny-shaped inhomogeneities and the corresponding energy disturbances, and in particular their macroscopic representation. Note that the penny-shaped crack or anticrack is just a special case of the penny-shaped inhomogeneity, as discussed in Section 4. Thus, the penny-shaped inhomogeneity furnishes a much more general microdefect model.

In this paper, asymptotic developments for small thickness ratio  $\xi$  is used to the Eshelby's tensor  $\mathcal{S}$  of spheroidal inclusions, i.e.,  $\mathcal{S} = \mathbf{S} + \mathcal{M}\xi + \mathbf{O}(\xi^2)$ . The tensor  $\mathbf{S}$  and  $\mathcal{M}$  are expressed in global coordinate systems. Furthermore, the Eshelby's equivalent eigenstrains of penny-shaped inhomogeneities are expressed explicitly by  $\mathcal{S}$ ,  $\mathbf{S}$ , and  $\mathcal{M}$  as tensorial entities.

The energy disturbance caused by a penny-shaped inhomogeneity is  $U^* = V/2\sigma^0 : \epsilon^*$  where  $\sigma^0$  is the farfield stress;  $V$  and  $\epsilon^*$  are the size and eigenstrain of the Eshelby's equivalent inclusion of the inhomogeneity, respectively. The inhomogeneity is geometrically identical to its Eshelby's equivalent inclusion. Since  $V \rightarrow 0$  due to  $\xi \rightarrow 0$ , for a macroscopically sensible penny-shaped inhomogeneity, the equivalent eigenstrain  $\epsilon^*$  should be singular, otherwise its energy disturbance is trivial and the inhomogeneity is macroscopically insensible. The asymptotic expressions of the singular equivalent eigenstrain  $\epsilon^*$  have been obtained in this paper, and the singular factors  $\eta$  contained in the singular  $\epsilon^*$  have been identified.

If moving the singular factor  $\eta$  from  $\epsilon^*$  to  $V$ , the equivalent size is  $V^{eq} = V\eta$  and the equivalent eigenstrain is  $\epsilon_{eq}^* = \epsilon^*/\eta$ . Both the equivalent size  $V^{eq}$  and eigenstrain  $\epsilon_{eq}^*$  are finite quantities, and the energy disturbance can be rewritten as  $U^* = V^{eq}/2\sigma^0 : \epsilon_{eq}^*$ . In Section 3.2, the so-called *energy-based inclusion method* is proposed. Unlike the Eshelby's equivalent inclusion  $\Omega$ , the energy-based inclusion  $\Omega^{eq}$  may be geometrically different with the corresponding inhomogeneity. For a penny-shaped inhomogeneity, the energy-based equivalent inclusion is of the size  $V^{eq}$  and prescribed by the energy-based equivalent strain  $\epsilon_{eq}^*$ . The energy-based equivalent inclusion and the Eshelby's

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 18, 2000; final revision, Jan. 8, 2001. Associate Editor: D. Kouris. Discussion on the paper should be addressed to the 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.

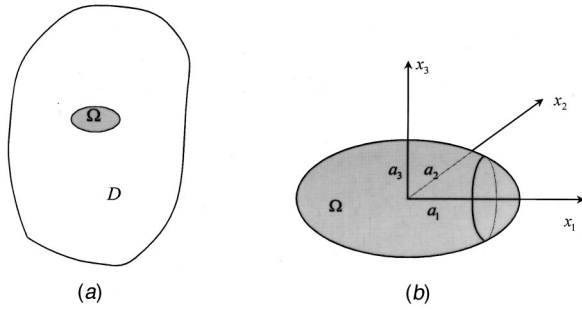


Fig. 1 (a) Inclusion or inhomogeneity  $\Omega$ ; (b) an ellipsoidal inclusion with principal half-axes  $a_1$ ,  $a_2$ , and  $a_3$

equivalent inclusion (prescribed by  $\epsilon^*$ ) cause different local stress fields but the same energy disturbance. It is pointed out in Section 8 that the size  $V^{eq}$  of the energy-based equivalent inclusion furnishes a generic damage measurement for different inhomogeneities.

## 2 Global Eshelby's Tensor

The Eshelby's tensor of a spheroidal inclusion ( $a_1 = a_2 = a$ ) takes the form  $S_{ijkl} = S_{ijkl}(\nu, \xi)$  where  $\xi = a_3/a$  as defined in Eqs. (A3), (A4), and (A5) in Appendix A, and is denoted symbolically as  $\mathcal{S}$ . For small thickness ratio  $\xi$ , the  $\mathcal{S}$  tensor can be expanded into power series with respect to  $\xi$  about  $\xi=0$

$$\mathcal{S} = \mathcal{S}(\nu, \xi) = \mathcal{S}|_{\xi=0} + \frac{\partial \mathcal{S}}{\partial \xi} \Big|_{\xi=0} \xi + \mathcal{O}(\xi^2) = \mathcal{S} + \mathcal{M}\xi + \mathcal{O}(\xi^2) \quad (3)$$

where

$$\mathcal{S} = \mathcal{S}|_{\xi=0}, \quad \mathcal{M} = \frac{\partial \mathcal{S}}{\partial \xi} \Big|_{\xi=0}. \quad (4)$$

The asymptotic expression

$$\mathcal{S} \rightarrow \mathcal{S} + \mathcal{M}\xi \quad (5)$$

is a good approximation of the  $\mathcal{S}$  tensor for  $\xi \rightarrow 0$  or penny-shaped inclusions.

**2.1 Properties of the Global  $\mathcal{S}$  Tensor.** Obviously, the  $\mathcal{S}$  tensor is just the limit of the Eshelby's tensor defined in Eq. (A8). Now we transform the  $\mathcal{S}$  tensor from the local coordinate system to a global coordinate system. Consider a local coordinate system  $x_1', x_2', x_3'$  in the global coordinate system  $x_1, x_2, x_3$  as shown in Fig. 2. The direction cosines of axes  $x_1', x_2', x_3'$  in global system are  $n_i^{(1)}$ ,  $n_i^{(2)}$ , and  $n_i^{(3)}$ , respectively, and the principal axes of a penny-shaped ellipsoidal inclusion coincide with the axes  $x_1', x_2', x_3'$ , as shown in Fig. 2. If the components of the Eshelby's tensor in the local coordinate system is  $S'_{ijkl}$ , its components in the global coordinate system  $S_{ijkl}$  can be determined by the transformation

$$S_{ijkl} = n_i^{(m)} n_j^{(n)} n_k^{(p)} n_l^{(q)} S'_{mnpq}. \quad (6)$$

Note that  $n_i^{(1)} n_j^{(1)} + n_i^{(2)} n_j^{(2)} + n_i^{(3)} n_j^{(3)} = \delta_{ij}$  and set

$$n_i^{(3)} = n_i = \mathbf{n}. \quad (7)$$

Then the  $\mathcal{S}$  tensor defined by Eq. (A8) in Appendix A can be expressed in the global coordinate system

$$\begin{aligned} \mathcal{S} = S_{ijkl} = & \frac{1}{2} (\delta_{ik} n_j n_l + \delta_{il} n_j n_k + \delta_{jk} n_i n_l + \delta_{jl} n_i n_k - 4 n_i n_j n_k n_l) \\ & + \frac{\nu}{1-\nu} n_i n_j (\delta_{kl} - n_k n_l) + n_i n_j n_k n_l \end{aligned} \quad (8)$$

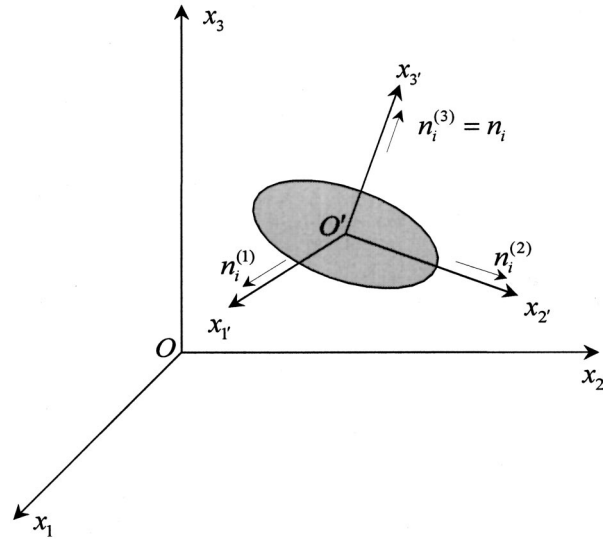


Fig. 2 Global and local coordinate system

see also Yang [12]. Evidently,  $S_{ijkl}$  reduces to Eq. (A8) if  $n_i = \{0, 0, 1\}$ . Equation (8) implies that  $\mathcal{S} = \mathcal{S}(\nu, \mathbf{n})$ . This is a very special property since in general,  $\mathcal{S} = \mathcal{S}[\nu, \mathbf{n}, \mathbf{n}^{(1)}, \mathbf{n}^{(2)}]$ . Here, we introduce the identity tensors of rank four and two

$$\mathbf{I} = I_{ijkl} = \frac{1}{2} (\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}), \quad \mathbf{I} = I_{ij} = \delta_{ij}, \quad (9)$$

and the  $\mathbf{N}$  and  $\mathbf{T}$  tensors

$$\begin{aligned} \mathbf{N} = \mathbf{nnnn}, \quad \mathbf{T} = T_{ijkl} = & \frac{1}{4} (n_i n_k \delta_{jl} + n_i n_l \delta_{jk} + n_j n_k \delta_{il} + n_j n_l \delta_{ik} \\ & - 4 n_i n_j n_k n_l). \end{aligned} \quad (10)$$

The  $\mathbf{N}$  and  $\mathbf{T}$  tensors are associated with the normal component  $\sigma$  and shear component  $\tau$  of the stress tensor  $\boldsymbol{\sigma}$  in the direction of  $\mathbf{n}$ , i.e.,

$$\sigma^2 = \boldsymbol{\sigma} : \mathbf{N} : \boldsymbol{\sigma}, \quad \tau^2 = \boldsymbol{\sigma} : \mathbf{T} : \boldsymbol{\sigma}. \quad (11)$$

Then the  $\mathcal{S}$  tensor can be expressed in terms of the  $\mathbf{N}$  and  $\mathbf{T}$  tensors

$$\mathcal{S} = 2\mathbf{T} + \frac{\nu}{1-\nu} \mathbf{nnl} + \frac{1-2\nu}{1-\nu} \mathbf{N}. \quad (12)$$

The transpose of  $S_{ijkl}$  for  $ij$  and  $kl$  is denoted  $\mathcal{S}^T = S_{ijkl}^T = S_{klij}$ ,

$$\mathcal{S}^T = 2\mathbf{T} + \frac{\nu}{1-\nu} \mathbf{l} \mathbf{nn} + \frac{1-2\nu}{1-\nu} \mathbf{N}. \quad (13)$$

The  $\mathcal{S}$  tensor has some interesting properties, e.g.,

$$S_{ijmn} S_{mnkl} = S_{ijkl} \quad \text{or} \quad \mathcal{S} : \mathcal{S} = \mathcal{S}, \quad (14)$$

and

$$(\mathcal{S} - \mathbf{I}) : \mathbf{n} = \mathbf{0}. \quad (15)$$

An important theorem is

$$(\mathbf{I} - \theta \mathcal{S})^{-1} = \mathbf{I} + \frac{\theta}{1-\theta} \mathcal{S} \quad (16)$$

for  $\theta \neq 1$ , since

$$(\mathbf{I} - \theta \mathcal{S}) : \left( \mathbf{I} + \frac{\theta}{1-\theta} \mathcal{S} \right) \equiv \left( \mathbf{I} + \frac{\theta}{1-\theta} \mathcal{S} \right) : (\mathbf{I} - \theta \mathcal{S}) \equiv \mathbf{I} \quad (17)$$

except  $\theta = 1$ . The  $\mathcal{S}$  tensor can be split into two parts:

$$\mathbf{S} = \mathbf{S}_n + \mathbf{S}_t, \quad \mathbf{S}_n = \frac{\nu}{1-\nu} \mathbf{nnl} + \frac{1-2\nu}{1-\nu} \mathbf{N}, \quad \mathbf{S}_t = 2\mathbf{T}. \quad (18)$$

It is easy to verify

$$\mathbf{S}_n : \mathbf{S}_t = \mathbf{S}_t : \mathbf{S}_n = 0, \quad \mathbf{S}_n : \mathbf{S}_n = \mathbf{S}_n, \quad \mathbf{S}_t : \mathbf{S}_t = \mathbf{S}_t. \quad (19)$$

Split the  $\mathbf{S}_n$  tensor into the unsymmetric and symmetric parts

$$\mathbf{S}_n = \mathbf{S}_n^{uns} + \mathbf{S}_n^{sym}, \quad \mathbf{S}_n^{uns} = \frac{\nu}{1-\nu} \mathbf{nnl}, \quad \mathbf{S}_n^{sym} = \frac{1-2\nu}{1-\nu} \mathbf{nnnn}. \quad (20)$$

Then, the following relations can be verified

$$\mathbf{S}_t : \mathbf{C} = \frac{1}{2\mu} \mathbf{S}_t, \quad \mathbf{S}_n : \mathbf{C} = \frac{1}{2\mu} \mathbf{S}_n^{sym} \quad (21)$$

where  $\mathbf{C}$  and  $\mu$  are the elastic compliance tensor and shear modulus, respectively,

$$\mathbf{C} = -\frac{\nu}{E} \mathbf{II} + \frac{1+\nu}{E} \mathbf{I}, \quad \mu = \frac{E}{2(1+\nu)} \quad (22)$$

where  $E$  is the Young's modulus. It is easy to verify

$$\mathbf{C}^* = \mathbf{S} : \mathbf{C} = \mathbf{C} : \mathbf{S}^T = \frac{1}{\mu} \left[ \frac{1-2\nu}{2(1-\nu)} \mathbf{N} + \mathbf{T} \right]. \quad (23)$$

Therefore,

$$\mathbf{C}^{-1} : \mathbf{S} : \mathbf{C} = \mathbf{S}^T. \quad (24)$$

**2.2 Properties of the Global  $\mathcal{M}$  Tensor.** Obviously, the tensor  $\mathbf{S} + \mathcal{M} : \xi$  is just the Eshelby's tensor for  $\xi \leq 1$ , as defined by Eq. (A7) in the local coordinate system. The  $\mathcal{M}$  can be split into two parts

$$\mathcal{M} = \tilde{\mathbf{M}} + \mathbf{M}, \quad \tilde{\mathbf{M}} = \tilde{\mathbf{M}}[\nu, \mathbf{n}^{(1)}, \mathbf{n}^{(2)}], \quad \mathbf{M} = \mathbf{M}(\nu, \mathbf{n}). \quad (25)$$

In view of the Eq. (A7), the  $\tilde{\mathbf{M}}$  tensor can be defined in the local coordinate system

$$\begin{aligned} \tilde{M}_{1111} = \tilde{M}_{2222} &= \frac{13-8\nu}{4} \gamma, \quad \tilde{M}_{1122} = \tilde{M}_{2211} = \frac{8\nu-1}{4} \gamma, \\ \tilde{M}_{1212} &= \frac{7-8\nu}{4} \gamma, \quad \gamma = \frac{\pi}{8(1-\nu)}. \end{aligned} \quad (26)$$

All other nonzero components are obtained by the cyclic permutation of (1, 2, 3). Similarly, the  $\mathbf{M}$  tensor is

$$M_{1313} = M_{2323} = (\nu-2) \gamma, \quad M_{3333} = 2(2\nu-1) \gamma, \quad (27)$$

$$M_{3311} = M_{3322} = -(4\nu+1) \gamma, \quad M_{1133} = M_{2233} = (2\nu-1) \gamma.$$

All other nonzero components are obtained by the cyclic permutation of (1, 2, 3). Using Eq. (6), the global  $\tilde{\mathbf{M}}$  tensor is obtained

$$\begin{aligned} \tilde{\mathbf{M}} &= \frac{13-8\nu}{4} \gamma [\mathbf{n}^{(1)} \mathbf{n}^{(1)} \mathbf{n}^{(1)} \mathbf{n}^{(1)} + \mathbf{n}^{(2)} \mathbf{n}^{(2)} \mathbf{n}^{(2)} \mathbf{n}^{(2)}] \\ &+ \frac{8\nu-1}{4} \gamma [\mathbf{n}^{(1)} \mathbf{n}^{(1)} \mathbf{n}^{(2)} \mathbf{n}^{(2)} + \mathbf{n}^{(2)} \mathbf{n}^{(2)} \mathbf{n}^{(1)} \mathbf{n}^{(1)}] \\ &+ \frac{7-8\nu}{4} \gamma [\mathbf{n}^{(1)} \mathbf{n}^{(2)} \mathbf{n}^{(1)} \mathbf{n}^{(2)} + \mathbf{n}^{(2)} \mathbf{n}^{(1)} \mathbf{n}^{(1)} \mathbf{n}^{(2)} \\ &+ \mathbf{n}^{(2)} \mathbf{n}^{(1)} \mathbf{n}^{(2)} \mathbf{n}^{(1)} + \mathbf{n}^{(1)} \mathbf{n}^{(2)} \mathbf{n}^{(2)} \mathbf{n}^{(1)}] \end{aligned} \quad (28)$$

and the global  $\mathbf{M}$  tensor is

$$\begin{aligned} \mathbf{M} = M_{ijkl} &= (\nu-2) \gamma (\delta_{ik} n_j n_l + \delta_{il} n_j n_k + \delta_{jk} n_i n_l + \delta_{jl} n_i n_k \\ &- 4 n_i n_j n_k n_l) - (4\nu+1) \gamma n_i n_j (\delta_{kl} - n_k n_l) \\ &+ (2\nu-1) \gamma n_k n_l (\delta_{ij} - n_i n_j) + 2(2\nu-1) \gamma n_i n_j n_k n_l \end{aligned} \quad (29)$$

or

$$\mathbf{M} = 4(\nu-2) \gamma \mathbf{T} - (4\nu+1) \gamma \mathbf{nnl} - (1-2\nu) \gamma \mathbf{lnn} + 6\nu \gamma \mathbf{N}. \quad (30)$$

It is easy to verify

$$\tilde{\mathbf{M}} : \mathbf{S} = \mathbf{0} \Rightarrow \mathbf{S} : \tilde{\mathbf{M}} : \mathbf{S} = \mathbf{0} \Rightarrow \mathbf{S} : \mathcal{M} : \mathbf{S} = \mathbf{S} : \mathbf{M} : \mathbf{S} \quad (31)$$

and

$$\mathbf{S} : \mathbf{M} = 4(\nu-2) \gamma \mathbf{T} - (4\nu+1) \gamma \mathbf{nnl} + \frac{7\nu-4\nu^2-1}{1-\nu} \gamma \mathbf{N} \quad (32)$$

$$\mathbf{S} : \mathbf{M} : \mathbf{S} = -2\gamma \left[ (2-\nu) \mathbf{S}_t + \frac{1-2\nu}{1-\nu} \mathbf{S}_n \right].$$

### 3 Equivalent Eigenstrain of Inhomogeneities

Equation (B4) in Appendix B could be rearranged as

$$[D_{ijmn}^0 - (D_{ijkl}^0 - D_{ijkl}^*) S_{klmn}] \epsilon_{mn}^* = (D_{ijkl}^0 - D_{ijkl}^*) \epsilon_{kl}^0 \quad (33)$$

where  $S_{klmn}$  represents the general Eshelby's tensor of ellipsoidal inclusions, and

$$D_{ijkl}^0 = \mathbf{D}^0 = \lambda^0 \mathbf{II} + 2\mu^0 \mathbf{I},$$

$$\lambda^0 = \frac{\nu^0 E^0}{(1+\nu^0)(1-2\nu^0)}, \quad \mu^0 = \frac{E^0}{2(1+\nu^0)} \quad (34)$$

and  $D_{ijkl}^* = \mathbf{D}^*$  possesses the same form with the parameters  $E^*, \nu^*$ . If assuming  $\nu^0 = \nu^*$ , then

$$\mathbf{D}^* = (1-\theta) \mathbf{D}^0, \quad \theta = 1 - \frac{E^*}{E^0}. \quad (35)$$

Thus Eq. (33) can be simplified as

$$(\mathbf{I} - \theta \mathbf{S}) : \epsilon^* = \theta \epsilon^0 \Rightarrow \epsilon^* = \theta (\mathbf{I} - \theta \mathbf{S})^{-1} : \epsilon^0. \quad (36)$$

An asymptotic expression of the inverse tensor is

$$(\mathbf{I} - \theta \mathbf{S})^{-1} \rightarrow \mathbf{B}^{-1} + \theta \mathbf{B}^{-1} : (\mathbf{S} - \mathbf{S}) : \mathbf{B}^{-1} \quad \text{if } \xi \rightarrow 0 \quad (37)$$

where

$$\mathbf{B} = \mathbf{I} - \theta \mathbf{S}, \quad \mathbf{B}^{-1} = (\mathbf{I} - \theta \mathbf{S})^{-1} = \mathbf{I} + \frac{\theta}{1-\theta} \mathbf{S} \quad (38)$$

due to Eq. (16). Note that, with the asymptotic expression Eq. (5) for  $\xi \rightarrow 0$ ,

$$\begin{aligned} &(\mathbf{I} - \theta \mathbf{S}) : [\mathbf{B}^{-1} + \theta \mathbf{B}^{-1} : (\mathbf{S} - \mathbf{S}) : \mathbf{B}^{-1}] \\ &\rightarrow (\mathbf{B} - \theta \mathcal{M} \xi) : [\mathbf{B}^{-1} + \theta \mathbf{B}^{-1} : \mathcal{M} : \mathbf{B}^{-1} \xi] \\ &= \mathbf{I} - \theta^2 \mathcal{M} : \mathbf{B}^{-1} : \mathcal{M} : \mathbf{B}^{-1} \xi^2 \rightarrow \mathbf{I}. \end{aligned} \quad (39)$$

Thus, Eq. (37) is proved or with an equivalent form

$$(\mathbf{I} - \theta \mathbf{S})^{-1} = \mathbf{B}^{-1} + \theta \mathbf{B}^{-1} : (\mathbf{S} - \mathbf{S}) : \mathbf{B}^{-1} + \mathbf{O}(\xi^2). \quad (40)$$

Then, the eigenstrain is obtained by Eq. (36),

$$\begin{aligned} \epsilon^* &= \theta (\mathbf{I} - \theta \mathbf{S})^{-1} : \epsilon^0 = \theta \mathbf{B}^{-1} : \epsilon^0 + \theta^2 \mathbf{B}^{-1} : (\mathbf{S} - \mathbf{S}) : \mathbf{B}^{-1} : \epsilon^0 \xi \\ &+ \mathbf{O}(\xi^2). \end{aligned} \quad (41)$$

According to Eq. (B5) of the Appendix, the energy disturbance caused by the existence of the penny-shaped inhomogeneity is

$$U^* = \frac{1}{2} \int_{\Omega} \sigma^0 : \epsilon^* dV = \frac{V}{2} \sigma^0 : \epsilon^* \quad (42)$$



since  $\epsilon^*$  is uniform in domain  $\Omega$ , where  $V$  is the size of the spheroidal inhomogeneity

$$V = \frac{4\pi}{3} a_1 a_2 a_3 = \frac{4\pi}{3} a^3 \frac{a_3}{a} = \frac{4\pi}{3} a^3 \xi. \quad (43)$$

Therefore, the energy disturbance is obtained,

$$U^* = \frac{2\pi}{3} a^3 \sigma^0 : \epsilon^* \xi. \quad (44)$$

Evidently, the equivalent eigenstrain  $\epsilon^*$  should be singular or infinitely large, otherwise  $U^* \rightarrow 0$  which implies that the penny-shaped inhomogeneity is insensible at macroscopic level. In order to make  $\epsilon^*$  singular, obviously, the only possibility is that the relative Young's modulus  $\Gamma$  is infinitesimal or infinitely large, i.e.,

$$\Gamma = \frac{E^*}{E^0} \rightarrow 0 \quad \text{or} \quad \Gamma = \frac{E^*}{E^0} \rightarrow \infty. \quad (45)$$

**3.1 Asymptotic Eigenstrain  $\epsilon^*$  for Infinitesimal  $\Gamma$ .** Note that  $1 - \theta = \Gamma$ , and if  $\Gamma \rightarrow 0$ , due to Eq. (38),

$$\theta \rightarrow 1, \quad \mathbf{B}^{-1} \rightarrow \frac{1}{\Gamma} \mathbf{S}. \quad (46)$$

In view of Eqs. (37) and (39), the asymptotic expression of the inverse tensor is

$$(\mathbf{I} - \theta \mathbf{S})^{-1} \rightarrow \frac{1}{\Gamma} \left[ \mathbf{S} + \frac{1}{\Gamma} \mathbf{S} : (\mathbf{S} - \mathbf{S}) : \mathbf{S} \right] \quad (47)$$

with the error

$$\theta^2 \mathcal{M} : \mathbf{B}^{-1} : \mathcal{M} : \mathbf{B}^{-1} \xi^2 \rightarrow \mathcal{M} : \mathbf{S} : \mathcal{M} : \mathbf{S} \left( \frac{\xi}{\Gamma} \right)^2 = \mathbf{M} : \mathbf{S} : \mathbf{M} : \mathbf{S} \xi_{eq}^2 \quad (48)$$

due to Eq. (31). In order to make the error sufficient small, it is required that the equivalent thickness ratio  $\xi_{eq}$  should be sufficient small, i.e.,

$$\xi_{eq} = \frac{\xi}{\Gamma} \ll 1. \quad (49)$$

Note that

$$\frac{1}{\Gamma} [\mathbf{S}(\xi) - \mathbf{S}] \rightarrow \frac{1}{\Gamma} \mathcal{M} \xi = \mathcal{M} \xi_{eq} \rightarrow \mathbf{S}(\xi_{eq}) - \mathbf{S}. \quad (50)$$

Then, Eq. (47) can be rewritten as

$$[\mathbf{I} - \theta \mathbf{S}(\xi)]^{-1} \rightarrow \frac{1}{\Gamma} \{ \mathbf{S} + \mathbf{S} : [\mathbf{S}(\xi_{eq}) - \mathbf{S}] : \mathbf{S} \} = \frac{1}{\Gamma} \mathbf{S} : \mathbf{S}(\xi_{eq}) : \mathbf{S}. \quad (51)$$

In view of Eq. (36), the equivalent eigenstrain is obtained,

$$\epsilon^* = \theta (\mathbf{I} - \theta \mathbf{S})^{-1} : \epsilon^0 \rightarrow (\mathbf{I} - \theta \mathbf{S})^{-1} : \epsilon^0 \rightarrow \frac{1}{\Gamma} \mathbf{S} : \mathbf{S}(\xi_{eq}) : \mathbf{S} : \epsilon^0. \quad (52)$$

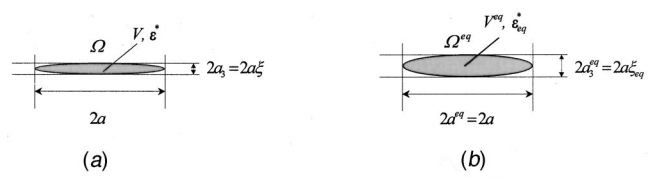
**3.2 Energy-Based Equivalent Inclusion Method.** For the penny-shaped inhomogeneity with  $\Gamma \rightarrow 0$ , its Eshelby's equivalent eigenstrain  $\epsilon^*$  contains a singular factor  $\eta$ , as shown in Eq. (52),

$$\eta = \frac{1}{\Gamma}. \quad (53)$$

The singular factor  $\eta$  ensures a finite energy disturbance  $U^*$ . The energy disturbance is obtained by Eqs. (42) and (43),

$$U^* = \frac{V}{2} \sigma^0 : \epsilon^* = \frac{2\pi}{3} a^3 \xi_{eq} \mathbf{S} : \mathbf{S}(\xi_{eq}) : \mathbf{S} : \epsilon^0 \quad (54)$$

or in an equivalent form



**Fig. 3 (a) The Eshelby's equivalent inclusion  $\Omega$  of the penny-shaped inhomogeneity; (b) the energy-based equivalent inclusion  $\Omega^{eq}$  of the same inhomogeneity**

$$U^* = \frac{V^{eq}}{2} \sigma^0 : \epsilon_{eq}^*, \quad V^{eq} = \eta V = \frac{4\pi}{3} a^3 \xi_{eq},$$

$$\epsilon_{eq}^* = \frac{1}{\eta} \epsilon^* = \mathbf{S} : \mathbf{S}(\xi_{eq}) : \mathbf{S} : \epsilon^0. \quad (55)$$

Both equations show that the energy disturbance of the penny-shaped inhomogeneity can be evaluated by introducing an inclusion prescribed by an fictitious eigenstrain. However, the Eshelby's equivalent inclusion method, corresponding to Eq. (54), is based on the inclusion  $\Omega$  which is geometrically identical to the inhomogeneity, as shown in Fig. 3(a). On the other hand, Eq. (55) represents a new approach, the *energy-based equivalent inclusion method*. The method yields the same energy disturbance but its inclusion  $\Omega^{eq}$  is geometrically different with the inhomogeneity, as shown in Fig. 3(b). Especially, due to  $\xi_{eq} \ll 1$ , the energy-based equivalent inclusion, is still penny-shaped and its spheroidal principal half-axes are  $a^{eq} = a$  and  $a_3^{eq} = a \xi_{eq} = \xi_{eq} / \xi a_3$ , as shown in Fig. 3(b). Thus, the energy-based equivalent inclusion  $\Omega^{eq}$  is much thicker than the Eshelby's equivalent inclusion  $\Omega$ .

The two methods are equivalent in the sense,

$$U^* = \frac{V}{2} \sigma^0 : \epsilon^* = \frac{V^{eq}}{2} \sigma^0 : \epsilon_{eq}^* \Rightarrow V \epsilon^* = V^{eq} \epsilon_{eq}^* \quad (56)$$

which shows that  $V^{eq}$  and  $\epsilon_{eq}^*$  are one pair of conjugate variables. Although  $V^{eq}$  and  $\epsilon_{eq}^*$  may have infinite combinations based on Eq. (56), it is required that the  $V^{eq}$  and  $\epsilon_{eq}^*$  are obtained from  $V$  and  $\epsilon^*$  by moving the singular factor from  $\epsilon^*$  to  $V$ , as shown in Eq. (55). That both  $V^{eq}$  and  $\epsilon_{eq}^*$  are finite quantities, is one of the essential ingredients of the energy-based equivalent inclusion method.

Note that  $\xi_{eq} \ll 1 \Rightarrow \mathbf{S}(\xi_{eq}) \approx \mathbf{S}$ . Then,

$$\epsilon_{eq}^* = \mathbf{S} : \mathbf{S}(\xi_{eq}) : \mathbf{S} : \epsilon^0 \approx \mathbf{S} : \mathbf{S} : \mathbf{S} : \epsilon^0 = \mathbf{S} : \epsilon^0. \quad (57)$$

Note that

$$\epsilon^0 = \mathbf{C}^0 : \sigma^0, \quad \mathbf{C}^0 = -\frac{\nu^0}{E^0} \mathbf{I} + \frac{1 + \nu^0}{E^0} \mathbf{I}. \quad (58)$$

In the finite domain  $\Omega^{eq}$ , the energy-based equivalent density of the energy disturbance is defined as

$$u^{eq} = \frac{U^*}{V^{eq}} = \frac{1}{2} \sigma^0 : \epsilon_{eq}^* \approx \frac{1}{2} \sigma^0 : \mathbf{S} : \mathbf{C}^0 : \sigma^0 = \frac{1}{2} \sigma^0 : \mathbf{C}^* : \sigma^0 \quad (59)$$

where  $\mathbf{C}^*$  is the equivalent elastic compliance, due to Eq. (23),

$$\mathbf{C}^* = \mathbf{S} : \mathbf{C}^0 = \frac{1}{\mu^0} \left[ \frac{1 - 2\nu^0}{2(1 - \nu^0)} \mathbf{N} + \mathbf{T} \right]. \quad (60)$$

Obviously,  $\mathbf{C}^*$  possesses the same symmetry as  $\mathbf{C}^0$ , i.e.,  $C_{ijkl}^* = C_{jikl}^* = C_{ijlk}^* = C_{klij}^*$  and is a positive definite tensor since

$$u^{eq} \approx \frac{1}{2} \sigma^0 : \mathbf{C}^* : \sigma^0 = \frac{1}{\mu^0} \left[ \frac{1 - 2\nu^0}{2(1 - \nu^0)} \sigma_0^2 + \tau_0^2 \right] \geq 0 \quad (61)$$

where  $\sigma_0$  and  $\tau_0$  are the normal and shear components, respectively, of the stress tensor  $\boldsymbol{\sigma}^0$  on the plane of the penny-shaped inhomogeneity,  $\sigma_0^2 = \boldsymbol{\sigma}^0 : \mathbf{N} : \boldsymbol{\sigma}^0$  and  $\tau_0^2 = \boldsymbol{\sigma}^0 : \mathbf{T} : \boldsymbol{\sigma}^0$ : see, e.g., Eq. (11).

### 3.3 Asymptotic Eigenstrain $\boldsymbol{\varepsilon}^*$ for Infinitely Large $\Gamma$ .

Note that  $\theta = 1 - \Gamma$ , and if  $\Gamma \rightarrow \infty$ ,

$$\theta \rightarrow -\Gamma \rightarrow -\infty, \quad \mathbf{B}^{-1} \rightarrow \mathbf{I} - \mathbf{S}. \quad (62)$$

In view of Eqs. (37) and (39), the asymptotic expression of the inverse tensor is

$$(\mathbf{I} - \theta \mathbf{S})^{-1} \rightarrow \mathbf{I} - \mathbf{S} - \Gamma(\mathbf{I} - \mathbf{S}) : (\mathbf{S} - \mathbf{S}) : (\mathbf{I} - \mathbf{S}) \quad (63)$$

with the error

$$\begin{aligned} \theta^2 \mathcal{M} : \mathbf{B}^{-1} : \mathcal{M} : \mathbf{B}^{-1} \xi^2 &\rightarrow \mathcal{M} : (\mathbf{I} - \mathbf{S}) : \mathcal{M} : (\mathbf{I} - \mathbf{S}) (\Gamma \xi)^2 \\ &= \mathcal{M} : (\mathbf{I} - \mathbf{S}) : \mathcal{M} : (\mathbf{I} - \mathbf{S}) \xi_{eq}^2. \end{aligned} \quad (64)$$

In order to make the error sufficient small, it is required that the equivalent aspect ratio  $\xi_{eq}$  should be sufficient small, i.e.,

$$\xi_{eq} = \Gamma \xi \ll 1. \quad (65)$$

Note that

$$\Gamma[\mathbf{S}(\xi) - \mathbf{S}] \rightarrow \Gamma \mathcal{M} \xi = \mathcal{M} \xi_{eq} \rightarrow \mathbf{S}(\xi_{eq}) - \mathbf{S}. \quad (66)$$

Then, Eq. (63) can be rewritten as

$$\begin{aligned} [\mathbf{I} - \theta \mathbf{S}(\xi)]^{-1} &\rightarrow \mathbf{I} - \mathbf{S} + (\mathbf{I} - \mathbf{S}) : [\mathbf{S}(\xi_{eq}) - \mathbf{S}] : (\mathbf{I} - \mathbf{S}) \\ &= \mathbf{I} - \mathbf{S} + (\mathbf{I} - \mathbf{S}) : \mathbf{S}(\xi_{eq}) : (\mathbf{I} - \mathbf{S}). \end{aligned} \quad (67)$$

In view of Eq. (36), the equivalent eigenstrain is obtained,

$$\boldsymbol{\varepsilon}^* = \theta(\mathbf{I} - \theta \mathbf{S})^{-1} : \boldsymbol{\varepsilon}^0 \rightarrow -\Gamma[\mathbf{I} - \mathbf{S} + (\mathbf{I} - \mathbf{S}) : \mathbf{S}(\xi_{eq}) : (\mathbf{I} - \mathbf{S})] : \boldsymbol{\varepsilon}^0. \quad (68)$$

The eigenstrain contains the singular factor  $\Gamma$  which ensures a finite energy disturbance  $U^*$ . Similarly, the energy disturbance takes the form

$$U^* = \frac{V}{2} \boldsymbol{\sigma}^0 : \boldsymbol{\varepsilon}^* = \frac{V^{eq}}{2} \boldsymbol{\sigma}^0 : \boldsymbol{\varepsilon}_{eq}^* \quad (69)$$

where

$$\eta = \Gamma, \quad (70a)$$

$$V^{eq} = \frac{4\pi}{3} a^3 \xi_{eq}, \quad (70b)$$

$$\boldsymbol{\varepsilon}_{eq}^* = -[\mathbf{I} - \mathbf{S} + (\mathbf{I} - \mathbf{S}) : \mathbf{S}(\xi_{eq}) : (\mathbf{I} - \mathbf{S})] : \boldsymbol{\varepsilon}^0, \quad (70c)$$

which is based on the energy-based equivalent inclusion method, as shown in Fig. 3. The Eshelby's equivalent inclusion and the corresponding energy-based equivalent inclusion can still be related by  $a^{eq} = a$  and  $a_3^{eq} = a \xi_{eq}$ . Note that,  $\xi_{eq} \ll 1 \Rightarrow \mathbf{S}(\xi_{eq}) \approx \mathbf{S}$ . Then, due to Eq. (70c),

$$\boldsymbol{\varepsilon}_{eq}^* \approx (\mathbf{S} - \mathbf{I}) : \boldsymbol{\varepsilon}^0. \quad (71)$$

Similar to Eq. (59), the equivalent density of the energy disturbance is

$$u^{eq} = \frac{U^*}{V^{eq}} = \frac{1}{2} \boldsymbol{\sigma}^0 : \boldsymbol{\varepsilon}_{eq}^* \approx \frac{1}{2} \boldsymbol{\sigma}^0 : (\mathbf{C}^* - \mathbf{C}^0) : \boldsymbol{\sigma}^0. \quad (72)$$

## 4 Asymptotic Equivalent Eigenstrain of Crack

If  $\mathbf{D}^* = \mathbf{0}$  and  $\xi \rightarrow 0$ , the inhomogeneity becomes a penny-shaped crack.<sup>1</sup> In view of Eq. (33), the eigenstrain of the crack is determined by

<sup>1</sup>If  $\mathbf{D}^* = \infty$  and  $\xi \rightarrow 0$ , the inhomogeneity becomes a penny-shaped anticrack, and the corresponding eigenstrain is determined by  $\mathbf{S} : \boldsymbol{\varepsilon}^* = -\boldsymbol{\varepsilon}^0$ .

$$(\mathbf{I} - \mathbf{S}) : \boldsymbol{\varepsilon}^* = \boldsymbol{\varepsilon}^0 \Rightarrow \boldsymbol{\varepsilon}^* = (\mathbf{I} - \mathbf{S})^{-1} : \boldsymbol{\varepsilon}^0. \quad (73)$$

An asymptotic solution of the inverse is

$$(\mathbf{I} - \mathbf{S})^{-1} \rightarrow \frac{1}{\xi} (c_1 \mathbf{S}_n + c_2 \mathbf{S}_t) + \mathbf{A} \quad (74)$$

where

$$c_1 = \frac{1}{2\gamma} \frac{1 - \nu^0}{1 - 2\nu^0}, \quad c_2 = \frac{1}{2\gamma} \frac{1}{2 - \nu^0}, \quad (75)$$

$$\mathbf{A} = \mathbf{I} - (1 - 2\nu^0) \gamma c_1 \mathbf{lnn} : \mathbf{S}_n.$$

Note that, if  $\xi \rightarrow 0$ ,

$$\begin{aligned} (\mathbf{I} - \mathbf{S}) : (c_1 \mathbf{S}_n + c_2 \mathbf{S}_t) \frac{1}{\xi} &\rightarrow (\mathbf{I} - \mathbf{S} - \mathcal{M} \xi) : (c_1 \mathbf{S}_n + c_2 \mathbf{S}_t) \frac{1}{\xi} \\ &= -\mathcal{M} : (c_1 \mathbf{S}_n + c_2 \mathbf{S}_t) \\ &= -\mathbf{M} : (c_1 \mathbf{S}_n + c_2 \mathbf{S}_t) \\ &= (1 - 2\nu^0) \gamma c_1 \mathbf{S}_n + 2(2 - \nu^0) \gamma c_2 \mathbf{S}_t \\ &\quad + (1 - 2\nu^0) \gamma c_1 \mathbf{lnn} : \mathbf{S}_n \end{aligned} \quad (76)$$

due to

$$(\mathbf{I} - \mathbf{S}) : \mathbf{S}_n = (\mathbf{I} - \mathbf{S}) : \mathbf{S}_t = \tilde{\mathbf{M}} : \mathbf{S}_n = \tilde{\mathbf{M}} : \mathbf{S}_t = 0, \quad (77)$$

and

$$\begin{aligned} (\mathbf{I} - \mathbf{S}) : \mathbf{A} &\rightarrow (\mathbf{I} - \mathbf{S} - \mathcal{M} \xi) : \mathbf{A} = \mathbf{A} - \mathbf{S} : \mathbf{A} - \mathcal{M} : \mathbf{A} \xi \\ &= \mathbf{A} - \mathbf{S} + (1 - 2\nu^0) \gamma c_1 \mathbf{S} : \mathbf{lnn} : \mathbf{S}_n - \mathcal{M} : \mathbf{A} \xi \\ &= \mathbf{I} - (1 - 2\nu^0) \gamma c_1 \mathbf{lnn} : \mathbf{S}_n - \mathbf{S} \\ &\quad + (1 - 2\nu^0) \frac{1 + \nu^0}{1 - \nu^0} \gamma c_1 \mathbf{S}_n - \mathcal{M} : \mathbf{A} \xi \end{aligned} \quad (78)$$

due to

$$\mathbf{S} : \mathbf{lnn} : \mathbf{S}_n = \mathbf{S}_n : \mathbf{lnn} : \mathbf{S}_n = \frac{1 + \nu^0}{1 - \nu^0} \mathbf{S}_n. \quad (79)$$

Therefore, for  $\xi \rightarrow 0$ ,

$$\begin{aligned} (\mathbf{I} - \mathbf{S}) : \left[ \frac{1}{\xi} (c_1 \mathbf{S}_n + c_2 \mathbf{S}_t) + \mathbf{A} \right] &\rightarrow \mathbf{I} - \mathbf{S} - \mathcal{M} : \mathbf{A} \xi + 2(2 - \nu^0) \gamma c_2 \mathbf{S}_t \\ &\quad + (1 - 2\nu^0) \gamma c_1 \mathbf{S}_n + (1 - 2\nu^0) \frac{1 + \nu^0}{1 - \nu^0} \gamma c_1 \mathbf{S}_n \\ &= \mathbf{I} - \mathbf{S} - \mathcal{M} : \mathbf{A} \xi + \mathbf{S}_t + \mathbf{S}_n \\ &= \mathbf{I} - \mathcal{M} : \mathbf{A} \xi \rightarrow \mathbf{I} \end{aligned} \quad (80)$$

or

$$(\mathbf{I} - \mathbf{S})^{-1} = \frac{1}{\xi} (c_1 \mathbf{S}_n + c_2 \mathbf{S}_t) + \mathbf{A} + \mathbf{O}(\xi). \quad (81)$$

The asymptotic expression of the eigenstrain of the crack is obtained,

$$\begin{aligned} \boldsymbol{\varepsilon}^* &= (\mathbf{I} - \mathbf{S})^{-1} : \boldsymbol{\varepsilon}^0 \rightarrow \left[ \frac{1}{\xi} (c_1 \mathbf{S}_n + c_2 \mathbf{S}_t) + \mathbf{A} \right] : \boldsymbol{\varepsilon}^0 \\ &\rightarrow \frac{1}{\xi} (c_1 \mathbf{S}_n + c_2 \mathbf{S}_t) : \boldsymbol{\varepsilon}^0. \end{aligned} \quad (82)$$

The energy disturbance caused by the presence of the crack is obtained,



**Fig. 4 (a) The Eshelby's equivalent inclusion of a crack; (b) the spherical energy-based equivalent inclusion of the same crack**

$$U_c^* = \frac{V}{2} \boldsymbol{\sigma}^0 : \boldsymbol{\epsilon}^0 = \lim_{\xi \rightarrow 0} \frac{1}{2} \frac{4\pi}{3} a^3 \xi \left[ \frac{1}{\xi} (c_1 \mathbf{S}_n + c_2 \mathbf{S}_t) : \boldsymbol{\epsilon}^0 \right] \\ = \frac{2\pi}{3} a^3 (c_1 \mathbf{S}_n + c_2 \mathbf{S}_t) : \boldsymbol{\epsilon}^0 \quad (83)$$

or as an equivalent form

$$U_c^* = \frac{V^{eq}}{2} \boldsymbol{\sigma}^0 : \boldsymbol{\epsilon}_{eq}^*, \quad \eta = \frac{1}{\xi}, \quad V^{eq} = \frac{4\pi}{3} a^3, \quad (84) \\ \boldsymbol{\epsilon}_{eq}^* = (c_1 \mathbf{S}_n + c_2 \mathbf{S}_t) : \boldsymbol{\epsilon}^0$$

which indicates that the energy-based equivalent inclusion of a crack is a sphere, i.e.,  $a_1^{eq} = a_2^{eq} = a_3^{eq} = a$ , as shown in Fig. 4.

In view of Eqs. (58) and (21), the energy disturbance  $U_c^*$  of the crack in Eq. (83) can be rewritten as

$$U_c^* = \frac{a^3}{E^0} [\sigma_0^2 f(\nu^0) + \tau_0^2 g(\nu^0)], \quad f(\nu^0) = \frac{8}{3} [1 - (\nu^0)^2], \quad (85) \\ g(\nu^0) = \frac{16[1 - (\nu^0)^2]}{3(2 - \nu^0)}$$

where  $\sigma_0^2 = \boldsymbol{\sigma}^0 : \mathbf{N} : \boldsymbol{\sigma}^0$  and  $\tau_0^2 = \boldsymbol{\sigma}^0 : \mathbf{T} : \boldsymbol{\sigma}^0$ . Evidently, it is exactly the result obtained by Budiansky and O'Connell [11]. The coefficients  $c_1$  and  $c_2$  are directly related with  $f(\nu^0)$  and  $g(\nu^0)$ . Here, we give the coefficients a new mathematical explanation, due to Eqs. (32) and (75),

$$\mathbf{S} : \mathbf{M} : \mathbf{S} = - \left( \frac{1}{c_1} \mathbf{S}_n + \frac{1}{c_2} \mathbf{S}_t \right) \Rightarrow \mathbf{S}_n : \mathbf{M} : \mathbf{S}_n = - \frac{1}{c_1} \mathbf{S}_n, \quad (86) \\ \mathbf{S}_t : \mathbf{M} : \mathbf{S}_t = - \frac{1}{c_2} \mathbf{S}_t.$$

## 5 Prime Eigenstrain

In view of Eqs. (3) and (41), the limits of the Eshelby's tensor and the eigenstrain for  $\xi \rightarrow 0$  are

$$\lim_{\xi \rightarrow 0} \mathbf{S} \rightarrow \mathbf{S}, \quad \lim_{\xi \rightarrow 0} \boldsymbol{\epsilon}^* = \theta \mathbf{B}^{-1} : \boldsymbol{\epsilon}^0 = \frac{\theta}{1 - \theta} [(1 - \theta) \mathbf{I} + \theta \mathbf{S}] : \boldsymbol{\epsilon}^0. \quad (87)$$

The limit of the eigenstrain is termed the prime eigenstrain and is still denoted by  $\boldsymbol{\epsilon}^*$  in this section. The asymptotic expressions of the prime eigenstrain are

$$\boldsymbol{\epsilon}^* \rightarrow \frac{1}{\Gamma} \mathbf{S} : \boldsymbol{\epsilon}^0, \quad \text{if } \Gamma \rightarrow 0; \quad \boldsymbol{\epsilon}^* \rightarrow \Gamma (\mathbf{S} - \mathbf{I}) : \boldsymbol{\epsilon}^0, \quad \text{if } \Gamma \rightarrow \infty. \quad (88)$$

As shown in Eqs. (52) and (68), the prime eigenstrain is the dominant part of the eigenstrain if  $\xi_{eq} \ll 1$ . In fact, the approximate asymptotic eigenstrains in Eqs. (57) and (71) are just the

asymptotic prime strains without the singular factors. Then, it is worthy to discuss the prime eigenstrain and some related properties.

**5.1 Local Prime Eigenstrain.** In the local coordinate system, i.e.,  $\mathbf{n} = \{0, 0, 1\}$ , the prime eigenstrain is determined by Eq. (87),

$$\boldsymbol{\epsilon}_{11}^* = \theta \boldsymbol{\epsilon}_{11}^0, \quad \boldsymbol{\epsilon}_{22}^* = \theta \boldsymbol{\epsilon}_{22}^0, \quad (89) \\ \boldsymbol{\epsilon}_{33}^* = \frac{\theta}{1 - \theta} \boldsymbol{\epsilon}_{33}^0 + \frac{\theta^2}{1 - \theta} \frac{\nu^0}{1 - \nu^0} (\boldsymbol{\epsilon}_{11}^0 + \boldsymbol{\epsilon}_{22}^0), \\ \boldsymbol{\epsilon}_{12}^* = \theta \boldsymbol{\epsilon}_{12}^0, \quad \boldsymbol{\epsilon}_{23}^* = \frac{\theta}{1 - \theta} \boldsymbol{\epsilon}_{23}^0, \quad \boldsymbol{\epsilon}_{31}^* = \frac{\theta}{1 - \theta} \boldsymbol{\epsilon}_{31}^0$$

which is consistent with Eqs. (22.29) and (22.19) of Mura [2]. For example, according to Eqn. (22.19) of Mura [2],

$$\boldsymbol{\epsilon}_{23}^* = \frac{\theta}{1 - 2\theta \mathbf{S}_{2323}} \boldsymbol{\epsilon}_{23}^0. \quad (90)$$

On the other hand, by Eq. (87),

$$\boldsymbol{\epsilon}_{23}^* = \theta \left( 1 + \frac{2\theta}{1 - \theta} \mathbf{S}_{2323} \right) \boldsymbol{\epsilon}_{23}^0. \quad (91)$$

Although they possess different forms, both of them yield the same result since  $\mathbf{S}_{2323} = 0.5$ .

**5.2 Stress of the Inhomogeneity.** The strain disturbance in the inhomogeneity  $\Omega$  is, due to Eqs. (A1) and (14),

$$\boldsymbol{\epsilon} = \mathbf{S} : \boldsymbol{\epsilon}^* = \mathbf{S} : \left( \theta \mathbf{I} + \frac{\theta^2}{1 - \theta} \mathbf{S} \right) : \boldsymbol{\epsilon}^0 = \frac{\theta}{1 - \theta} \mathbf{S} : \boldsymbol{\epsilon}^0. \quad (92)$$

The stress in the penny-shaped inhomogeneity  $\Omega$  is uniform and can be found due to Eqs. (B4) and (24)

$$\hat{\boldsymbol{\sigma}} = \mathbf{D}^0 : (\boldsymbol{\epsilon}^0 + \boldsymbol{\epsilon} - \boldsymbol{\epsilon}^*) = \mathbf{D}^0 : [(1 - \theta) \mathbf{I} + \theta \mathbf{S}] : \boldsymbol{\epsilon}^0 \\ = \mathbf{D}^0 : [(1 - \theta) \mathbf{I} + \theta \mathbf{S}] : \mathbf{C}^0 : \boldsymbol{\sigma}^0 \\ = [(1 - \theta) \mathbf{I} + \theta \mathbf{S}^T] : \boldsymbol{\sigma}^0. \quad (93)$$

The stress components of the inhomogeneity in the local coordinate system,  $\mathbf{n} = \{0, 0, 1\}$ , is determined by Eq. (93)

$$\hat{\sigma}_{11} = (1 - \theta) \sigma_{11}^0 + \frac{\theta \nu^0}{1 - \nu^0} \sigma_{33}^0, \quad \hat{\sigma}_{22} = (1 - \theta) \sigma_{22}^0 + \frac{\theta \nu^0}{1 - \nu^0} \sigma_{33}^0, \quad (94) \\ \hat{\sigma}_{33} = \sigma_{33}^0, \quad \hat{\sigma}_{12} = (1 - \theta) \sigma_{12}^0, \quad \hat{\sigma}_{23} = \sigma_{23}^0, \quad \hat{\sigma}_{31} = \sigma_{31}^0.$$

The inhomogeneity stress vector on the plane of the penny-shaped inhomogeneity is

$$\mathbf{n} \cdot \hat{\boldsymbol{\sigma}} = \{\hat{\sigma}_{31}, \hat{\sigma}_{32}, \hat{\sigma}_{33}\} = \{\sigma_{31}^0, \sigma_{32}^0, \sigma_{33}^0\} = \mathbf{n} \cdot \boldsymbol{\sigma}^0 \quad (95)$$

which indicates that the inhomogeneity stress vector on the inhomogeneity plane is equal to the farfield stress vector on the same plane and has nothing to do with the elastic modulus  $E^*$  or  $E^0$ . The key leading to Eq. (95) is Eq. (15). By using Eqs. (93) and (15),

$$\mathbf{n} \cdot \hat{\boldsymbol{\sigma}} = \mathbf{n} \cdot [(1 - \theta) \mathbf{I} + \theta \mathbf{S}^T] : \boldsymbol{\sigma}^0 = \mathbf{n} \cdot \boldsymbol{\sigma}^0 + \theta \mathbf{n} \cdot (-\mathbf{I} + \mathbf{S}^T) : \boldsymbol{\sigma}^0 \\ = \mathbf{n} \cdot \boldsymbol{\sigma}^0 + \theta \boldsymbol{\sigma}^0 : (\mathbf{S} - \mathbf{I}) \cdot \mathbf{n} = \mathbf{n} \cdot \boldsymbol{\sigma}^0. \quad (96)$$

## 6 Influences of Unequal Poisson's Ratios

During the preceding deduction, the equality of  $\nu^0 = \nu^*$  is required to achieve  $\mathbf{D}^* = (1 - \theta) \mathbf{D}^0$  which simplifies the problem significantly. Now we discuss the influences on the deduced expressions owing to the deviation from the assumption, i.e.,  $\nu^0 \neq \nu^*$ . Note that

$$\mathbf{C}^0:\mathbf{D}^* = \frac{\nu^* - \nu^0}{(1-2\nu^*)(1+\theta^0)} \frac{\mu^*}{\mu^0} \mathbf{I} + \frac{\mu^*}{\mu^0} \mathbf{I} = \theta_\nu \Gamma \mathbf{I} + \Gamma \mathbf{I} \quad (97)$$

where

$$\theta_\nu = \frac{\nu^* - \nu^0}{(1-2\theta^*)(1+\nu^0)}, \quad \Gamma = \frac{\mu^*}{\mu^0}. \quad (98)$$

Then Eq. (33) can be recast as

$$(\mathbf{I} - \theta \mathbf{S}) : \boldsymbol{\varepsilon}^* = \theta \boldsymbol{\varepsilon}^0 - k \mathbf{I}, \quad (99a)$$

where

$$\theta = 1 - \Gamma, \quad (99b)$$

$$k = \theta_\nu \Gamma (\mathbf{I} : \mathbf{S} : \boldsymbol{\varepsilon}^* + \text{tr} \boldsymbol{\varepsilon}^0). \quad (99c)$$

Note that the  $\theta$  defined in Eq. (99) is consistent with the  $\theta$  defined in Eq. (35) due to  $\nu^0 = \nu^*$  there. As compared with Eq. (36), the unequal  $\nu^0$  and  $\nu^*$  will create an additional term  $kl$  which is a spherical tensor and disappears if  $\nu^* = \nu^0$ . Then, the eigenstrain is obtained,

$$\boldsymbol{\varepsilon}^* = \boldsymbol{\varepsilon}_\mu^* + \boldsymbol{\varepsilon}_{gn}^*, \quad (100a)$$

$$\boldsymbol{\varepsilon}_\mu^* = \theta (\mathbf{I} - \theta \mathbf{S})^{-1} : \boldsymbol{\varepsilon}^0, \quad (100b)$$

$$\boldsymbol{\varepsilon}_\nu^* = -k (\mathbf{I} - \theta \mathbf{S})^{-1} : \mathbf{I}. \quad (100c)$$

The leading term  $\boldsymbol{\varepsilon}_\mu^*$  is exactly the eigenstrain  $\boldsymbol{\varepsilon}^*$  for equal  $\nu^*$  and  $\nu^0$ , as discussed in Section 3. The energy disturbance can also be written as a similar form

$$U^* = \frac{V}{2} \boldsymbol{\sigma}^0 : \boldsymbol{\varepsilon}^* = U_\mu^* + U_\nu^*, \quad U_\mu^* = \frac{V}{2} \boldsymbol{\sigma}^0 : \boldsymbol{\varepsilon}_\mu^*, \quad U_\nu^* = \frac{V}{2} \boldsymbol{\sigma}^0 : \boldsymbol{\varepsilon}_\nu^*. \quad (101)$$

**6.1 Equal Shear Moduli.** Consider the case of  $\mu^0 = \mu^*$  and  $\nu^0 \neq \nu^*$ . Since  $\theta = 0$  due to  $\mu^0 = \mu^*$ , one obtains by Eq. (100)

$$\boldsymbol{\varepsilon}_\mu^* = \mathbf{0}, \quad (102a)$$

$$\boldsymbol{\varepsilon}_\nu^* = -k \mathbf{I}. \quad (102b)$$

Substituting Eq. (102b) into Eq. (99c) then

$$k = \theta_\nu (-\mathbf{I} : \mathbf{S} : k \mathbf{I} + \text{tr} \boldsymbol{\varepsilon}^0) \Rightarrow k = \frac{\theta_\nu}{1 + \mathbf{I} : \mathbf{S} : \mathbf{I} \theta_\nu} \text{tr} \boldsymbol{\varepsilon}^0 = w \text{tr} \boldsymbol{\varepsilon}^0 \quad (103)$$

where

$$w = \frac{\theta_\nu}{1 + \mathbf{I} : \mathbf{S} : \mathbf{I} \theta_\nu}, \quad \mathbf{I} : \mathbf{S} : \mathbf{I} \approx \mathbf{I} : \mathbf{S} : \mathbf{I} = \frac{1 + \nu^0}{1 - \nu^0}. \quad (104)$$

Then,  $U_\mu^* = 0$  due to  $\boldsymbol{\varepsilon}_\mu^* = \mathbf{0}$ , and the energy disturbance contributed by the unequal Poisson's ratios is

$$\begin{aligned} U_\nu^* &= \frac{V}{2} \boldsymbol{\sigma}^0 : \boldsymbol{\varepsilon}_\nu^* = -\frac{2\pi}{3} a^3 w \text{tr} \boldsymbol{\sigma}^0 \text{tr} \boldsymbol{\varepsilon}^0 \xi \\ &= -\frac{2\pi(1-2\nu^0)}{3E^0} a^3 w (\text{tr} \boldsymbol{\sigma}^0)^2 \xi \end{aligned} \quad (105)$$

due to

$$\text{tr} \boldsymbol{\varepsilon}^0 = \frac{1-2\nu^0}{E^0} \text{tr} \boldsymbol{\sigma}^0. \quad (106)$$

Obviously,  $w$  is a finite quantity for any combination of  $\nu^*$  and  $\nu^0$ , and even for  $\nu^* \rightarrow 0.5$  since

$$\nu^* \rightarrow 0.5 \Rightarrow \theta_\nu \rightarrow \infty \Rightarrow w \rightarrow \frac{1-\nu^0}{1+\nu^0}. \quad (107)$$

Therefore,  $U_\nu^* \rightarrow 0$  if  $\xi \rightarrow 0$ , which indicates again that the energy disturbance is trivial if the difference between  $E^0$  and  $E^*$  is limited.

**6.2 Infinitesimal Relative Shear Modulus  $\Gamma$ .** For  $\Gamma \rightarrow 0$  or  $\theta \rightarrow 1$ , we adopt the following asymptotic expressions

$$\boldsymbol{\varepsilon}^* \rightarrow \frac{1}{\Gamma} \mathbf{S} : \boldsymbol{\varepsilon}^0, \quad (\mathbf{I} - \theta \mathbf{S})^{-1} \rightarrow \frac{1}{\Gamma} \mathbf{S}, \quad \mathbf{S} \rightarrow \mathbf{S}. \quad (108)$$

They are the prime eigenstrain and corresponding quantities; see, e.g., Eqs. (88) and (47). These asymptotic quantities dominate the behaviors of the penny-shaped inhomogeneity if  $\xi_{eq} = \xi/\Gamma \ll 1$ . Note that

$$\mathbf{S} : \boldsymbol{\varepsilon}^* \rightarrow \mathbf{S} : \left( \frac{1}{\Gamma} \mathbf{S} : \boldsymbol{\varepsilon}^0 \right) = \frac{1}{\Gamma} \mathbf{S} : \boldsymbol{\varepsilon}^0. \quad (109)$$

Therefore,

$$k = \theta_\nu \Gamma (\mathbf{I} : \mathbf{S} : \boldsymbol{\varepsilon}^0 + \text{tr} \boldsymbol{\varepsilon}^0) \rightarrow \theta_\nu \Gamma \left( \frac{1}{\Gamma} \mathbf{I} : \mathbf{S} : \boldsymbol{\varepsilon}^0 + \text{tr} \boldsymbol{\varepsilon}^0 \right) \rightarrow \theta_\nu \mathbf{I} : \mathbf{S} : \boldsymbol{\varepsilon}^0. \quad (110)$$

Note that, due to Eq. (60),

$$\mathbf{I} : \mathbf{S} : \mathbf{C}^0 = \frac{1-2\nu^0}{1-\nu^0} \frac{1}{2\mu^0} \mathbf{nn}. \quad (111)$$

Therefore,

$$k \rightarrow \theta_\nu \mathbf{I} : \mathbf{S} : \boldsymbol{\varepsilon}^* = \theta_\nu \mathbf{I} : \mathbf{S} : \mathbf{C}^0 : \boldsymbol{\sigma}^0 = \frac{1-2\nu^0}{1-\nu^0} \frac{\theta_\nu}{2\mu^0} \mathbf{nn} : \boldsymbol{\sigma}^0. \quad (112)$$

Note that

$$(\mathbf{I} - \theta \mathbf{S})^{-1} : \mathbf{I} \rightarrow \frac{1}{\Gamma} \mathbf{S} : \mathbf{I} = \frac{1+\nu^0}{1-\nu^0} \frac{1}{\Gamma} \mathbf{nn}. \quad (113)$$

Then, the eigenstrain contributed by the unequal Poisson's ratios is obtained by Eq. (100c)

$$\boldsymbol{\varepsilon}_\nu^* = -k (\mathbf{I} - \theta \mathbf{S})^{-1} : \mathbf{I} \rightarrow -\frac{1}{\Gamma} \frac{(1-2\nu^0)(1+\nu^0)}{(1-\nu^0)^2} \frac{\theta_\nu}{2\mu^0} \mathbf{N} : \boldsymbol{\sigma}^0 \quad (114)$$

with the singular factor  $1/\Gamma$ . The corresponding energy disturbance is

$$U_\nu^* = \frac{V}{2} \boldsymbol{\sigma}^0 : \boldsymbol{\varepsilon}_\nu^* \rightarrow -\frac{2\pi}{3} a^3 \xi_{eq} \frac{\theta_\nu^*}{2\mu^0} \sigma_0^2, \quad (115)$$

$$\theta_\nu^* = \frac{(1-2\nu^0)(1+\nu^0)}{(1-\nu^0)^2} \theta_\nu$$

or in an equivalent form based on the energy-based equivalent inclusion method,

$$U_\nu^* = \frac{V^{eq}}{2} \boldsymbol{\sigma}^0 : \boldsymbol{\varepsilon}_{\nu,eq}^*, \quad V^{eq} = \frac{4\pi}{3} a^3 \xi_{eq}, \quad \boldsymbol{\varepsilon}_{\nu,eq}^* \rightarrow -\frac{\theta_\nu^*}{2\mu^0} \mathbf{N} : \boldsymbol{\sigma}^0. \quad (116)$$

**6.3 Infinitely Large Relative Shear Modulus  $\Gamma$ .** For  $\Gamma \rightarrow \infty$  or  $\theta \rightarrow -\infty$ , we adopt the following asymptotic expressions

$$\boldsymbol{\varepsilon}^* \rightarrow \Gamma (\mathbf{S} - \mathbf{I}) : \boldsymbol{\varepsilon}^0, \quad (\mathbf{I} - \theta \mathbf{S})^{-1} \rightarrow \mathbf{I} - \mathbf{S}, \quad \mathbf{S} \rightarrow \mathbf{S}. \quad (117)$$

They are the prime eigenstrain and corresponding quantities; see, e.g., Eqs. (88) and (63). These asymptotic quantities dominate the behaviors of the penny-shaped inhomogeneity if  $\xi_{eq} = \Gamma \xi \ll 1$ . Note that

$$\mathbf{S} : \boldsymbol{\varepsilon}^* \rightarrow \mathbf{S} : [\Gamma (\mathbf{S} - \mathbf{I}) : \boldsymbol{\varepsilon}^0] = \mathbf{0}. \quad (118)$$

Therefore,



$$k = \theta_\nu \Gamma (\mathbf{I} : \boldsymbol{\varepsilon}^0 + \text{tr} \boldsymbol{\varepsilon}^0) \rightarrow \theta_\nu \Gamma \text{tr} \boldsymbol{\varepsilon}^0 = \frac{1-2\nu^0}{E^0} \theta_\nu \Gamma \text{tr} \boldsymbol{\sigma}^0. \quad (119)$$

The eigenstrain contributed by the unequal Poisson's ratios is

$$\boldsymbol{\varepsilon}_\nu^* = -k(\mathbf{I} - \theta \boldsymbol{\varepsilon})^{-1} : \mathbf{I} \rightarrow -\frac{1-2\nu^0}{E^0} \theta_\nu \Gamma \text{tr} \boldsymbol{\sigma}^0 (\mathbf{I} - \boldsymbol{\varepsilon}) : \mathbf{I} \quad (120)$$

and the corresponding energy disturbance

$$U_\nu^* = \frac{V}{2} \boldsymbol{\sigma}^0 : \boldsymbol{\varepsilon}_\nu^* \rightarrow -\frac{2\pi a^3 \xi_{eq}}{3} \frac{1-2\nu^0}{E^0} \theta_\nu \text{tr} \boldsymbol{\sigma}^0 \left( \text{tr} \boldsymbol{\sigma}^0 - \frac{1+\nu^0}{1-\nu^0} \sigma_0 \right). \quad (121)$$

It should be noted that the deduced asymptotic eigenstrains in Eqs. (114) and (120) is based on the assumption that the prime eigenstrain is the dominant part of the eigenstrain. Thus, these asymptotic expressions is valid if  $|\nu^* - \nu^0| \ll 1$ .

## 7 Signs of Energy Disturbances

As discussed in Section 3 and Section 6, the energy disturbance of penny-shaped inhomogeneities is trivial unless the relative shear modulus  $\Gamma$  is sufficient small or large. In this section, we discuss the sign of energy disturbance of the penny-shaped inhomogeneity for the two cases. In view of Eqs. (61) and (115), the energy disturbance takes following signs, if  $\Gamma \rightarrow 0$ ,

$$U_\mu^* = V^{eq} u^{eq}, \quad u^{eq} \approx \frac{1}{2} \boldsymbol{\sigma}^0 : \mathbf{C}^* : \boldsymbol{\sigma}^0 > 0; \quad U_\nu^* \begin{cases} > 0 & \text{if } \nu^0 > \nu^* \\ = 0 & \text{if } \nu^0 = \nu^* \\ < 0 & \text{if } \nu^0 < \nu^* \end{cases} \quad (122)$$

for nonzero  $\boldsymbol{\sigma}^0$ . In the case  $\Gamma \rightarrow \infty$ , Eq. (72) can be rewritten as

$$U_\mu^* = V^{eq} u^{eq}, \quad u^{eq} \approx \frac{1}{2} \boldsymbol{\sigma}^0 : (\mathbf{C}^* - \mathbf{C}^0) : \boldsymbol{\sigma}^0 \leq 0 \quad (123)$$

since the tensor  $\mathbf{C}^0 - \mathbf{C}^*$  is positive semidefinite, as discussed in Appendix C. The sign of  $U_\nu^*$  for  $\Gamma \rightarrow \infty$ , as shown in Eq. (121), depends on not only the  $\nu^* - \nu^0$  but also the stress state.

The different signs of the energy disturbances  $U_\mu^*$  for  $\Gamma \rightarrow 0$  and  $\Gamma \rightarrow \infty$  just indicate a simple fact, through the variation of the complimentary energy, that a weak inhomogeneity ( $\Gamma < 1$ ) surrounded by a solid weakens the solid, and a strong inhomogeneity ( $\Gamma > 1$ ) strengthens the solid.

**7.1 Energy Release Rates.** The rate of the energy disturbance is

$$\dot{U}^* = \dot{U}^*(a, \Gamma, \mathbf{n}, \dots) = G_a \dot{a} + G_\Gamma \dot{\Gamma} + G_{\mathbf{n}} \dot{\mathbf{n}} + \dots \quad (124)$$

where

$$G_a = \frac{\partial U^*}{\partial a}, \quad G_\Gamma = \frac{\partial U^*}{\partial \Gamma}, \quad G_{\mathbf{n}} = \frac{\partial U^*}{\partial \mathbf{n}}, \dots \quad (125)$$

are the energy release rates due to the variations of  $a$ ,  $\Gamma$ ,  $\mathbf{n}$ ,  $\dots$ , respectively. Evidently

$$U^* = V^{eq} u^{eq} \Rightarrow G_a = S^{eq} u^{eq}, \quad (126a)$$

$$S^{eq} = \frac{\partial V^{eq}}{\partial a}. \quad (126b)$$

Especially, for a penny-shaped inhomogeneity and  $\xi_{eq} \ll 1$ ,

$$V^{eq} = \frac{4\pi}{3} a^3 \xi_{eq} \Rightarrow S^{eq} = 4\pi a^2 \xi_{eq}, \quad (127)$$

and for a crack

$$V^{eq} = \frac{4\pi}{3} a^3 \Rightarrow S^{eq} = 4\pi a^2. \quad (128)$$

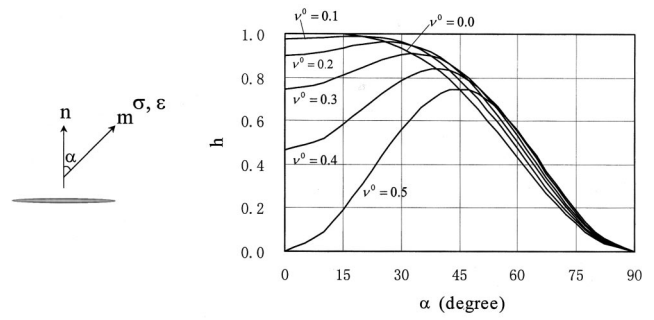


Fig. 5 The variation of  $h$  with respect to  $\alpha$  and  $\nu^0$

Thus,  $S^{eq}$  is exactly the surface area of the energy-based equivalent inclusion  $\Omega^{eq}$ . Equation (126b) shows that the sign of the energy release rate  $G_a$  is fully dependent on the sign of  $u^{eq}$ . Note that a penny-shaped inhomogeneity approaches a penny-shaped crack or anticrack if  $\Gamma \rightarrow 0$  or  $\Gamma \rightarrow \infty$ , as discussed in Section 4. Then, the different signs of  $G_a$  for  $\Gamma \rightarrow 0$  and  $\Gamma \rightarrow \infty$  are consistent with the fact that the energy release rates of cracks and anticracks have different signs, as pointed out by Hurtado [13].

**7.2 Apparent Young's Modulus of  $\mathbf{C}^*$ .** Both of the Eqs. (122) and (123) show that the compliance tensor  $\mathbf{C}^*$  plays a key role in the equivalent densities of the energy disturbances. It is worthy to discuss the property of the tensor  $\mathbf{C}^*$ . Consider a body subjected uniaxial stress  $\sigma$  in the direction  $\mathbf{m}$ . Here,  $\varepsilon$  denotes the normal strain in the same direction  $\mathbf{m}$  caused by the uniaxial stress  $\sigma$ . If the body is of the compliance tensor  $\mathbf{C}^0$ , obviously, the stress and strain are related by  $\sigma = E^0 \varepsilon$ . If the body is of the compliance tensor  $\mathbf{C}^*$ , then  $\sigma$  and  $\varepsilon$  are related by the apparent Young's modulus  $E$ ,

$$\sigma = E \varepsilon, \quad E = E(\mathbf{m}). \quad (129)$$

Consider the problem in the coordinate systems as shown in Fig. 2. If the direction  $\mathbf{m}$  coincides with the  $x_3$ -axis, the direction cosines of  $\mathbf{n}$  can be expressed in the global coordinate system,  $\mathbf{n} = \{n_1, n_2, n_3\}$ , where  $n_3 = \mathbf{m} \cdot \mathbf{n} = \cos \alpha$ . In view of Eq. (60), one obtains

$$\frac{1}{E(\mathbf{m})} = C_{3333} = \frac{1}{\mu^0} \left[ \frac{1-2\nu^0}{2(1-\nu^0)} N_{3333} + T_{3333} \right] \quad (130)$$

where  $N_{3333} = n_3^4 = \cos^4 \alpha$  and  $T_{3333} = n_3^2 - n_3^4 = \cos^2 \alpha \sin^2 \alpha$ , due to Eq. (10). Therefore,

$$E(\mathbf{m}) = E(\alpha) = \frac{E^0}{h}, \quad (131)$$

$$h = 2(1+\nu^0) \cos^2 \alpha \left[ \frac{1-2\nu^0}{2(1-\nu^0)} \cos^2 \alpha + \sin^2 \alpha \right].$$

The relation  $h = h(\alpha, \nu^0)$  is illustrated in Fig. 5. Evidently,  $0 \leq h \leq 1.0 \Rightarrow E \geq E^0$ .

## 8 Generic Damage Measurement

As shown Eq. (55), the size of the energy-based equivalent inclusion  $\Omega^{eq}$  generally takes the form

$$V^{eq} = \eta V. \quad (132)$$

For a penny-shaped inhomogeneity, its size is  $V = 4\pi/3 a^3 \xi$  with the singular factor  $\xi$ , so the product  $\xi \eta$  should be finite to ensure a finite  $V^{eq}$ . For an inhomogeneity of finite size, its Eshelby's equivalent eigenstrain should contain no singular factor, otherwise its energy disturbance will be trivial or infinitely large. Let's con-

**Table 1 Some characteristic quantities of different inhomogeneities**

Inhomogeneities	$\frac{V}{V^{sp}}$	$\frac{V^{eq}}{V^{sp}}$	$\eta$	$\Pi$
Penny-shaped inhomogeneity				
for $\Gamma \rightarrow 0, \xi_{eq} = \frac{\xi}{\Gamma} \ll 1$	$\xi$	$\xi_{eq}$	$\frac{1}{\Gamma}$	<b>S</b>
Penny-shaped inhomogeneity				
for $\Gamma \rightarrow \infty, \xi_{eq} = \xi \Gamma \ll 1$	$\xi$	$\xi_{eq}$	$\Gamma$	<b>S-I</b>
Penny-shaped crack	$\xi$	1	$\frac{1}{\xi}$	$c_1 \mathbf{S}_n + c_2 \mathbf{S}_t$
Spherical cavity	1	1	1	$\frac{3(1-\nu_0)}{7-5\nu^0} \left[ \frac{1-\nu^0}{2(1-2\nu^0)} \Pi + 5\mathbf{I} \right]$
Rigid sphere	1	1	1	$-\frac{3(1-\nu^0)}{2(4-5\nu^0)} \left( \frac{1-5\nu^0}{1+\nu^0} \Pi + 5\mathbf{I} \right)$

sider spherical inhomogeneities. If the inhomogeneity is a spherical cavity, its Eshelby's equivalent eigenstrain can be determined by

$$(\mathbf{I} - \tilde{\mathbf{S}}) : \boldsymbol{\varepsilon}^* = \boldsymbol{\varepsilon}^0 \Rightarrow \boldsymbol{\varepsilon}^* = \frac{3(1-\nu^0)}{7-5\nu^0} \left[ \frac{1-\nu^0}{2(1-2\nu^0)} \Pi + 5\mathbf{I} \right] : \boldsymbol{\varepsilon}^0 \quad (133)$$

where  $\tilde{\mathbf{S}}$  is the Eshelby's tensor of spherical inclusions expressed in Eq. (A6). If the inhomogeneity is a rigid sphere, its equivalent eigenstrain can be determined by

$$\tilde{\mathbf{S}} : \boldsymbol{\varepsilon}^* = -\boldsymbol{\varepsilon}^0 \Rightarrow \boldsymbol{\varepsilon}^* = -\frac{3(1-\nu^0)}{2(4-5\nu^0)} \left( \frac{1-5\nu^0}{1+\nu^0} \Pi + 5\mathbf{I} \right) : \boldsymbol{\varepsilon}^0. \quad (134)$$

Naturally,  $\eta=1$  for the inhomogeneities of finite size, and then  $V^{eq}=V$ . Generally speaking, the Eshelby's and energy-based equivalent eigenstrain can be expressed as

$$\boldsymbol{\varepsilon}^* = \eta \boldsymbol{\varepsilon}_{eq}^*, \quad \boldsymbol{\varepsilon}_{eq}^* = \Pi : \boldsymbol{\varepsilon}^0. \quad (135)$$

Some quantities related to different inhomogeneities are listed in Table 1, where  $V^{sp}=4\pi/3a^3$  is a reference size of a sphere with the radius  $a$ .

As shown in Table 1, all the  $\Pi$  tensors of the different inhomogeneities are of the same order of magnitude. Note that  $U^* = \frac{1}{2} V^{eq} \boldsymbol{\sigma}^0 : \boldsymbol{\varepsilon}^{eq} = 1/2 V^{eq} \boldsymbol{\sigma}^0 : \Pi : \boldsymbol{\varepsilon}^0$ . Therefore, the size  $V^{eq}$  of the energy-based equivalent inclusion  $\Omega^{eq}$  becomes the characteristic

geometric measurement of the energy disturbance caused by the presence of the inhomogeneity. Naturally, the dimensionless damage variable can be defined as (assuming only one single inhomogeneity)

$$\omega = \frac{V^{eq}}{V^0} \quad (136)$$

where  $V^0$  is the volume of a representative unit cell in the microstructure. If the inhomogeneity is a crack, the damage variable is identical to that of Budiansky and O'Connell [11] defined in Eq. (2) just scaled by  $4\pi/3$ . For an inhomogeneity of finite size,  $\eta=1$ , then

$$\omega = \frac{V^{eq}}{V^0} = \frac{V}{V^0} \quad (137)$$

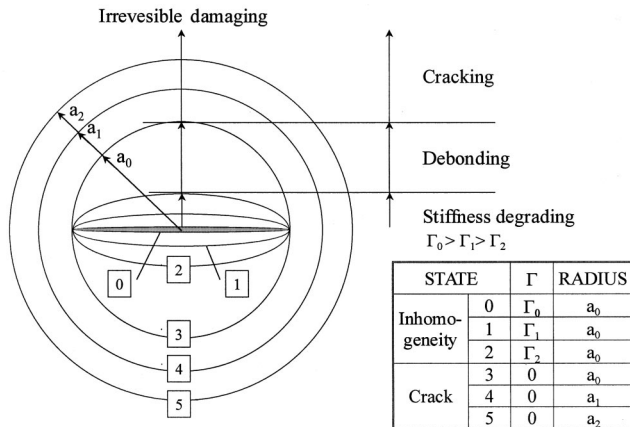
which is just Gurson's damage variable ([14]), the void volume fraction of a porous solid.

We call  $V^{eq}$  the generic damage measurement because it is applicable to all types of inhomogeneities: cavities, weak and strong inhomogeneities, and rigid inhomogeneities of infinitesimal or finite sizes. It is also convenient to depict the damaging processes of inhomogeneities by  $V^{eq}$ . Consider a penny-shaped inhomogeneity with the relative modulus  $\Gamma_0$  and radius  $a_0$  (state 0), as shown in Fig. 6. The damaging process would be: (a) stiffness degrading of the inhomogeneity, state 0 $\rightarrow$ 2; (b) debonding between the inhomogeneity and matrix, state 2 $\rightarrow$ 3; and (c) cracking, state 3 $\rightarrow$ 5. All the state variations, mechanically or geometrically, can be illustrated unitarily by a series of spheroids or spheres of the energy-based equivalent inclusions, as shown in Fig. 6. The irreversible energy dissipation implies that the current spheroid or sphere is always expanding during the damaging process.

As shown in Eq. (126),  $U^* = V^{eq} u^{eq} \Rightarrow G_a = S^{eq} u^{eq}$ . Similarly, the  $S^{eq}$ , which is the surface area of the energy-based equivalent inclusion  $\Omega^{eq}$ , is the generic measurement of the energy release rate for the radius propagation. Note that the surface area of a spheroid or sphere is four times of the area of the corresponding ellipse or circle in Fig. 6. From the thermodynamic point of view, the  $G_a$  is the driving force behind the radius propagation. Since the current spheroid or sphere is always expanding, the driving force is always increasing, too.

## 9 Conclusion

The Eshelby's problem generally results in complicated formulation even in a well-chosen local coordinate system. It looks quite difficult or even impossible to express the solutions explicitly in Eshelby's tensor as an entity in global coordinate system. The present work demonstrates that such an approach is possible



**Fig. 6 The evolution of the energy-based equivalent inclusion  $\Omega^{eq}$  during the damaging process of a penny-shaped inhomogeneity: stiffness degrading, debonding, and cracking**

at least for penny-shaped inhomogeneities including cracks. The deduced eigenstrain and their asymptotic expressions are all in tensorial form which help us gain deeper insight into their subtle relations.

In order to describe the macroscopic behavior of microdefects, this work focus on the tensorial representation and the analysis of energy disturbances caused by presence of microdefects. Furthermore, the so-called *energy-based equivalent inclusion method* is introduced, as a certain generalization of Eshelby's equivalent inclusion method. For the inhomogeneities of finite size, both methods are equivalent to each other. For the inhomogeneities of infinitesimal size, e.g., penny-shaped inhomogeneities, the energy-based equivalent inclusion method can effectively remove the singularities of the size and eigenstrain of the Eshelby's equivalent inclusion. The energy-based equivalent inclusion  $\Omega^{eq}$  is always of finite size, and may be geometrically different with the corresponding inhomogeneity, unlike the Eshelby's equivalent inclusion  $\Omega$ . For a penny-shaped inhomogeneity, the energy-based equivalent method cannot describe the local stress field but can accurately yield the energy disturbance, which is sufficient for damage modeling.

The energy-based equivalent inclusions of penny-shaped inhomogeneities are spheroids. Especially, the  $\Omega^{eq}$  of a crack is a sphere. The size of energy-based equivalent inclusions,  $V^{eq}$ , can be used as the generic damage measurement for all types of inhomogeneities: cavities, weak and strong inhomogeneities, and rigid inhomogeneities of infinitesimal or finite sizes. The  $V^{eq}$ -based volume fraction is consistent with the well-known damage variables of Budiansky and O'Connell [11] and Gurson [14].

## Acknowledgments

The work reported here was jointly supported by the National Science Foundation of China with grant No. 59879005 and the Austrian National Science Foundation under contract No. S08004-TEC.

## Appendix

**A Eshelby's Tensor.** When an strain  $\varepsilon_{ij}^*$  is prescribed in a finite subdomain  $\Omega$  in a homogeneous material  $D$  as shown in Fig. 1(a) and it is zero in the matrix  $D - \Omega$ , then  $\Omega$  is called an inclusion. This type of prescribed strain  $\varepsilon_{ij}^*$  is called "eigenstrain." If an ellipsoidal inclusion, as shown in Fig. 1(b), is prescribed by uniform eigenstrain  $\varepsilon_{ij}^*$ , the strain and stress fields become uniform for the interior points inside the inclusion ([1,2]). The elastic strain  $\varepsilon_{ij}$  in interior point is

$$\varepsilon_{ij} = S_{ijkl} \varepsilon_{kl}^* \quad (A1)$$

where  $S_{ijkl}$  is called Eshelby's tensor and has the symmetry

$$S_{ijkl} = S_{jikl} = S_{ijlk} \quad (A2)$$

The Eshelby's tensor of an ellipsoidal inclusion is generally associated with some standard elliptic integrals. For a spheroidal inclusion, i.e.,  $a_1 = a_2 = a$ , the Eshelby's tensor can be simplified as

$$S_{ijkl} = S_{ijkl}(\nu, \xi), \quad \xi = \frac{a_3}{a} \quad (A3)$$

where  $\nu$  is the Poisson's ratio of the material, for example,

$$S_{1111} = S_{2222} = \frac{1}{4(1-\nu)} \left[ \frac{3\xi^2}{2(\xi^2-1)} - \left( 1-2\nu - \frac{9}{4(\xi^2-1)} \right) g \right] \quad (A4)$$

$$S_{3333} = \frac{1}{2(1-\nu)} \left[ 1-2\nu + \frac{3\xi^2-1}{\xi^2-1} + \left( 1-2\nu + \frac{3\xi^2}{\xi^2-1} \right) g \right]$$

where

$$g = \begin{cases} \frac{\xi}{(\xi^2-1)^{3/2}} [\xi(\xi^2-1)^{1/2} - \coth^{-1} \xi] & \text{if } \xi > 1 \\ \frac{\xi}{(1-\xi^2)^{3/2}} [\cos^{-1} \xi - \xi(1-\xi^2)^{1/2}] & \text{if } \xi < 1 \end{cases} \quad (A5)$$

If  $\xi = 1$ , i.e.,  $a_1 = a_2 = a_3 = a$ , the Eshelby's tensor of the spherical inclusion is simply an isotropic tensor

$$S_{ijkl} = \frac{5\nu-1}{15(1-\nu)} \delta_{ij} \delta_{kl} + \frac{4-5\nu}{15(1-\nu)} (\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}) \quad (A6)$$

The spheroidal inclusion approaches a penny-shaped inclusion if  $\xi \ll 1$  or  $a \gg a_3$ , and then the Eshelby's tensor can be further simplified as

$$\begin{aligned} S_{1111} = S_{2222} &= \frac{13-8\nu}{32(1-\nu)} \pi \xi, & S_{3333} &= 1 - \frac{1-2\nu}{1-\nu} \frac{\pi}{4} \xi \\ S_{1122} = S_{2211} &= \frac{8\nu-1}{32(1-\nu)} \pi \xi, & S_{1133} = S_{2233} &= \frac{2\nu-1}{1-\nu} \frac{\pi}{8} \xi \\ S_{3311} = S_{3322} &= \frac{\nu}{1-\nu} \left( 1 + \frac{4\nu+1}{8\nu} \pi \xi \right) \\ S_{1313} = S_{2323} &= \frac{1}{2} \left( 1 + \frac{\nu-2}{1-\nu} \frac{\pi}{4} \xi \right), & S_{1212} &= \frac{7-8\nu}{32(1-\nu)} \pi \xi. \end{aligned} \quad (A7)$$

All other nonzero components are obtained by the cyclic permutation of (1, 2, 3). If  $\xi \rightarrow 0$  or  $a_3 \rightarrow 0$ , the Eshelby's tensor approaches the limit

$$\begin{aligned} S_{2323} = S_{2332} = S_{3223} = S_{3232} &= 0.5, \\ S_{1313} = S_{1331} = S_{3113} = S_{3131} &= 0.5, \\ S_{3311} = S_{3322} &= \nu/(1-\nu), \\ S_{3333} &= 1, \text{ and all other } S_{ijkl} = 0. \end{aligned} \quad (A8)$$

**B Eshelby's Equivalent Inclusion Method.** Eshelby [2] solved the problem of an ellipsoidal inhomogeneity by *equivalent inclusion method*. Consider an infinitely extended material with the elastic moduli  $D_{ijkl}^0$ , containing an ellipsoidal inhomogeneity domain  $\Omega$  as shown in Fig. 1(b) with the elastic moduli  $D_{ijkl}^*$ . Let us denote the applied stress at infinity by  $\sigma_{ij}^0$  and the corresponding strain by  $\varepsilon_{ij}^0$ .  $\varepsilon_{ij}^*$  has been introduced here arbitrarily in order to simulate the inhomogeneity problem by use of the inclusion method. Such an eigenstrain is called an equivalent eigenstrain. The stress disturbance and the strain disturbance are denoted by  $\sigma_{ij}$  and  $\varepsilon_{ij}$ , respectively. The total stress or actual stress is  $\sigma_{ij}^0 + \sigma_{ij}$ , and the total strain is  $\varepsilon_{ij}^0 + \varepsilon_{ij}$ . Hooke's law is written as

$$\begin{aligned} \sigma_{ij}^0 + \sigma_{ij} &= D_{ijkl}^* (\varepsilon_{kl}^0 + \varepsilon_{kl}) \quad \text{in } \Omega \\ \sigma_{ij}^0 + \sigma_{ij} &= D_{ijkl}^0 (\varepsilon_{kl}^0 + \varepsilon_{kl}) \quad \text{in } D - \Omega. \end{aligned} \quad (B1)$$

Now consider an infinitely extended homogeneous material with the elastic  $D_{ijkl}^0$  everywhere, containing domain  $\Omega$  with an eigenstrain  $\varepsilon_{ij}^*$ . When this homogeneous material is subjected to the applied strain  $\varepsilon_{ij}$  at infinity, the resulting total stress and strain, respectively, are  $\sigma_{ij}^0 + \sigma_{ij}$ ,  $\varepsilon_{ij}^0 + \varepsilon_{ij}$ , and  $\varepsilon_{ij}^0 + \varepsilon_{ij} - \varepsilon_{ij}^*$  in  $\Omega$ . Then, Hooke's law yields

$$\begin{aligned} \sigma_{ij}^0 + \sigma_{ij} &= D_{ijkl}^0 (\varepsilon_{kl}^0 + \varepsilon_{kl} - \varepsilon_{kl}^*) \quad \text{in } \Omega \\ \sigma_{ij}^0 + \sigma_{ij} &= D_{ijkl}^0 (\varepsilon_{kl}^0 + \varepsilon_{kl}) \quad \text{in } D - \Omega. \end{aligned} \quad (B2)$$

The necessary and sufficient condition for the equivalence of the stresses and strains in the above two problems of inhomogeneity and inclusion is

$$D_{ijkl}^* (\varepsilon_{kl}^0 + \varepsilon_{kl}) = D_{ijkl}^0 (\varepsilon_{kl}^0 + \varepsilon_{kl} - \varepsilon_{kl}^*) \quad \text{in } \Omega. \quad (B3)$$

If  $\sigma_{ij}^0 = D_{ijkl}^0 \varepsilon_{kl}^0$  is a uniform stress,  $\varepsilon_{ij}^*$  is also uniform in  $\Omega$  and can be determined by Eq. (A1). Therefore,

$$\hat{\sigma}_{ij} = D_{ijkl}^* (\varepsilon_{kl}^0 + S_{klmn} \varepsilon_{mn}^*) = D_{ijkl}^0 (\varepsilon_{kl}^0 + S_{klmn} \varepsilon_{mn}^* - \varepsilon_{kl}^*) \quad (B4)$$

from which  $\varepsilon_{ij}^*$  is determined. Evidently,  $\hat{\sigma}_{ij} = \sigma_{ij}^0 + \sigma_{ij}$  in  $\Omega$ . The elastic strain energy for this system is

$$\begin{aligned} U &= \frac{1}{2} \int_D (\sigma_{ij}^0 + \sigma_{ij}) (\varepsilon_{ij}^0 + \varepsilon_{ij}) dD \\ &= \frac{1}{2} \int_D \sigma_{ij}^0 \varepsilon_{ij}^0 dD + \frac{1}{2} \int_D \sigma_{ij}^0 \varepsilon_{ij}^* dD = U^0 + U^*. \end{aligned} \quad (B5)$$

In this paper,  $U^*$  is also termed the energy disturbance,

$$U^* = \frac{1}{2} \int_D \sigma_{ij}^0 \varepsilon_{ij}^* dD. \quad (B6)$$

**C Positive Definiteness of the  $\mathbf{C}-\mathbf{C}^*$  Tensor.** The  $\mathbf{C}$  and  $\mathbf{C}^*$  tensors are defined in Eqs. (22) and (23), respectively. In the local coordinate system  $x_1, x_2, x_3$ , as shown in Fig. 2, the components of the stress tensor  $\boldsymbol{\sigma}$  are denoted

$$\boldsymbol{\sigma} = \begin{bmatrix} \sigma_x & \tau_{xy} & \tau_{xz} \\ \tau_{yx} & \sigma_y & \sigma_{yz} \\ \tau_{zx} & \tau_{zy} & \sigma_z \end{bmatrix} \quad (C1)$$

where  $\sigma_x, \sigma_y, \sigma_z$  are along the  $x_1, x_2, x_3$  axes, respectively. In view of Eq. (23),

$$u^* = \frac{1}{2} \boldsymbol{\sigma} : \mathbf{C}^* : \boldsymbol{\sigma} = \frac{w}{2E} \sigma_z^2 + \frac{1}{2\mu} (\tau_{zx}^2 + \tau_{zy}^2) \quad (C2)$$

where

$$w = \frac{(1-2\nu)(1+\nu)}{1-\nu}, \quad 0 \leq w \leq 1. \quad (C3)$$

Note that

$$\begin{aligned} u_3 &= \frac{1}{2} \boldsymbol{\sigma}_3 : \mathbf{C} : \boldsymbol{\sigma}_3 = \frac{1}{2E} \sigma_z^2 + \frac{1}{2\mu} (\tau_{zx}^2 + \tau_{zy}^2), \\ \boldsymbol{\sigma}_3 &= \begin{bmatrix} 0 & 0 & \tau_{xz} \\ 0 & 0 & \tau_{yz} \\ \tau_{zx} & \tau_{zy} & \sigma_z \end{bmatrix}. \end{aligned} \quad (C4)$$

The effect of the  $\mathbf{C}^*$  tensor can be shown by comparing Eqs. (C2) and (C4). Evidently,  $u_3 \geq u^*$ ,  $u_1^* = 1/2(\boldsymbol{\sigma} - \boldsymbol{\sigma}_3) : \mathbf{C}^* : (\boldsymbol{\sigma} - \boldsymbol{\sigma}_3) = 0$ . Note that

$$\begin{aligned} u &= \frac{1}{2} \boldsymbol{\sigma} : \mathbf{C} : \boldsymbol{\sigma} = \frac{1}{2} \boldsymbol{\sigma}_n : \mathbf{C} : \boldsymbol{\sigma}_n + \frac{1}{2\mu} (\tau_{xy}^2 + \tau_{zx}^2 + \tau_{zy}^2), \\ \boldsymbol{\sigma}_n &= \begin{bmatrix} \sigma_x & 0 & 0 \\ 0 & \sigma_y & 0 \\ 0 & 0 & \sigma_z \end{bmatrix}. \end{aligned} \quad (C5)$$

Therefore

$$\begin{aligned} u - u^* &= \frac{1}{2} \boldsymbol{\sigma} : (\mathbf{C} - \mathbf{C}^*) : \boldsymbol{\sigma} = \frac{1}{2} \boldsymbol{\sigma}_n : \mathbf{C} : \boldsymbol{\sigma}_n + \frac{1}{2\mu} \tau_{xy}^2 - \frac{w}{2E} \sigma_z^2 \\ &\geq \frac{1}{2} \boldsymbol{\sigma}_n : \mathbf{C} : \boldsymbol{\sigma}_n - \frac{w}{2E} \sigma_z^2 \\ &= \frac{1}{2E} \{\sigma_x, \sigma_y, \sigma_z\} \begin{bmatrix} 1 & -\nu & -\nu \\ -\nu & 1 & -\nu \\ -\nu & -\nu & 1-w \end{bmatrix} \begin{Bmatrix} \sigma_x \\ \sigma_y \\ \sigma_z \end{Bmatrix}. \end{aligned} \quad (C6)$$

The  $3 \times 3$  matrix has three eigenvalues:

$$\lambda_1 = 0, \quad \lambda_2 = \frac{3\nu^2 - 2\nu + 1}{1-\nu} > 0, \quad \lambda_3 = 1 + \nu > 0 \quad \text{for } 0 \leq \nu \leq 0.5, \quad (C7)$$

so the matrix is positive semidefinite. Therefore, the tensor  $\mathbf{C}-\mathbf{C}^*$  is also a positive semidefinite tensor.

## References

- [1] Mura, T., 1987, *Micromechanics of Defects in Solids*, 2nd Ed., Martinus Nijhoff, Dordrecht, The Netherlands.
- [2] Eshelby, J. D., 1957, "The Determination of the Elastic Field of an Ellipsoidal Inclusion and Related Problems," *Proc. R. Soc. London, Ser. A*, **A241**, pp. 376–396.
- [3] Hurtado, J. A., Dundurs, J., and Mura, T., 1996, "Lamellar Inhomogeneities in a Uniform Stress Field," *J. Mech. Phys. Solids*, **44**, pp. 1–21.
- [4] Homocovschi, D., and Dascalu, C., 1996, "Uniform Asymptotic Solutions for Lamellar Inhomogeneities in Plane Elasticity," *J. Mech. Phys. Solids*, **44**, pp. 1–21.
- [5] Zhao, Y. H., and Weng, G. J., 1996, "Plasticity of a Two-Phase Composite With Partially Debonded Inclusions," *Int. J. Plast.*, **12**, pp. 781–804.
- [6] Shafiro, B., and Kachanov, M., 1999, "Solids With Non-Spherical Cavities: Simplified Representations of Cavity Compliance Tensors and the Overall Anisotropy," *J. Mech. Phys. Solids*, **47**, pp. 877–898.
- [7] Kachanov, M., 1980, "A Continuum Model of Medium With Cracks," *J. Eng. Mech.*, **106**, pp. 1039–1051.
- [8] Swoboda, G., and Yang, Q., 1999, "An Energy-Based Damage Model of Geomaterials—I. Formulation and Numerical Results," *Int. J. Solids Struct.*, **36**, pp. 1719–1734.
- [9] Swoboda, G., and Yang, Q., 1999, "An Energy-Based Damage Model of Geomaterials—II. Deduction of Damage Evolution Laws," *Int. J. Solids Struct.*, **36**, pp. 1735–1755.
- [10] Yang, Q., Zhou, W. Y., and Swoboda, G., 1999, "Micromechanical Identification of Anisotropic Damage Evolution Laws," *Int. J. Fract.*, **98**, pp. 55–76.
- [11] Budiansky, B., and O'Connell, R. J., 1976, "Elastic Moduli of a Cracked Solid," *Int. J. Solids Struct.*, **12**, pp. 81–97.
- [12] Yang, Q., 1996, "Numerical Modeling for Discontinuous Geomaterials Considering Damage Propagation and Seepage," Ph.D. thesis, Faculty of Architecture and Civil Engineering, University of Innsbruck, Austria.
- [13] Hurtado, J. A., 1997, "Estudio de fibras de forma laminar en un campo de tensión uniforme: Grietas, anti-grietas y cuasi-grietas," *Anales de Mecánica de la Fractura*, **14**, pp. 105–110.
- [14] Gurson, A. L., 1977, "Continuum Theory of Ductile Rupture by Void Nucleation and Growth, I. Yield Criteria and Flow Rules for Porous Ductile Media," *ASME J. Eng. Mater. Technol.*, **99**, pp. 2–15.



# On the Simulation of Discontinuous Functions

J. D. Turner

Mem. ASME

National Advanced Driving Simulator  
and Simulation Center,  
The University of Iowa,  
2401 Oakdale Boulevard,  
Iowa City, IA 52242-5003  
e-mail: jturner@nads-sc.uiowa.edu

*Discontinuous function constraints arise during the calculation of surface contact, stiction, and friction effects in studies of the behavior of complex systems. These nonlinear effects are mathematically defined by inequality constraints of the form  $0 \geq g(x(t), t)$ . The unknown in the problem is the time,  $t^*$ , when the equality condition is reached. This paper presents an exact solution for  $t^*$ , which is obtained by introducing a slack variable that replaces time as the independent variable, leading to an extended state-space that is noniteratively integrated to the constraint surface. Several applications are presented to demonstrate the method. [DOI: 10.1115/1.1387022]*

## 1 Introduction

Discontinuous system dynamics models describe many practical problems in science and engineering. Mathematically, the discontinuous system models are modeled by inequality constraints. As shown in Fig. 1 the motion is divided into different motion phases, where different dynamical models are valid. Computationally the problem is to find the time,  $t^*$ , when the old motion phase is replaced by a new motion phase (Carver [1], Ellison [2], and Enright et al. [3]). Examples of discontinuous constraints include: (i) the height above a surface, (ii) the switching of the sign of a velocity component, (iii) impact phenomena, (iv) hysteresis, (v) structure varying systems, or (vi) a force component exceeding a critical value. Detecting when a discontinuous event takes place is straightforward; one simply monitors sign changes in the constraints. The challenge is finding the time,  $t^*$ , because it is implicitly defined as a function of the constraint condition. Finding  $t^*$  is important, because the governing differential equations are assumed to change discontinuously at  $t^*$ .

Several conventional approaches are presented for handling nonlinear problems subject to discontinuous inequality constraint conditions. The strengths and weaknesses of these approaches are discussed. The main contribution of this paper is the presentation of a new method for noniteratively localizing  $t^*$ . This approach permits an accurate integration of the response of discontinuous systems. The advantages of the new method are that (1) the integration algorithm never passes over a singular event (thereby avoiding numerical instabilities), (2) no iteration is required, (3) conventional integration routines can be used, and (4) little computational overhead is required to implement the algorithm. Discontinuous system behaviors that can be handled include material properties, nonlinear behaviors, geometry, kinematic effects, impact, intermittent contact, and other sources.

The paper consists of six major sections. Current approaches for handling discontinuous system behaviors are presented in Section 2. The slack variable algorithm is presented in Section 3. This section covers mathematical models, the slack variable algorithm, the differential equation for  $dt/ds$ , and the extended state-space model. Two applications are presented in Section 4. One model consists of a simple one-dimensional problem that can be handled analytically. The second application considers a toy woodpecker problem that is highly nonlinear in its behavior because the number of degrees-of-freedom vary and the system is subjected to

multiple impulsive inputs. A strategy is presented for handling problems characterized by intermittent contact, where multiple constraints can be active at one time in Section 5. Conclusions are presented in Section 6.

## 2 Current Approaches

Several techniques have been used for handling problems with nonlinear and discontinuous behaviors. The computational problem is that standard numerical integration algorithms have been derived under the assumption that the variables being integrated are continuous through all the derivatives being sampled (Filippov [4], Halin [5], and Pfeiffer [6]). Discontinuous problems, however, violate this basic assumption by introducing discontinuous changes into the first or higher derivatives of the system models. (see Fig. 2(a)). As a result, it is not surprising that standard numerical methods have problems handling discontinuous problems. A further complication is that a discontinuous event can occur at any time. As a result, algorithms cannot be designed in advance to change models at predetermined times to avoid discontinuous system behaviors.

A discontinuous event is detected during an integration time step by monitoring sign changes in the inequality constraints (see Fig. 3). The most naïve strategy for handling nonlinear problems is simply to ignore the existence of the discontinuous system changes. A potential risk of this approach is that the integration algorithm can become unstable or be subject to unknown large integration errors. Clearly, this approach is unacceptable for applications where high-fidelity behavior predictions are required.

The ideal solution is to noniteratively integrate the solution to the switch time  $t^*$ , change the equations being integrated, and restart the integration. This approach avoids having the numerical integration algorithm sample the discontinuous changes in the governing differential equations. The challenge is finding  $t^*$ , because it is implicitly defined by the inequality constraint and the nonlinear nature of the governing differential equations.

Four approaches are frequently used for simulating systems with discontinuous behaviors. First, one can use a very small step-size or variable step-size control algorithm. This approach brings the system dynamics very close to  $t^*$ . This approach is successful for many applications; however, very long run times are possible for sensitive problems. Second, root finding strategies or derivative-based iterative algorithms can be introduced for local-

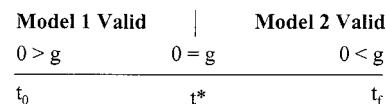
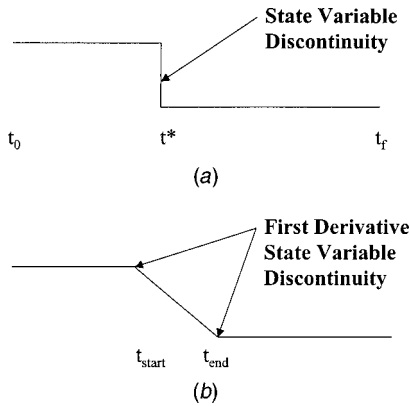
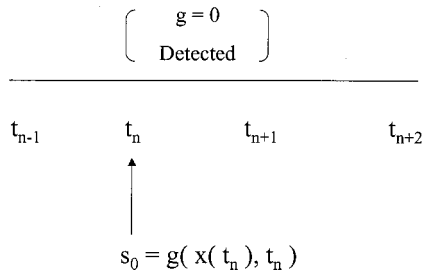


Fig. 1 Discontinuous model motion phases

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 12, 2000; final revision, Apr. 16, 2001. Associate Editor: N. C. Perkins. Discussion on the paper should be addressed to the 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.



**Fig. 2 (a) State variable discontinuity, (b) smoothed state variable discontinuity**



**Fig. 3 Discontinuity detected in the time interval and slack variable initial condition**

izing  $t^*$ . This approach increases algorithm complexity, and many iterations may be required for predicting  $t^*$ . Third, a low-order low-accuracy integration algorithm that is insensitive to the presence of a discontinuity can be used. The risk of this approach is that large undetected errors can develop in the integrated solution, yielding meaningless results. Fourth, smoothing algorithms, which attempt to reduce the strength of the discontinuity by pushing the discontinuous changes in the system behaviors into higher derivatives, are introduced. Figures 2(a) and 2(b) present an example of a control-smoothing algorithm. In this example, the instantaneous jump nonlinearity is replaced by a transition region that spreads the discontinuous change over a finite time interval (see Fig. 2(b)). The problem with this approach is that the real physics is not modeled and that the resulting solution may be meaningless. In all four cases the normal guarantees of solution accuracy are potentially compromised. All of these topics are thoroughly reviewed in Chapter 6 of Soellner and Fuhrer [7].

In summary, the basic problem is that  $t^*$  is implicitly defined by the constraint and the nonlinear character of the governing equations. The implicit nature of the problem makes the problem hard to deal with numerically. What is well known is the change in the inequality constraint that is required to reach the equality condition that defines  $t^*$ . This observation motivates a change of variables solution technique in Section 3. The new algorithm replaces time as the independent variable, with a slack variable that measures the distance to the equality condition for the constraint. The advantage of this approach is that the slack variable problem formulation permits the nonlinear system to be noniteratively integrated to the constraint surface.

### 3 Slack Variable Method

During a simulation, when a sign change is detected in a constraint, a change of variables is introduced that replaces the independent variable time,  $t$ , with an exact distance-like slack variable,

$s$ , that measures the distance to the equality constraint condition (see Fig. 3). The slack variable is described as a distance-like variable because in different problem formulations  $s$  can have units of length, velocity, force, work, potential, or any valid generalized coordinate. There are two significant benefits of this transformation. First, the inequality constraint is converted into an equality constraint. Second, the new equality constraint defines a functional relationship between time and the slack variable. The key step in the algorithm is that the functional relationship between time and the slack variable can be differentiated to provide a differential equation for  $dt/ds$ . The equation for  $dt/ds$  permits the original *implicit*  $n$ -dimensional problem for solving for  $t^*$  to be embedded in an *explicit*  $(n+1)$ -dimensional problem. The extended state-space model permits time to be integrated as a function of the slack variable, thereby eliminating the need for an iterative root solving procedure.

In the  $(n+1)$ -dimension problem formulation time,  $t$ , is replaced with the slack variable,  $s$ , as the independent variable, so that  $t=t(s)$ . The limits of integration for the  $(n+1)$ -dimensional problem formulation are defined as  $s=s_0$  to  $s=0$ , where the upper limit of integration,  $s=0$ , denotes that the system is on the constraint surface at the end of the integration. With the state on the constraint surface, the discontinuous changes in the system dynamics are introduced. The integration process is then restarted by using  $x(t^*(0))$  as the initial condition for the numerical method. The integration process then continues until either a new discontinuity is encountered or the end of the integration interval has been reached. The flow diagram for the algorithm is presented in Fig. 4.

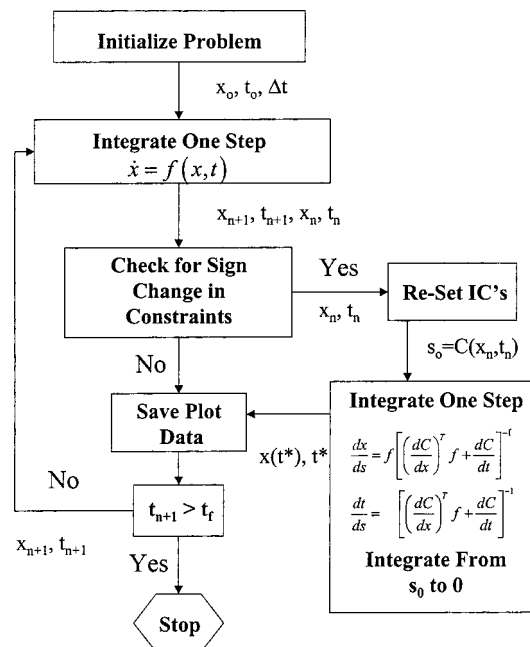
**3.1 Mathematical Model.** The constrained equations of motion for a physical system can be cast in the first-order form

$$\frac{dx}{dt} = f(x, t), \quad x_0 = b \quad (1a)$$

where

$$C(x(t), t) \geq 0 \quad (1b)$$

**Flow Diagram for Slack Variable Integration for Discontinuous Functions**



**Fig. 4 Flow diagram for the slack variable integration algorithm**

where  $x$  denotes the  $n \times 1$  state vector,  $x_0$  denotes the initial condition vector for the state,  $t$  denotes time,  $d(\cdot)/dt$  denotes the derivative with respect to time, and  $C(\cdot)$  defines the constraint surface for the inequality constraint.

Equation (1b) is monitored during each time interval  $[t_n, t_n + h]$  for a change in sign for  $C(\cdot)$  ([7]) (see Fig. 3). As long as  $C(\cdot) \neq 0$ , the numerical integration process is straightforward and accurate. Detecting the sign change for  $C(\cdot)$  is important because numerical algorithms are not designed to handle discontinuous changes in the derivatives being sampled. Section 3.2 introduces a slack variable algorithm that permits  $t^*$  to be determined noniteratively.

**3.2 Slack Variable Algorithm.** Assuming that a sign change has been detected during the current integration time step (i.e.,  $[t_n, t_n + h]$ ), the solution for the pair  $(x, t)$  is reset to  $(x_n, t_n)$ . The inequality constraint of Eq. (1b) is *locally redefined* as an equality constraint by introducing a slack variable,  $s$ , leading to

$$s - C(x(t), t) = 0 \quad (2)$$

where  $s$  represents a distance-like measure to the equality condition of Eq. (1b). The initial value for  $s$  is defined as  $s_0$ , where  $s_0 = C(x(t_n), t_n)$ . The desired value for the slack variable is  $s = 0$ , which corresponds to the condition  $C(x(t^*), t^*) = 0$ , where  $t^*$  is the unknown time for the discontinuous event.

**3.3.1 Change of Variables Transformation.** Equation (2) defines a functional relationship between  $t$  and  $s$ . With  $t$  as the independent variable, Eq. (1) is hard to solve. However, if  $s$  becomes the independent variable then Eq. (1) is easy to solve. This observation motivates us to introduce  $s$  as the independent variable by assuming that  $t = t(s)$  in Eq. (2), leading to

$$s = C(x(t(s)), t(s)). \quad (3)$$

Two steps are required for transforming Eq. (1) so that  $t = t(s)$ . First, Eq. (3) is used to create a differential equation for  $dt/ds$ . Second, Eq. (1a) is transformed so that  $s$  becomes the independent variable. For these transformations to be well defined, it is assumed that Eq. (3) has continuous first-order partial derivative for  $s$ ,  $x$ , and  $t$ . Equation (3) is used in Section 3.3.2 for deriving a differential equation for  $dt/ds$ .

**3.3.2 Differential Equation for  $t(s)$ .** Assuming that Eq. (3) possesses the required continuous partial derivatives, a differential equation is obtained for  $dt/ds$ , by using the chain rule of calculus to differentiate Eq. (3) with respect to  $s$ , one obtains

$$1 = \left( \frac{\partial C}{\partial x} \right)^T \frac{dx}{dt} \frac{dt}{ds} + \frac{\partial C}{\partial t} \frac{dt}{ds}$$

where the solution for  $dt/ds$  follows as

$$\frac{dt}{ds} = \left[ \left( \frac{\partial C}{\partial x} \right)^T \frac{dx}{dt} + \frac{\partial C}{\partial t} \right]^{-1}. \quad (4)$$

The solution for Eq. (4) is valid as long as the  $[*] \neq 0$ . Equation (4) is further simplified by recalling Eq. (1a), leading to

$$\frac{dt}{ds} = \left[ \left( \frac{\partial C}{\partial x} \right)^T f + \frac{\partial C}{\partial t} \right]^{-1}. \quad (5)$$

Equation (5) defines the rate of change of time as a function of the distance to the constraint surface. What is still unknown is  $dx/ds$ , which defines how the state changes as a function of the distance to the constraint surface. Equations (1a) and (5) are used to define an extended state-space model in Section 3.3.3.

**3.3.3 Extended State-Space Model.** With  $s$  as the new independent variable an extended  $(n+1)$ -dimensional state-space is defined by the following variable:

$$y(s) = [x(t(s)), t(s)] \quad (6)$$

where  $y$  is the  $(n+1) \times 1$  extended state-space model and  $t(s)$  is now treated as a dependent variable. The initial condition for  $t$  is  $t(s_0) = t_n$  (see Fig. 3). The initial condition for  $y$  is  $y(s_0) = [x(t_n), t_n]$ . The interval for the slack variable integration is  $s \in [s_0, 0]$ .

Two steps are required for defining the differential equation for  $y$ . First, from Eq. (5) the differential equation  $dt/ds$  is already available. Second, the chain rule of calculus is applied to transform Eq. (1a) from a function of  $t$  to a function of  $s$ , as follows:

$$\frac{dx}{dt} = \frac{dx}{ds} \frac{ds}{dt} = f$$

leading to the transformed equation

$$\frac{dx}{ds} = f \frac{dt}{ds} = f \left[ \left( \frac{\partial C}{\partial x} \right)^T f + \frac{\partial C}{\partial t} \right]^{-1}. \quad (7)$$

Equations (5) and (7) provide the required differential equations for solving the slack variable algorithm. By combining Eqs. (5) and (7), the extended state-space differential equations can be cast in the form

$$\frac{dy}{ds} = \begin{pmatrix} \frac{dx}{ds} \\ \frac{dt}{ds} \end{pmatrix} = \begin{pmatrix} f \\ 1 \end{pmatrix} \left[ \left( \frac{\partial C}{\partial x} \right)^T f + \frac{\partial C}{\partial t} \right]^{-1} \quad (8)$$

which is subject to the following initial conditions for  $x(s)$  and  $t(s)$ :

$$\begin{pmatrix} x(s_0) \\ t(s_0) \end{pmatrix} = \begin{pmatrix} x(t_n) \\ t_n \end{pmatrix}. \quad (9)$$

The solution for  $y(s)$  is integrated to the constraint surface by solving the following set of equations:

**Table 1 Equations for time localization algorithm for discontinuous functions**

**Original State-Space Model:**

Independent variable:  $t$

(Nominal governing equations)

**Extended State-Space Model:**

Independent variable:  $s$

$$x(t) = x(0) + \int_{t_0}^t f dt$$

subject to:  $0 \geq C(x(t), t)$

iterative methods required to find  $t^*$

$$\begin{pmatrix} x(0) \\ t(0) \end{pmatrix} = \begin{pmatrix} x(s_0) \\ t(s_0) \end{pmatrix} + \int_{s_0}^0 \begin{pmatrix} f \\ 1 \end{pmatrix} \left[ \left( \frac{\partial C}{\partial x} \right)^T f + \frac{\partial C}{\partial t} \right]^{-1} ds$$

Non-iterative method to find  $t(0) = t^*$

(Equations for locally integrating to a constraint surface.)

Note  $t(0)$  is the discontinuous event time.)

$$\begin{pmatrix} x(0) \\ t(0) \end{pmatrix} = \begin{pmatrix} x(s_0) \\ t(s_0) \end{pmatrix} + \int_{s_0}^0 \begin{pmatrix} f \\ 1 \end{pmatrix} \left[ \left( \frac{\partial C}{\partial x} \right)^T f + \frac{\partial C}{\partial t} \right]^{-1} ds. \quad (10)$$

The left-hand side of Eq. (10) provides  $x(0)$  on the constraint surface and  $t(0)=t^*$ , which is the originally unknown discontinuous event time. Since  $s_0$  is typically small, a single integration step in Eq. (9) is all that is generally required. The left-hand side of Eq. (10) also defines the initial conditions for restarting the integration of Eq. (1a) after the discontinuous changes have been introduced into the differential equations (see Fig. 4).

Equations (1a) and (10) provide a complete algorithm for numerically integrating the response of systems subject to discontinuous behaviors. The governing equations are summarized in Table 1. This algorithm is particularly useful for problems characterized by friction, stiction, variable topology, and multiple/intermittent contact dynamics.

#### 4 Applications

Two applications are presented that describe how the slack variable transformation can be applied in different situations. The first problem considers a one-dimensional linear problem where the constraint depends on a specific location being reached. The second problem considers a generalization to a variable topology problem with two to four degrees-of-freedom, while being subject to multiple discontinuous impulsive events.

**4.1 One-Dimensional Problem.** These ideas are made more concrete by considering the following problem. Assuming that Eq. (1a) is defined by the following one-degree-of-freedom system

$$\dot{x} = ax, \quad x_0 = b/10; \quad a, b \geq 0$$

where the inequality constraint of Eq. (1b) becomes

$$0 \geq b - x.$$

The inequality constraint requires  $x$  to be less than  $b$ . Introducing the slack variable,  $s$ , the transformed inequality constraint is cast in the form

$$s = b - x.$$

At  $t=t_0$ , from the initial condition for  $x_0$ , it follows that the distance to the constraint surface is  $s_0=9b/10$ . Assuming that  $s$  replaces  $t$  as the independent variable (i.e.,  $t=t(s)$ ), the constraint is transformed as follows:

$$s = b - x(t(s)). \quad (11)$$

The differential equation for  $t(s)$  is obtained by differentiating Eq. (11) with respect to  $s$ , leading to

$$1 = \frac{ds}{dx} \frac{dx}{dt} \frac{dt}{ds} = (-1) * (ax) * \frac{dt}{ds}$$

which yields the following differential equation for  $t(s)$ :

$$\frac{dt}{ds} = \frac{-1}{ax}. \quad (12)$$

From Eqs. (10) and (12) the general solution for the extended state-space is cast in the form

$$\begin{pmatrix} x(0) \\ t(0) \end{pmatrix} = \begin{pmatrix} b/10 \\ 0 \end{pmatrix} + \int_{9b/10}^0 \begin{pmatrix} ax \\ 1 \end{pmatrix} [(-1)(ax) + 0]^{-1} ds$$

or

$$\begin{pmatrix} x(0) \\ t(0) \end{pmatrix} = \begin{pmatrix} b/10 \\ 0 \end{pmatrix} + \int_{9b/10}^0 \begin{pmatrix} -1 \\ -1/ax \end{pmatrix} ds.$$

Integrating the trivial top equation leads to

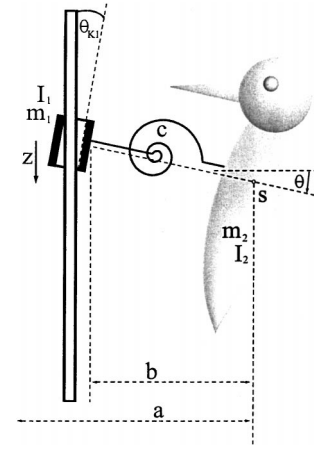


Fig. 5 Toy woodpecker problem

$$x(0) = b/10 + \int_{9b/10}^0 (-1) ds = b$$

where  $x(0)=b$  is the desired boundary condition for  $x$ . Leaving only the second integral to be evaluated as

$$t(0) = \frac{-1}{a} \int_{9b/10}^0 \frac{ds}{x} = \frac{-1}{a} \int_{9b/10}^0 \frac{ds}{b-s} = - \left( -\ln(b) + \ln\left(\frac{b}{10}\right) \right) / a = \ln(10)/a$$

where the equality constraint has been used to simplify the integral. The final result is that  $t(0)=\ln(10)/a$  is the predicted time for constraint of Eq. (11) to be identically zero. This prediction can be checked in this special case by analytically integrating the original differential equation and imposing the initial condition, yielding the solution

$$x(t) = x_0 e^{at} = \frac{b}{10} e^{at}.$$

This equation is solved for the time when  $x=b$ , yielding

$$t = \ln(10)/a$$

which agrees with the prediction obtained from integrating the extended state-space formulation.

**4.2 A Toy Woodpecker Modeling Problem.** The algorithm of Section 3 is applied to the problem of modeling the motions of a toy woodpecker that is attached to a sleeve, while it slides down a rod. The problem nonlinearities arise because (1) the toy experiences multiple impulse loads, and (2) the number of degrees-of-freedom are either free or fixed depending on the angular motion of the toy. This problem is taken from Pfeiffer [6] and Soellner and Führer [7] (see Fig. 5). The toy woodpecker model consists of

- a rod where the woodpecker slides down,
- a sleeve that glides down the rod with some play allowed, and
- a the woodpecker is connected to the sleeve by a torsional spring.

The nonlinear system behaviors arise from several sources. The system has two degrees-of-freedom. One rotational,  $\theta$ , defines the angle between the rod and the sleeve. One translational,  $z$ , defines the motion of the combined sleeve/woodpecker when sliding motions are possible down the rod. The complexity of the problem arises because the sleeve sliding motion stops for the sleeve/woodpecker system when  $|\theta| > |\theta_{K1}|$ . Another complexity is that the system experiences an impulsive load when the woodpecker strikes the rod when  $\theta = -\theta_{K2}$ , where  $\theta_{K1} < \theta_{K2}$ . The sleeve mo-



tion is modeled as a massless joint (Pfeiffer [6]). Only small oscillations are considered. The governing differential equations are Equation 1: ( $\theta$  is the only degree-of-freedom,  $z=\text{constant}$ ,  $dz/dt=0$ )

$$(I_2 + m_2 b^2) \ddot{\theta} = -c \theta + m_2 b g$$

Equation 2: (two degrees-of-freedom)

$$\begin{bmatrix} I_2 + m_2 b^2 \left(1 - \frac{m_2}{m_1 + m_2}\right) & 0 \\ m_2 b & (m_1 + m_2) \end{bmatrix} \begin{pmatrix} \ddot{\theta} \\ \ddot{z} \end{pmatrix} = \begin{pmatrix} -c \theta \\ (m_1 + m_2) g \end{pmatrix}.$$

The impact of the woodpecker bill is elastic. The other impacts are nonelastic and lead to jumps in the angular velocity,  $\dot{\theta}$ . The impulse changes to the angular velocity are defined by

- impact at the top of the sleeve

$$\dot{\theta}^+ = (1 - d_3) \left( \dot{\theta}^- + \frac{m_2 b}{I_2 + m_2 b^2} \dot{z}^- \right),$$

impact at the bottom of the sleeve

$$\dot{\theta}^+ = (1 - d_1) \left( \dot{\theta}^- + \frac{m_2 b}{I_2 + m_2 b^2} \dot{z}^- \right),$$

impact of the woodpecker bill

$$\dot{\theta}^+ = -\dot{\theta}^-$$

where  $\dot{\theta}^-$ ,  $\dot{z}^-$  are the velocities before the impact and  $\dot{\theta}^+$ ,  $\dot{z}^+$  are the velocities after the impact.

The constants for the problem have been taken from Pfeiffer [6]:  $b=0.0015$  m,  $a=0.025$  m,  $m_1=0.003$  kg,  $m_2=0.0045$  kg,  $I_2=7 \times 10^{-7}$  kg-m<sup>2</sup>,  $c=0.0056$  Nm,  $d_3=0.04766$ ,  $d_1=0.18335$ ,  $g=9.81$  m/sec<sup>2</sup>,  $\theta_{k1}=10$  deg,  $\theta_{k2}=12$  deg.

**4.2.1 Distinct Motion Phases.** The motion is divided into five distinct phases. The angular motion of the woodpecker defines each phase. The motion phases are monitored during a simulation to determine when the number of degrees-of-freedom change and when impulse loads are applied to system. The five motion phases are described as follows:

**Phase 1**  $\theta > \theta_{k1}$ : The woodpecker swings to the right until it reaches its maximum amplitude and then swings back to the left. During this phase the translational degree-of-freedom is locked.

**Phase 2**  $-\theta_{k1} < \theta \leq \theta_{k1}$ : Phase 1 ends when  $\theta = \theta_{k1}$  and the translational degree-of-freedom is released so that the sleeve/woodpecker system can move downward. The system experiences an impulse when the translational motion starts.

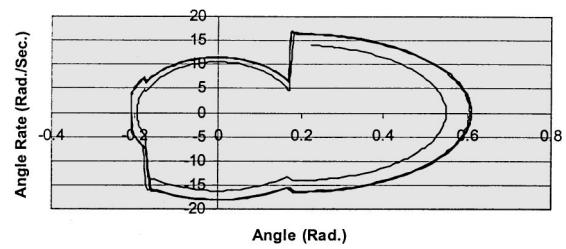
**Phase 3**  $-\theta_{k2} < \theta \leq -\theta_{k1}$ : Phase 2 ends when  $\theta = -\theta_{k1}$ , the rotational velocity impulsively changes, and the impact stops the sleeve motion.

**Phase 4**  $-\theta_{k2} \leq \theta < -\theta_{k1}$ : Phase 3 ends when  $\theta = -\theta_{k2}$ , the woodpecker touches the rod with its bill, and the rotational velocity impulsively reverses direction. The sleeve motion remains locked.

**Table 2** Number of nonlinear event as a function of the number of integration steps

Number of Integration Steps	Number of Events Detected
100	10
200	25
500	25
1000	30
2000	31
5000	31
10,000	31

**Angular Motion Phase Diagram**



**Fig. 6** Phase variable plot for woodpecker angle-angular rate plot

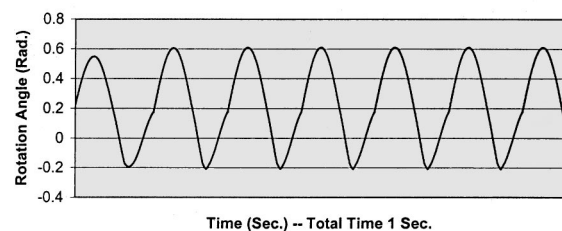
**Phase 5**  $-\theta_{k1} \leq \theta < \theta_{k1}$ : Phase 4 ends when  $\theta = -\theta_{k1}$ , the rotational velocity is impulsively changed, the sleeve motion is released and the sleeve/woodpecker again moves down the rod. This phase continues until  $\theta = \theta_{k1}$ , the rotational motion is impulsively changed, the sleeve motion stops, and Phase 1 begins again.

**4.2.2 Numerical Simulation Results for the Woodpecker Rotational/Translational Impact Problem.** A standard fourth-order Runge-Kutta integration algorithm is used with a fixed integration step size to generate the simulation results (Hildbrand [8], and Halin [5]). Several simulations have been run to identify the number of nonlinear events that occur during a one-second simulation. As shown in Table 2, 2000 integration time-steps are required before the number of nonlinear events becomes insensitive to the integration step-size.

Simulation results are presented in Figs. 6–12. A phase diagram of the angular woodpecker motion is presented in Fig. 6. The motion is seen to complete approximately one complete cycle before the motion history repeats in the phase space.

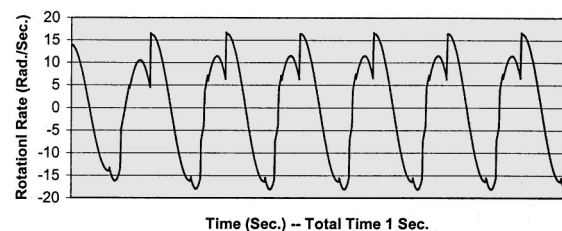
The angular time history motion is presented in Fig. 7. The motion is seen to be nearly periodic. The lower limit of the angular motion is governed by the Woodpecker's bill striking the rod and reversing the direction of the motion. The motion is basically smooth and continuous.

**Woodpecker Rotation History**



**Fig. 7** Woodpecker angle time history

**Woodpecker Rotational Rate History**



**Fig. 8** Woodpecker angular rate time history

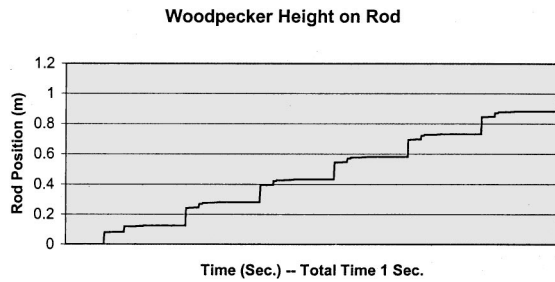


Fig. 9 Woodpecker height on rod

The angular rate time history is presented in Fig. 8. Discontinuous changes in the angular rates are present. The motion is not symmetric with respect to the time axis. These differences are attributed to the two impulsive impacts that take place when the angle is negative, versus the single impulsive impact that takes place when the angle is positive.

The woodpecker's height on the rod is presented in Fig. 9. The motion is seen to highly discontinuous. Multiple stops and starts occur during the motion.

The woodpecker's translational velocity relative to the rod is presented in Fig. 10. Very sharp changes in the translational velocity are observed.

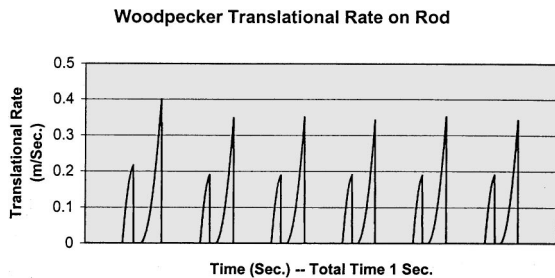


Fig. 10 Woodpecker translational velocity relative to the rod

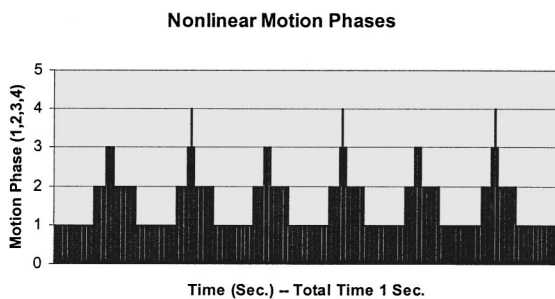


Fig. 11 Nonlinear phase motion time history

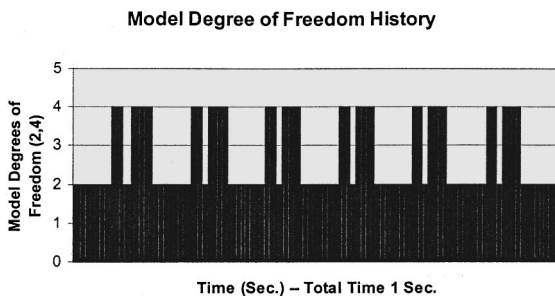


Fig. 12 Model degree-of-freedom time history

The coupled rotational-translational motion goes through four phases of motion that are defined by the angular motion and impulsive constraint inputs. There is a fifth motion phase, but it is solely defined by motion reversal when the woodpecker bill hits the rod. The motion phase history is presented in Fig. 11. The degree-of-freedom model change during the motion and Fig. 12 presents the changes in the number of model degrees-of-freedom.

A discontinuous event has typically been detected with  $|s_0| \sim 0.005$ . After the one-step integration algorithm of Eq. (10) has been applied the value for the constraint has been found to be  $|C| \sim 10^{-12}$ .

## 5 Intermittent Constraint Handling

The solution for Eq. (1) is complicated when the physical application must track several constraints simultaneously. The slack variable formulation only depends on identifying the next constraint to reach zero. As a result, if one or more constraints pass through zero during the sampling of Eq. (1), an estimate is made by extrapolating an estimate for  $t^*$  from all of the active constraints. Taylor expanding slack variable for each active constraint leads to

$$s_i(\bar{t}_i^*) \approx s_i(t_n) + \frac{ds_i}{dt}(\bar{t}_i^* - t_n) \quad (13)$$

where  $\bar{t}_i^*$  is the estimated time for  $i$ th constraint to reach zero. Equation (13) is solved by setting the left-hand side to zero, leading to the following Newton-like estimate for  $\bar{t}_i^*$ :

$$\bar{t}_i^* \approx t_n - s_i / ds_i / dt.$$

The smallest estimate for  $\bar{t}_i^*$  defines the constraint equation to be used for the extended state-space transformation of Eq. (10).

## 5 Conclusions

This paper has presented a slack variable formulation for handling inequality constraints. The new algorithm eliminates the need for iterative approaches for solving constraints for the exact time that a constraint is exactly satisfied. A slack variable that converts the inequality constraint into an equality constraint is introduced. The key step in the algorithm is that the slack variable replaces time as the independent variable. The transformed constraint equation is used to define a differential equation for time as a function of the slack variable. The original differential equations are transformed to become a function of the slack variable. An extended state-space is defined that permits the state to be noniteratively integrated to the constraint surface. The advantages of the new method are that (1) the integration algorithm never passes over a singular event (thereby, avoiding numerical instabilities), (2) no iteration is required, (3) conventional integration routines can be used, and (4) little computational overhead is required to implement the method. These algorithms support interdisciplinary problems in solid mechanics, structures, dynamics and control, where damping, friction, stiction, and variable topology problems arise. Two examples are presented that demonstrate the effectiveness of the slack variable formulation.

## References

- [1] Carver, M. B., 1978, "Efficient Integration Over Discontinuities in Ordinary Differential Equation Simulations," *Math. Comput. Simul.*, **20**, No. 3, pp. 190–196.
- [2] Ellison, D., 1981, "Efficient Automatic Integration Of Odes With Discontinuities," *Math. Comput. Simul.*, **23**, No. 1, pp. 12–20.

- [3] Enright, W. H., Jackson, K. R., Nørsett, S. P., and Thomson, P. G., 1986, "Effective Solution of Discontinuous Ivps Using a Runge-Kutta Formula Pair With Interpolants," *Appl. Math. Comput.*, **27**, pp. 313–335.
- [4] Filippov, A. F., 1964, "Differential Equations With Discontinuous Right Hand Side," *ASM Trans.*, **42**, pp. 199–231.
- [5] Halin, H. J., 1976, "Integration of Ordinary Differential Equations Containing Discontinuities," *Proceedings of the Summer Computer Simulation Conference* 976, La Jolla, SCI Press, La Jolla, CA, pp. 46–53.
- [6] Pfeiffer, F., 1984, "Mechanische Systeme mit unstetigen Übergängen," *Ingenieurarchiv*, **54**, pp. 232–240.
- [7] Eich-Soellner, Edda, and Fuhrer, Claus, 1998, *Numerical Methods in Multibody Dynamics*, B. G. Teubner, Stuttgart.
- [8] Hildbrand, F. B., 1962, *Advanced Calculus for Applications*, Prentice-Hall, Englewood Cliffs, NJ.

# Relationship Between the External Force and the Specific Work of Fracture $R$ for Steady-State Tearing and Peeling of Elastoplastic Materials

J. H. Liu<sup>1</sup>

A. G. Atkins

G. Jeronimidis

Department of Engineering,  
University of Reading,  
Whiteknights,  
P. O. Box 225,  
Reading RG6 2AY, UK

*A simple relationship is obtained between the external force  $F$  and the fracture toughness  $R$  for thin sheets in steady state elastoplastic combined tearing and peeling along self-similar paths. The relationship depends only on the material properties ( $E$ ,  $\sigma_y$ , and  $\alpha$  for an elastoplastic material with linear hardening) and strip cross section ( $B$  and  $H$ ). An earlier analysis (which incorporates transient tearing and peeling) requires lengthy computations over the whole length of the strip. The present analysis avoids that complication. Experiments in steady-state agree with the theory. [DOI: 10.1115/1.1387020]*

## 1 Introduction

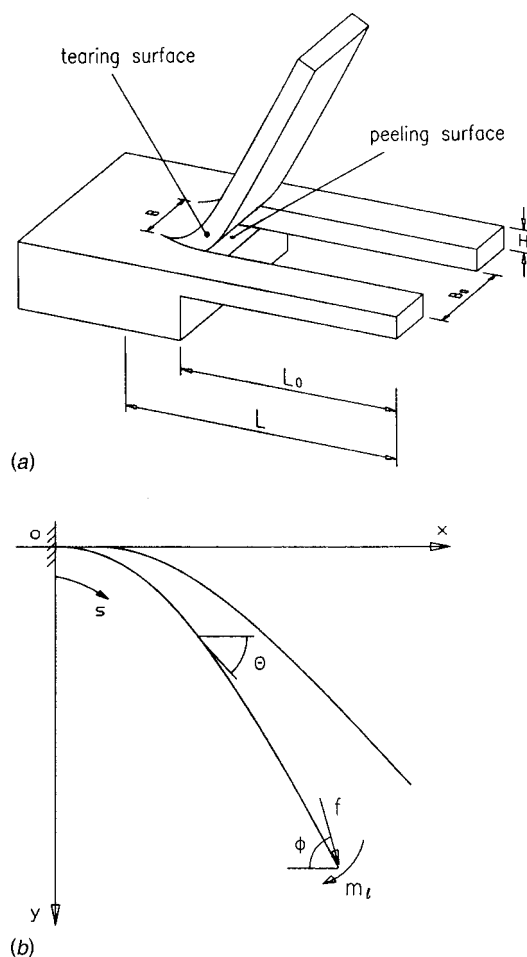
The relationship between the external loads and the specific work of fracture  $R$  (fracture toughness) for  $j$  strips tearing and peeling (e.g., as shown in Fig. 1(a)) can be obtained from ([1,2])

$$\sum_{i=1}^j F_i du_i + \sum_{i=1}^j (M_i)_i d\theta_i = \sum_{i=1}^j (d\Lambda_i + d\Gamma_i) + \sum_{i=1}^j R_i dA_i, \quad (1)$$

where  $F$  and  $M_i$  are the external force and moment, and  $du$  and  $d\theta$  are the incremental displacement and rotation angle at the loading point,  $d\Lambda$  and  $d\Gamma$  are the incremental elastic and plastic work in each strip,  $dA$  is the incremental area of fracture surface and  $RdA$  the incremental fracture work. The strip is commonly treated as a cantilever in the analysis for tearing and peeling, as shown in Fig. 1(b), with axial stretching being neglected ([2–6]).

For moment loading at the tip of the cantilever, the analysis of deformation is quite straightforward and the elastic and plastic work is easy to calculate, since the bending moment and curvature are the same at any point of the cantilever. For tearing and peeling along self-similar paths (without the change of strip cross section), the relationship between the moment and fracture toughness is given in Eq. (14a) by Liu et al. [2] for elastic deformation and, for strips in elastoplastic deformation, it can be obtained from Eqs. (10b) and (14b) in Liu et al. [2]. However, experimentally, it is very difficult to achieve a pure bending condition: A double roller experimental system was used by Yu et al. [6] and Muscat-Fenech and Atkins [5] to test the tearing fracture toughness of thin plate, but it was found that the tear legs disengage from the roller to form their own natural radii of curvature and the curvature is no longer constant near the tearing front. For tearing and peeling of thin plate with a concentrated force at the end of the strip, the analysis of strip deflection is very complicated since it is usually associated with large deflections ([2,7]). Even so, the full analysis

has been presented by Liu et al. [2] for the tearing and peeling of an end-loaded cantilever strip, to include both the transient conditions of increasing force during tearing from initial loading, up



**Fig. 1 (a) tearing and peeling of thin sheets, (b) deformation of strip during the tearing and peeling propagation**

<sup>1</sup>Current address: Advantica Technologies, Ltd., Ashby Road, Loughborough, Leicestershire LE11 3GA, UK.

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. 21, 1998; final revision, May 17, 2001. Associate Editor: J. W. Ju. Discussion on the paper should be addressed to the 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.



to steady-state tearing at constant force. The steady-state condition will be attained *providing* the “trouser leg” is long enough (i.e., a long enough starting length) or that the tear propagates to a sufficiently long length. Magnitudes of the length of tear and peel required to achieve a steady state are discussed later in this paper. The steady-state condition for tearing and peeling appears to start just after the rotation angle of the cantilever at the loading point at the end of the beam, reaches the force loading angle ( $\phi$  in Fig. 1(b)) whence it remains unchanged subsequently. This observation has important implications, since for an increment of tearing length  $dL$ , (i) the work due to the external force  $F$  is  $FdL$ , (ii) the incremental elastic and plastic work in the strip is equal to the work of the newly increased beam length  $dL$  bending with curvature from zero to its maximum value and returning to zero, and (iii) the incremental fracture work is equal to the fracture toughness times the increased fracture surface associated with  $dL$ . Therefore, instead of calculating the change of elastic and plastic work along the whole cantilever (as done by Liu et al. [2]), only the work in the newly created part of the strip  $dL$  during steady-state fracture propagation is needed in Eq. (1) to obtain the relationship between the external force and the fracture toughness.

The present paper discusses the relationship between the external force and fracture toughness in steady-state tearing and peeling thin sheets made of elastoplastic material with linear hardening shown in Fig. 2. It is assumed that all strips are identical and tearing and peeling occurs in a self-similar manner, i.e., there is no “tearing to a point” ([5]) so that the cross section of the strips remains the same throughout. Figure 3 shows the arrangement of (a) two strips tearing (“trouser tearing”) and (b) two strips peeling. The relationship between bending moment and curvature for an elastoplastic cantilever is given by Liu et al. [7]; the relationship between the fracture toughness and curvature in the fracture front is obtained by Liu et al. [2] and the magnitude of this curvature is maximum during fracture propagation.

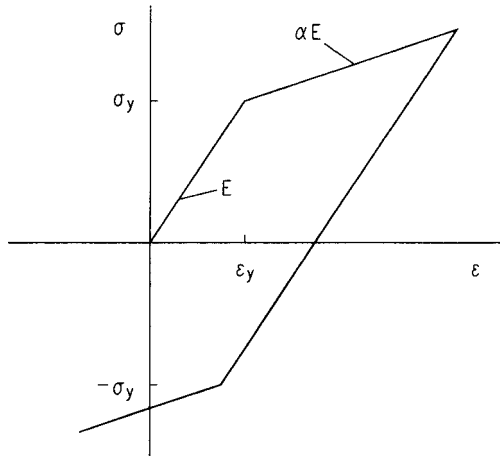


Fig. 2 Stress-strain relationship of strip material

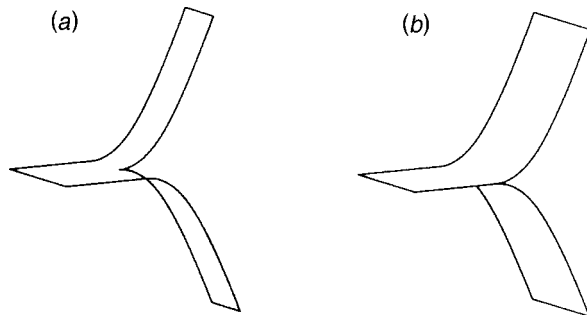


Fig. 3 The arrangement for two strips tearing and peeling

## 2 Elastic and Plastic Work

When a beam length  $dL$  is bent by a moment  $M$  with curvature changes from zero to  $d\theta/dS$ , the elastic work  $E_b$  is

$$E_b = \int_0^\theta M d\theta = \int_0^{d\theta/dS} M d\left(\frac{d\theta}{dS}\right) dL = \frac{M_e}{L_0} \int_0^\kappa m d\kappa dL, \quad (2a)$$

where

$$\theta = \frac{d\theta}{dS} dL, \quad (2b)$$

$$\kappa = \frac{d\theta}{ds} = L_0 \frac{d\theta}{dS}, \quad (2c)$$

$$s = \frac{S}{L_0}, \quad (2d)$$

$$m = \frac{M}{M_e}. \quad (2e)$$

$L_0$  is the original length of beam (the length before the fracture starts to propagate),  $S$  is the intrinsic coordinate of a section,  $M_e$  is the bending moment at the elastic limit and equals  $BH^2\sigma_y/6$  for rectangular cross section,  $B$  and  $H$  are the width and depth of beam, respectively, and  $\sigma_y$  is the yield stress.

For an elastoplastic material with linear strain hardening, the relationship between nondimensional bending moment  $m$  and curvature  $\kappa$  was obtained by Liu et al. [7], viz:

(i) elastic deformation,  $\kappa \leq \beta$ , or  $\alpha = 1$

$$m = \frac{\kappa}{\beta}, \quad (3a)$$

(ii) plastic loading,  $\kappa > \beta$

$$m = \frac{1}{2}(1-\alpha)(3-\beta^2\kappa^{-2}) + \frac{\alpha\kappa}{\beta}, \quad (3b)$$

(iii) elastoplastic deformation with partial linear unloading,  $\kappa \geq (1-\alpha)\kappa^* - (2-\alpha)\beta$

$$m = \frac{1}{2}(1-\alpha)\left(3-\beta^2\kappa^{*-2} - 2\frac{\kappa^*}{\beta}\right) + \frac{\kappa}{\beta}, \quad (3c)$$

(iv) elastoplastic deformation with partial reverse plastic loading,  $\kappa < (1-\alpha)\kappa^* - (2-\alpha)\beta$

$$m = \frac{1}{2}(1-\alpha)(2-\alpha)^3\beta^2A^{-2} - \frac{\alpha A}{\beta} - \frac{1}{2}(1-\alpha)\beta^2\kappa^{*-2} - \frac{3}{2}(1-\alpha)^2, \quad (3d)$$

in all of which four relations

$$\beta = \frac{M_e L_0}{EI}, \quad (3e)$$

$$A = (1-\alpha)\kappa^* - \kappa. \quad (3f)$$

$E$  is the Young's modulus,  $\alpha$  is the ratio of plastic modulus to elastic Young's modulus and  $\kappa^*$  is the maximum nondimensional curvature when the unloading starts.

Substituting Eqs. (3) into Eq. (2a), we obtain

(i) for  $0 \leq \kappa \leq \beta$ , or  $\alpha = 1$

$$E_b = \frac{M_e dL}{2L_0} \frac{\kappa^2}{\beta}, \quad (4a)$$

(ii) for  $\beta \leq \kappa \leq \kappa^*$

$$E_b = \frac{M_e dL}{2L_0} \left[ \alpha \frac{\kappa^2}{\beta} + 3(1-\alpha)\kappa + (1-\alpha)\beta^2\kappa^{-1} - 3(1-\alpha)\beta \right], \quad (4b)$$

(iii) for  $(1-\alpha)\kappa^* - (2-\alpha)\beta \leq \kappa \leq \kappa^*$

$$E_b = \frac{M_e dL}{2L_0} \left\{ \frac{\kappa^2}{\beta} + (1-\alpha) \left[ \left( 3 - \beta^2\kappa^{*-2} - 2\frac{\kappa^*}{\beta} \right) \kappa + \frac{\kappa^{*2}}{\beta} + 2\beta^2\kappa^{*-1} - 3\beta \right] \right\}, \quad (4c)$$

(iv) for  $\kappa < (1-\alpha)\kappa^* - (2-\alpha)\beta < \kappa^*$

$$E_b = \frac{M_e dL}{2L_0} \left\{ \frac{\alpha}{\beta} A^2 + (1-\alpha) \left[ (2-\alpha)^3 \beta^2 A^{-1} - \beta^2 \kappa^{*-2} \kappa - 3(1-\alpha)\kappa + \alpha \frac{\kappa^{*2}}{\beta} + 3(1-\alpha)(2-\alpha)\kappa^* + 2\beta^2 \kappa^{*-1} - 3(5-4\alpha+\alpha^2)\beta \right] \right\}. \quad (4d)$$

Therefore, when the part of beam of length  $dL$  is bent with curvature changing from zero to  $\kappa^*$  then returning to zero, we have

$$E_b = 0, \quad \text{for } \kappa^* \leq \beta, \text{ or } \alpha = 1 \quad (5a)$$

$$E_b = \frac{M_e dL}{2L_0} (1-\alpha) \left[ \frac{\kappa^{*2}}{\beta} + 2\frac{\beta^2}{\kappa^*} - 3\beta \right], \quad \text{for } \beta < \kappa^* \leq \frac{2-\alpha}{1-\alpha} \beta \quad (5b)$$

and

$$E_b = \frac{M_e dL}{2L_0} \left\{ (1-\alpha) \left[ \alpha(2-\alpha) \frac{\kappa^{*2}}{\beta} + 3(1-\alpha)(2-\alpha)\kappa^* - 3(5-4\alpha+\alpha^2)\beta \right] + [(2-\alpha)^3 + 2(1-\alpha)] \frac{\beta^2}{\kappa^*} \right\}, \quad (5c)$$

for  $\kappa^* > \frac{2-\alpha}{1-\alpha} \beta$ .

Equations (5) show that when the curvature of a part of beam  $dL$  changes from zero to a value  $\kappa^* \leq \beta$  then returns to zero, all the elastic energy in the beam is released with  $E_b = 0$ . However, after it enters plastic deformation with  $\kappa^* > \beta$ , a reverse bending moment is required to bend the beam back to zero curvature for which the associated work  $E_b > 0$ .

### 3 External Force $F$ and Fracture Toughness $R$

It has been shown ([2,4]) that during steady-state fracture the external force at the tip of a strip is a constant. It appears that the incremental elastic ( $d\Lambda$ ) and plastic ( $d\Gamma$ ) work in a strip is equal to that of the newly created strip length  $dL$  bending with curvature from zero to the maximum value  $\kappa^*$  and returning to zero, where  $d\Lambda + d\Gamma = E_b$  in Section 2 above, or  $d\Lambda + d\Gamma = E_b$  in Eq. (1). It is clear ([2]) that the curvature at the fracture front increases from zero to the value  $\kappa_0$  for fracture and then decreases during propagation, if the fracture path is self-similar and the fracture toughness is constant. Under these conditions, the maximum value of curvature  $\kappa^*$  is a constant and  $\kappa^* = \kappa_0$  along a newly created part of beam  $dL$  due to fracture. The relationship between the external force and the fracture toughness from Eq. (1) is

$$f = \frac{E_b L_0}{M_e dL} + \frac{1}{2} D \beta, \quad (6a)$$

for  $j$  strips,  $n_1$  tearing fracture surfaces and  $n - n_1$  peeling fracture surfaces, where

$$f = \frac{F L_0}{M_e}, \quad (6b)$$

$$D = \frac{6[n_1 r_t + (n - n_1) r_p]}{j}, \quad (6c)$$

$$r_t = \frac{E R_t}{B \sigma_y^2}, \quad (6d)$$

$$r_p = \frac{E R_p}{H \sigma_y^2}, \quad (6e)$$

$$E_b = d\Lambda + d\Gamma. \quad (6f)$$

$R_t$  and  $R_p$  are the tearing fracture and peeling fracture toughness, respectively.

The relationship between curvature and fracture toughness is given in Eqs. (10) of Liu et al. [2] with  $\kappa_0 = \kappa^*$  and

$$D = \frac{\kappa^{*2}}{\beta^2} \quad \text{for } \kappa^* \leq \beta, \text{ or } \alpha = 1 \quad (7a)$$

and

$$D = \alpha \frac{\kappa^{*2}}{\beta^2} - 2(1-\alpha) \frac{\beta}{\kappa^*} + 3(1-\alpha), \quad \text{for } \kappa^* > \beta. \quad (7b)$$

The relationship between the external force  $f$  and the fracture toughness  $D$  for steady-state fracture process can be obtained from Eqs. (5), (6), and (7) with  $E_b = d\Lambda + d\Gamma$ .

When  $\kappa^* > \beta$ , Eq. (7) gives

$$\frac{\kappa^*}{\beta} = \frac{2}{3-D} \quad \text{for } \alpha = 0 \quad (8a)$$

and

$$\frac{\kappa^*}{\beta} = \left[ \frac{1-\alpha}{\alpha} + \sqrt{\Delta} \right]^{1/3} + \left[ \frac{1-\alpha}{\alpha} - \sqrt{\Delta} \right]^{1/3} \quad \text{for } \alpha > 0 \text{ and } \Delta \geq 0 \quad (8b)$$

or

$$\frac{\kappa^*}{\beta} = 2 \sqrt{\frac{1}{3\alpha} [D - 3(1-\alpha)] \cos \omega}, \quad \text{for } \alpha > 0 \text{ and } \Delta \leq 0 \quad (8c)$$

where

$$\Delta = \left[ \frac{3(1-\alpha) - D}{3\alpha} \right]^3 + \frac{(1-\alpha)^2}{\alpha^2} \quad (8d)$$

and

$$\omega = \frac{1}{3} \cos^{-1} \left[ \frac{3(1-\alpha)}{D - 3(1-\alpha)} \sqrt{\frac{3\alpha}{D - 3(1-\alpha)}} \right]. \quad (8e)$$

(a) For elastic deformation  $\kappa^* \leq \beta$ ,  $f \leq \beta/2$ , or  $\alpha = 1$ .

When  $\kappa^* \leq \beta$ , or  $\alpha = 1$ ,  $d\Lambda + d\Gamma = E_b = 0$  and Eq. (6a) gives

$$f = \frac{1}{2} D \beta \quad (9a)$$

and

$$D = \frac{2f}{\beta}. \quad (9b)$$

In this case, Eq. (7a) with  $\kappa^* \leq \beta$  gives  $D \leq 1$ , therefore,  $f \leq \beta/2$ .

(b) For  $\beta < \kappa^* \leq (2-\alpha)\beta/(1-\alpha)$  or  $\beta/2 < f \leq \beta(2-\alpha)^2/2(1-\alpha)^2$ .

If  $\kappa^* > \beta$ , the strip enters elastoplastic deformation before unloading starts. Substituting Eqs. (5b) and (7b) into Eq. (6a) with  $E_b = d\Lambda + d\Gamma$ , we have

$$f = \frac{1}{2}(1-\alpha) \left[ \frac{\kappa^{*2}}{\beta} + 2 \frac{\beta^2}{\kappa^*} - 3\beta \right] + \frac{\beta}{2} \left[ \alpha \frac{\kappa^{*2}}{\beta^2} - 2(1-\alpha) \frac{\beta}{\kappa^*} + 3(1-\alpha) \right] = \frac{\kappa^{*2}}{2\beta} \quad (10a)$$

where  $\kappa^*/\beta$  is given by Eqs. (8), and

$$D = \alpha \frac{2f}{\beta} - 2(1-\alpha) \sqrt{\frac{\beta}{2f}} + 3(1-\alpha). \quad (10b)$$

Equation (9a) with  $\kappa^* \leq (2-\alpha)\beta/(1-\alpha)$  gives

$$f \leq \frac{(2-\alpha)^2 \beta}{2(1-\alpha)^2}. \quad (10c)$$

(c) For  $\kappa^* > (2-\alpha)\beta/(1-\alpha)$  or  $f > \beta(2-\alpha)^2/2(1-\alpha)^2$ .

Equations (5c), (6a), and (7b) give

$$f = \frac{\beta}{2} \left\{ e \frac{\kappa^{*2}}{\beta^2} + 3(1-\alpha)^2(2-\alpha) \frac{\kappa^*}{\beta} - 3(1-\alpha)(2-\alpha)^2 + (2-\alpha)^3 \frac{\beta}{\kappa^*} \right\} \quad (11a)$$

and

$$D = au^2 + 3(1-\alpha) - 2(1-\alpha)u^{-1}, \quad (11b)$$

where

$$u = \frac{1}{12} \left[ 12 + \frac{2f}{\beta} + \sqrt{\left( \frac{2f}{\beta} \right)^2 + 24 \frac{2f}{\beta} - 48} \right], \quad \text{for } \alpha = 0 \quad (11c)$$

$$u = \frac{1}{e} \left\{ 2 \sqrt{(1-\alpha)(2-\alpha)^2 + \frac{e}{3} \frac{2f}{\beta}} \cos \psi - (1-\alpha)^2(2-\alpha) \right\}, \quad \text{for } 0 < \alpha < 1 \quad (11d)$$

$$\psi = \frac{\pi}{3} - \frac{1}{3}$$

$$\times \arccos \left\{ \frac{(2-\alpha)^3 [1 + (1-\alpha)^3] + (1-\alpha)^2(2-\alpha) e \frac{2f}{\beta}}{2 \sqrt{\left[ (1-\alpha)(2-\alpha)^2 + \frac{e}{3} \frac{2f}{\beta} \right]^3}} \right\}, \quad (11e)$$

$$e = \alpha + \alpha(1-\alpha)(2-\alpha) = 1 - (1-\alpha)^3 \quad (11f)$$

and  $\kappa^*/\beta$  is given by Eq. (8).

#### 4 Calculation and Discussion

During the self-similar tearing and peeling of thin sheet, the rotation angle at the tip of strips,  $\theta_l$ , may reach and subsequently remain at the force loading angle  $\phi$ , shown in Fig. 1(b), if specimens are longer enough. In this case, tearing and peeling becomes a steady-state process with constant external force. Under these conditions when the beam increases from  $L$  to  $L + dL$  due to fracture propagation, the unloading work along the whole beam is equal to that of rebending the newly created length  $dL$  from  $\kappa = \kappa^*$  to  $\kappa = 0$ . Therefore, instead of calculating the deformation along the whole beam with a great deal of complexity, the incremental work in each strip can be obtained merely from the bending and rebending back of the newly created beam length  $dL$ .

The beam length corresponding to  $\theta_l = \phi$  will depend on the material properties, beam cross-section size, and the fracture

toughness. The stiffer the beam, the longer the length needed to achieve the steady-state process; and, the smaller the value of fracture toughness, the longer the length required to achieve a steady state. The relationship between the beam length for steady state and the fracture toughness is plotted in Fig. 4(a) for elastic materials ( $\alpha = 1$ ). Similar calculations were performed on beams made of bilinear elastic materials, i.e.,  $\alpha < 1$  (as for a plastic material, but where the unloading path reverses down the loading path towards the origin  $\sigma = 0$  at  $\varepsilon = 0$ ). It shows that the relationship is almost the same as that shown in Fig. 4, except for  $\alpha = 0$  with a sharp drop around  $D = 3$ .

Even with the external force increasing, *unloading* may occur in a cross section of beam at large deflections due to the shortening moment arm ([7]). For elastoplastic material, the required length for steady-state fracture is less since the curvature in the plastic unloading region is larger than that of bilinear elastic materials. Equation (3c) with  $\kappa^* = \text{constant}$  shows that the moment rate  $\partial m / \partial s$  is the same as that of elastic beams along the beam length created by fracture,  $0 \leq S \leq L - L_0$ . When the steady-state fracture, which originates with  $\theta = \theta_l = \phi$  at the beam tip, propagates to  $\theta = \phi$  at  $S = L - L_0$ , a beam length relationship between elastic beams and elastoplastic beams is obtained. The derivation is given in the Appendix and plotted in Fig. 4(b). It shows that the required beam length decreases with decrease of  $\alpha$ . For an estimated (or known) fracture value  $D$  with load force angle  $\phi = \pi/2$ , we can obtain the required elastic beam length for steady-state fracture from Fig. 4(a) and the relationship between elastic and elastoplastic beam length,  $(L_p - L_0)/(L_e - L_0)$  from Fig. 4(b) with a given value of  $\alpha$ . Therefore, the required length of a test specimen made of elastoplastic material can be estimated from the combination of Fig. 4(a) and Fig. 4(b), especially when the original length of beam (the length before the fracture starts to propagate),  $L_0$ , is relatively small.

The relationship between the fracture toughness and the external force is given in equations (9b), (10b), and (11b), or

$$D = \frac{2f}{\beta}, \quad \text{for } \frac{2f}{\beta} \leq 1, \quad \text{or } \alpha = 1 \quad (12a)$$

$$D = \alpha \frac{2f}{\beta} - 2(1-\alpha) \sqrt{\frac{\beta}{2f}} + 3(1-\alpha) \quad \text{for } 1 < \frac{2f}{\beta} \leq \frac{(2-\alpha)^2}{(1-\alpha)^2} \quad (12b)$$

and

$$D = au^2 + 3(1-\alpha) - 2(1-\alpha)u^{-1}, \quad \text{for } \frac{2f}{\beta} > \frac{(2-\alpha)^2}{(1-\alpha)^2} \quad (12c)$$

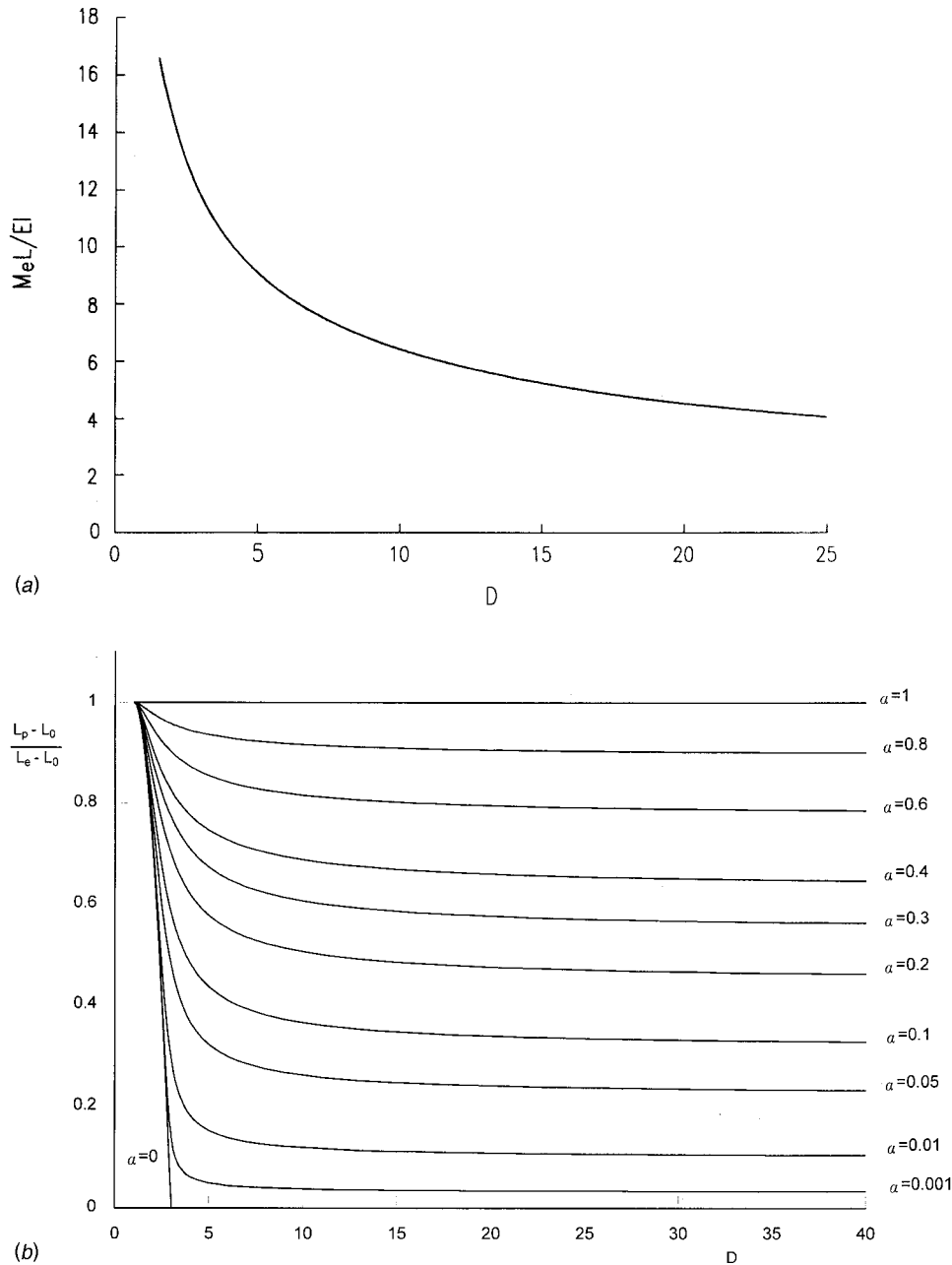
where

$$u = \frac{1}{12} \left[ 12 + \frac{2f}{\beta} + \sqrt{\left( \frac{2f}{\beta} \right)^2 + 24 \frac{2f}{\beta} - 48} \right], \quad \text{for } \alpha = 0 \quad (12d)$$

$$u = \frac{1}{e} \left\{ 2 \sqrt{(1-\alpha)(2-\alpha)^2 + \frac{e}{3} \frac{2f}{\beta}} \cos \psi - (1-\alpha)^2(2-\alpha) \right\}, \quad \text{for } 0 < \alpha < 1 \quad (12e)$$

$$\psi = \frac{\pi}{3} - \frac{1}{3}$$

$$\times \arccos \left\{ \frac{(2-\alpha)^3 [1 + (1-\alpha)^3] + (1-\alpha)^2(2-\alpha) e \frac{2f}{\beta}}{2 \sqrt{\left[ (1-\alpha)(2-\alpha)^2 + \frac{e}{3} \frac{2f}{\beta} \right]^3}} \right\} \quad (12f)$$



**Fig. 4 (a) The required length for steady-state tearing and/or peeling specimens made of elastic materials (for loading angle  $\phi=\pi/2$ ), (b) relationship between elastic and elastoplastic beam length required for steady state tearing and/or peeling**

and

$$e = 1 - (1 - \alpha)^3. \quad (12g)$$

It is clear that the relationship is only related to the material properties ( $E, \alpha, \sigma_y$ ) and dimensions of strips (width  $B$  and depth  $H$ ). During experimental tests, these parameters are usually known. Therefore, the fracture toughness follows directly from experimental values of external force and can be easily obtained from Eqs. (12). The variation of the nondimensional fracture toughness,  $D$  (given by Eq. 6(c)), with the nondimensional external force,  $2f/\beta = 2FEI/M_e^2$ , is plotted in Fig. 5 for different values of  $\alpha$ . It shows that greater fracture toughness,  $D$ , requires increased external force. Figure 5 also shows that with lowered ratio  $\alpha$  of plastic modulus to elastic Young's modulus in the trouser materials, the load increases for the same toughness. For  $\alpha=0$ ,

$D \rightarrow 3$  when the curvature tends to infinity, or when the bending moment tends to the value for plastic collapse, i.e.,  $M = BH^2\sigma_y/4$ . Therefore, when  $D$  is about 3 in magnitude, a small increase in fracture toughness produces a big change of curvature and the external force increases sharply ( $D$  cannot be larger than 3 for a perfectly plastic material). We employ Fig. 5 to find the value of fracture toughness, from the known external force in experiments on known materials of given strip dimensions.

Trouser tear tests shown in Fig. 3(a) are often employed to obtain the tearing fracture toughness  $R_t$ . If the strip deformation is in the elastic range ( $\alpha=1$ ), Eq. (12a) with  $j=2$ ,  $n=n_1=1$  and  $D=3ER_t/B\sigma_y^2$  gives the well-known relation

$$R_t = \frac{2F}{H}, \quad (13a)$$



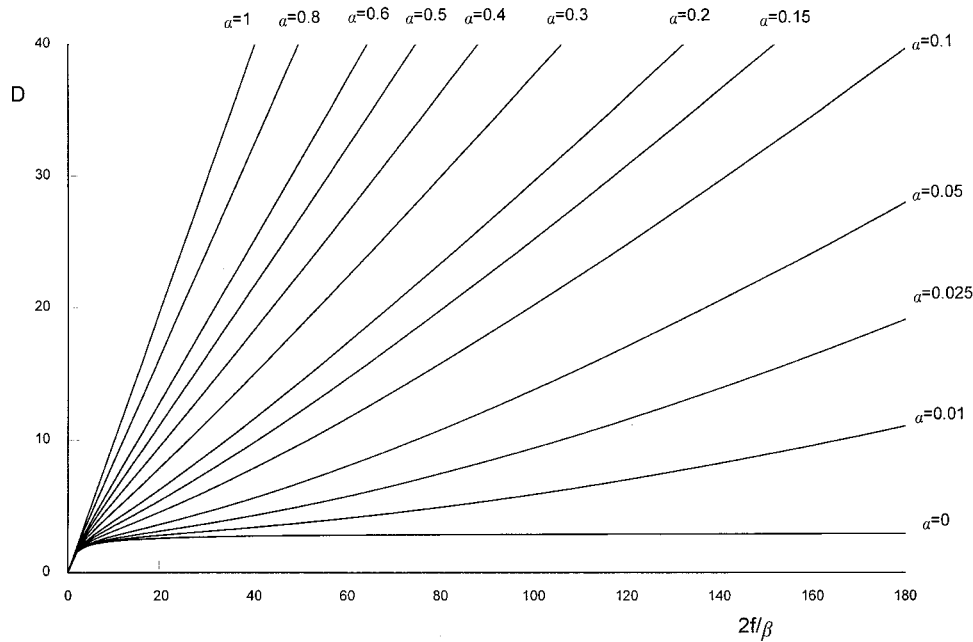


Fig. 5 Variation of the mixed nondimensional fracture toughness with the nondimensional external force,  $2f\beta = 2FE/HM_e^2$

since

$$\frac{2f}{\beta} = \frac{6FE}{BH\sigma_y^2} = D = \frac{3ER_t}{B\sigma_y^2}. \quad (13b)$$

The above equations are only valid when  $B \geq 3ER_t/B\sigma_y^2$ . Unfortunately, when  $B$  is large, twist of the legs may affect the test results.

For elastoplastic strip deformation at smaller values of  $B$ , the  $R_t$ - $F$  relationship can be obtained from Eqs. (12b)-(12f) with  $D = 3ER_t/B\sigma_y^2$  and  $2f/\beta = 6FE/BH\sigma_y^2$ . Figure 6 shows the comparison between the theoretical results obtained from Eqs. (12) (solid lines) and experimental results ([3,4]). These tearing tests were steady state. The theoretical results from Eqs. (12) fit well with experimental results when the fracture toughness was chosen as  $R_t = 44 \text{ kJ/m}^2$  for steel ( $\alpha = 0.00234$  and  $\beta = 0.45$ );  $R_t = 52 \text{ kJ/m}^2$  for aluminum ( $\alpha = 0.00551$  and  $\beta = 0.37714$ ); and  $R_t = 120 \text{ kJ/m}^2$  for brass ( $\alpha = 0.12025$  and  $\beta = 0.31634$ ). Instead of very complicated and lengthy computing calculations for the data of Fig. 9 in Liu et al. [2] only a few minutes of computing time were used to obtain the data in Fig. 6.

The method of plotting force/thickness versus width of strip employ in Fig. 6 had been used by, for example, by Mai and Cotterell [3] and Atkins and Mai [1] to obtain the fracture toughness using a simple analysis for tearing which predicted that  $F/H$  increases linearly with the increase of strip width  $B$ . However, as shown in Fig. 6, experimental data over much wider ranges of  $B$  than used by these authors show that  $F/H$  is not a linear function of  $B$ , nor does it always increase with increase of  $B$ . Furthermore, even when the nonlinear shape of  $F/H$  versus  $B$  plots shown in Fig. 6 had been correctly treated by Liu et al. [2] the values of fracture toughness had to be obtained by fitting the experimental data to the algebra by a trial and error method to find the toughness which gave a best fit. It is clear from the present analysis that the fracture toughness can be found directly from Eqs. (12), or from employing some master graphs similar to Fig. 5, using only the force obtained from experimental results.

It should be emphasized that  $F$  here is the *steady-state* force for tearing and/or peeling. Using the external force at the *initiation* of fracture in Eqs. (12a) and (13a) will *overestimate* the tearing frac-

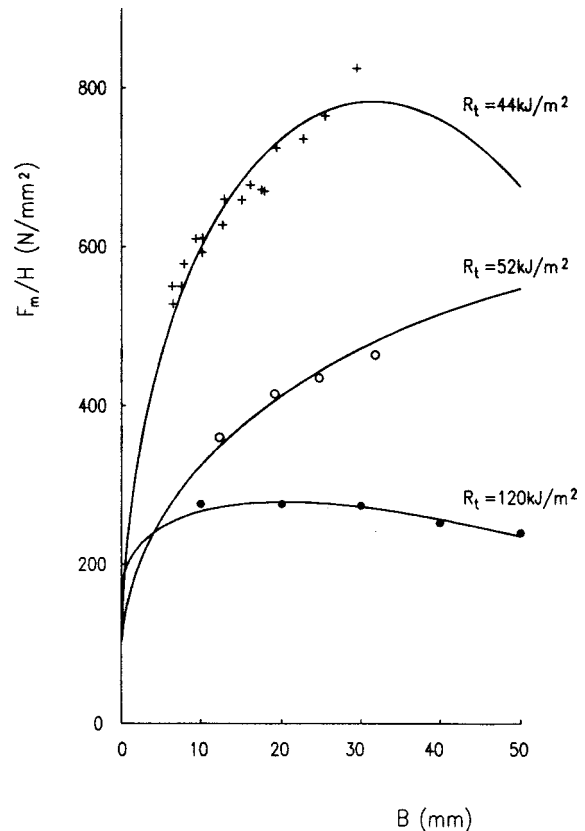
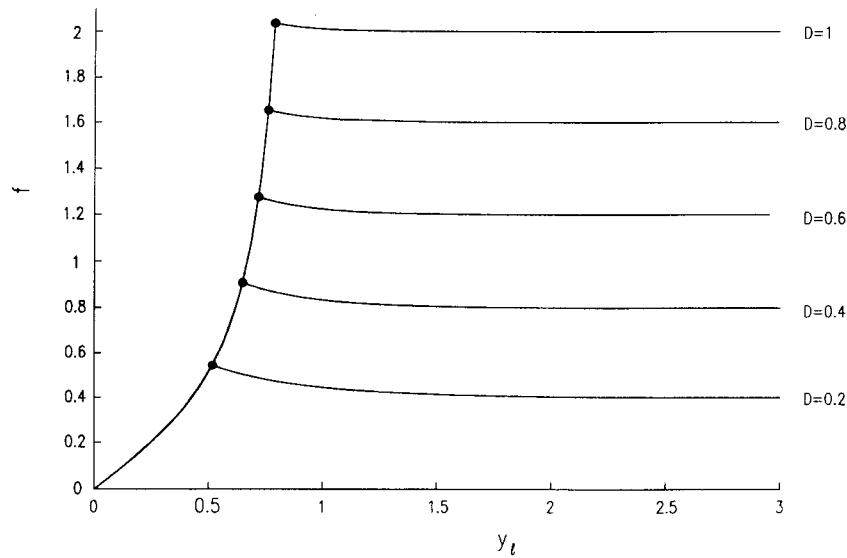


Fig. 6 Variation of the external force  $F_m/H$  with the beam width  $B$ . —: Eqs. (9), (10), and (11); +: experimental result of steel from [2]; O: 5251 aluminum alloy ([2]) and ●: brass ([3]).



**Fig. 7 Variation of nondimensional external force  $f$  with nondimensional displacement at the tip of beam  $y_t$  for elastic tearing and/or peeling.  $\phi = \pi/2$ ; ●: initiation point of fracture.**

ture toughness in elastic fracture, if the rotation angle at the tip of the leg,  $\theta_t$ , is less than the force loading angle,  $\phi$ , (i.e., the original length of the strip,  $L_0$ , is not long enough or the fracture toughness is too small) at the beginning of fracture propagation. Figure 7 shows that after initiation in elastic fracture, the force will decrease to the constant (steady-state) value during propagation. For elastoplastic materials it is not possible to say whether the use of the initiation force will overestimate or underestimate the true fracture toughness: as shown in Figs. 6 of Liu et al. [2] relating to elastoplastic propagation, before the external force reaches the steady-state value, it may continuously decrease for small values of  $D$ ; decrease then increase for intermediate values of  $D$ ; or continuously increase for big  $D$ .

## 5 Conclusions

Self-similar tearing and/or peeling of thin sheets (without change of strip cross section) occurs in steady state with constant external force when the "legs" of the testpieces are long enough. In this case, the incremental elastic and plastic energy in the deformed strip is equal to that of the newly created length  $dL$  with curvature change from zero to a maximum value  $\kappa^*$  and returning to zero. The relationship between the external force and fracture toughness depends only on the material properties ( $\alpha$ ,  $E$ , and  $\sigma_y$ ) and cross section ( $B$  and  $H$ ). For a given value of the external force, the fracture toughness can be obtained from Eqs. (12), or from a master graph like Fig. 5 in this paper. Since, in elastic tearing and peeling, the initiation load is greater than the steady-state load, true toughness is overestimated if the load at first cracking is used. In elastoplastic tearing and/or peeling, the steady-state load may be smaller or greater than the initiation load, depending on the value of the nondimensional fracture toughness  $D$ . In consequence, it is not possible to say in this case whether the toughness will be less or greater than that obtained using the initiation load in place of the steady-state load.

## Appendix

For elastoplastic materials, Eq. (3c) with  $\kappa^* = \text{constant}$  gives

$$\frac{\partial m}{\partial s} = \frac{1}{\beta} \frac{\partial \kappa}{\partial s} = \frac{1}{\beta} \frac{d^2 \theta}{ds^2} \quad (A1)$$

in the newly created beam length due to fracture,  $0 \leq s \leq L - L_0$  shown in Fig. 1(a), when  $\kappa \geq (1 - \alpha)\kappa^* - (2 - \alpha)\beta$ . Equation (12a)

shows that the moment rate along the beam is the same as that in elastic beams. The moment given from the equilibrium condition is ([7])

$$m = f \cos \phi (x_t - x) - f \sin \phi (y_t - y), \quad (A2a)$$

or

$$\frac{\partial m}{\partial s} = -f \sin(\phi - \theta). \quad (A2b)$$

The curvature,  $\kappa$ , can be obtained from Eqs. (A1) and (A2b), which is

$$\kappa = \frac{d\theta}{ds} = \sqrt{\kappa^{*2} + 2f\beta[\cos \phi - \cos(\phi - \theta)]}. \quad (A3a)$$

Therefore, the length from the fracture front  $s$  is

$$s = \int_0^\theta \frac{d\theta}{\sqrt{\kappa^{*2} + 2f\beta[\cos \phi - \cos(\phi - \theta)]}}, \quad (A3b)$$

where  $\phi$  is the loading angle of the external force shown in Fig. 1(a),  $\kappa^*$  is the curvature at the fracture front and  $\kappa^* = \kappa_e^*$  is given by Eq. (7a) for elastic material, or  $\kappa^* = \kappa_p^*$  is given by Eq. (7b) for elastoplastic material.

When  $\theta = \phi$  and  $\kappa = 0$  at  $S = L - L_0$ , Eq. (A3a) and (A3b) give

$$2f\beta = \frac{\kappa^{*2}}{(1 - \cos \phi)}$$

and

$$\kappa^* \frac{L - L_0}{L_0} = \int_0^\phi \frac{d\theta}{\sqrt{1 + [\cos \phi - \cos(\phi - \theta)]/(1 - \cos \phi)}}. \quad (A4)$$

Equation (A4) is valid for both elastic beams and elastoplastic beams with  $\kappa \geq (1 - \alpha)\kappa^* - (2 - \alpha)\beta$  in the newly created beam length due to fracture  $0 \leq s \leq L - L_0$ . For the same force loading angle  $\phi$ , the right-hand side of Eq. (A4) is constant. Therefore, we have

$$\frac{L_p - L_0}{L_e - L_0} = \frac{\kappa_e^*}{\kappa_p^*}, \quad (A5)$$

where  $L_e$  and  $L_p$  are total elastic and elastoplastic beam lengths,  $\kappa_e^*$  and  $\kappa_p^*$  are the curvatures for elastic and elastoplastic materials and can be obtained from Eqs. (7a) and (7b), respectively, with a given value of fracture toughness.

## References

- [1] Atkins, A. G., and Mai, Y. W., 1988, *Elastic and Plastic Fracture—Metals, Polymers, Ceramics, Composites, Biological Materials*, Ellis Horwood, Chichester, UK.
- [2] Liu, J. H., Atkins, A. G., and Jeronimidis, G., 1996, "Theoretical Analysis of Large Deformation Simultaneous Tearing and Peeling of Elastoplastic Materials," *Proc. Inst. Mech. Eng.*, **210**, pp. 433–444.
- [3] Mai, Y. W., and Cotterell, B., 1984, "The Essential Work of Fracture for Tearing of Ductile Metals," *Int. J. Fract.*, **24**, pp. 229–236.
- [4] Muscat-Fenech, C. M., and Atkins, A. G., 1994, "Elastoplastic Trouser Tear Testing of Sheet Materials," *Int. J. Fract.*, **67**, pp. 69–80.
- [5] Muscat-Fenech, C. M., and Atkins, A. G., 1994, "Elastoplastic Convergent and Divergent Crack Paths in Tear Testing of Sheet Materials," *Fatigue Fract. Eng. Mater. Struct.*, **17**, pp. 133–143.
- [6] Yu, T. X., Zhang, D. J., Zhang, Y., and Zhou, Q., 1988, "A Study of the Quasi-Static Tearing of Thin Metal Sheets," *Int. J. Mech. Sci.*, **30**, pp. 193–202.
- [7] Liu, J. H., Atkins, A. G., and Pretlove, A. J., 1995, "The Effect of Inclined Loads on the Large Deflection Behavior of Elastoplastic Workhardening Straight and Pre-Bent Cantilevers," *Proc. Inst. Mech. Eng.*, **209**, pp. 87–96.

N. M. Kinkaid

O. M. O'Reilly<sup>1</sup>

Assoc. Mem. ASME  
e-mail: oreilly@me.berkeley.edu

Department of Mechanical Engineering,  
University of Berkeley,  
Berkeley, CA 94720-1740

J. S. Turcotte

AFRL/VAS,  
Building 45,  
2130 Eighth Street, Suite 1,  
Wright-Patterson AFB, OH 45433-7765  
Assoc. Mem. ASME

# On the Steady Motions of a Rotating Elastic Rod

*In this paper, a model for the deformation of a rotating prismatic rod-like body is developed and analyzed. The novel feature of the model is its incorporation of the Poisson effect. As a result, it provides realistic solutions for the deformed states of steadily rotating rods. The model presented in this paper is also simplified to a nonlinear version of a classical model for rotating rods. Numerical continuation is used to solve the boundary value problems associated with these models. [DOI: 10.1115/1.1381003]*

## 1 Introduction

Motivated in part by applications in turbomachinery and helicopter rotors, there have been numerous studies of the dynamics of rotating rods (cf. Antman [1], Leissa [2], Rao [3], and references therein). A major portion of this work is concerned with the vibrational response of an elastic rod which is being whirled about a fixed axis (cf. Fig. 1). The vibrations of interest are usually considered to be small amplitude disturbances which are superimposed on the steady deformation of the rod induced by the rotation. In general, the partial differential equations governing the vibrations are nonhomogeneous and depend on the steady motion. These equations have been used to show the variation of the flexural frequencies as the rotational speed increases (cf. Wright et al. [4] and references therein). Similar models have also been proposed to analyze this issue in ever broadening types of rods and boundary conditions.

If one considers the system shown in Fig. 1 and supposes that the rod is composed of an isotropic elastic material, then it is easy to visualize that as the rod rotates, its length will increase. Furthermore, this increase in length will, because of the Poisson effect, be accompanied by a contraction of the cross section. However, the model most frequently used in the literature to predict the steady motion only considers the longitudinal displacement. We shall refer to this model as the uniaxial model. If a linear uniaxial stress-strain constitutive relation is used to establish this model, then Bhuta and Jones [5] and Brunelle [6] observed that above a certain critical speed of rotation, solutions to this model ceased to exist. Hodges and Bless [7] later showed how this lack of existence could be eliminated using nonlinear constitutive relations.

In contrast to the vast majority of existing works, in the present paper we determine the steady motions of the rod by using a model which accommodates the Poisson effect. The resulting model is based on a geometrically exact elastic rod. The results of our work provide a more realistic solution to this problem. For sufficiently large angular speeds, we also find physically unrealistic solutions; however, their nature is not the same as those found in the uniaxial model mentioned earlier. We also point out how certain modifications to the constitutive equations can be

used to eliminate the existence of these solutions. By following the earlier work of O'Reilly and Turcotte [8], the steady motions we calculate can be used to determine the equations governing the small-amplitude vibrations of the rotating rod.

The rod theory on which we base our model was developed in a series of works by A. E. Green and P. M. Naghdi and several of their co-workers dating to 1966 (cf. [9–11] and references therein). It is a Cosserat rod theory which includes, as particular constrained cases, most existing rod theories. Green and Naghdi's rod theory has been used to examine various problems where the Poisson effect is important. These include contact problems ([12,13]) and wave propagation and vibration ([14,15]).

An outline of this paper is as follows. First, a brief discussion of the rod theory is contained in Section 2. Then, in Section 3, the model for the whirling rod is presented. Most of the developments in Section 3 follow [8]. In Section 4, the boundary value problem governing the steady motions is discussed. Next, in Section 5, the steady motions predicted by the model are presented. Because the boundary value problem involves a stiff set of ordinary differential equations, finding solutions numerically proved to be a delicate matter. The numerical continuation algorithm for boundary value problems used by AUTO97 [16] was employed. As discussed by Doedel [17], a numerical continuation method is particularly suited to our needs.<sup>2</sup> We discuss the uniaxial model in Section 6. There, we show how it can be obtained from the model discussed in this paper, and how it relates to other models which have appeared in the literature. The closing section of this paper discusses future work.

## 2 Preliminaries on the Model Development

The whirling rod-like body is modeled in this paper as a material curve and two directors  $\mathbf{d}_\alpha$ .<sup>3</sup> That is, we are using a Cosserat (or directed) rod theory to model the body. As shown in Fig. 2, the reference state of the material curve is chosen to coincide with the line connecting the centers of area of each cross section of the undeformed rod. The points on this line are identified by the arc-length coordinate  $\xi$ . In addition, the reference values of the directors,  $\mathbf{D}_1(\xi)$  and  $\mathbf{D}_2(\xi)$ , are chosen to be unit orthonormal vectors which span the cross section. In the deformed state of the rod at time  $t$ , the position vector of a point of the material curve is defined by the position vector  $\mathbf{r}=\mathbf{r}(\xi,t)$  and the directors are vector-valued functions:  $\mathbf{d}_\alpha=\mathbf{d}_\alpha(\xi,t)$ . Here, the coordinate  $\xi$  is

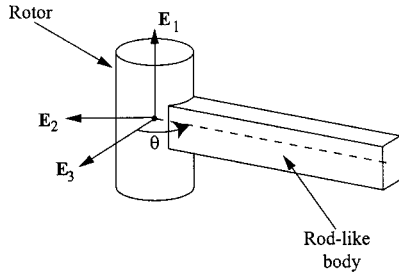
<sup>1</sup>To whom correspondence should be addressed.

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, August 16, 2000; final revision, January 2, 2001. Associate Editor: N. C. Perkins. Discussion on the paper should be addressed to the 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.

<sup>2</sup>An introduction to, and an overview of, continuation methods can be found in Seydel [18].

<sup>3</sup>In this paper, lowercase Latin indices range in value from 1 to 3 while lowercase Greek indices range from 1 to 2. These indices are summed when repeated. The set of vectors  $\{\mathbf{E}_i\}$  denotes a fixed right-handed Cartesian basis.





**Fig. 1 A rod-like body rotating with an angular speed  $\Omega$  about the  $E_1$ -axis. The angle  $\theta$  of rotation is such that  $\dot{\theta} = \Omega$ .**

used as a convected coordinate for the present configuration. In contrast to other rod theories, both the magnitude and direction of the directors are free to change.

To determine the equations governing  $\mathbf{r}$  and  $\mathbf{d}_\alpha$ , it is necessary to postulate balance laws and constitutive relations. These can be found in Naghdi [10]. We recall, from his review article, that the nontrivial balance laws for an elastic rod are the balance of linear momentum and two balances of director momentum:

$$\begin{aligned} \mathbf{n}' + \lambda \mathbf{f} &= \lambda \ddot{\mathbf{r}} + \lambda y^\beta \ddot{\mathbf{d}}_\beta, \\ \mathbf{m}^{\alpha'} + \lambda \mathbf{l}^\alpha - \mathbf{k}^\alpha &= \lambda y^\alpha \ddot{\mathbf{r}} + \lambda y^{\alpha\beta} \ddot{\mathbf{d}}_\beta. \end{aligned} \quad (1)$$

In these equations, the superposed dot is used to denote the partial derivative with respect to  $t$ , while the prime denotes the partial derivative with respect to  $\xi$ . Furthermore,  $\mathbf{n}$  is the contact force,  $\mathbf{m}^\alpha$  are the contact director forces,  $\mathbf{k}^\alpha$  are the intrinsic director forces,  $\mathbf{f}$  is the assigned force per unit mass, and  $\mathbf{l}^\alpha$  are the assigned director forces per unit mass. The constant inertia parameters are the mass per unit undeformed length,  $\lambda$ , and the inertias,  $y^\alpha$  and  $y^{\alpha\beta} = y^{\beta\alpha}$ . All of the aforementioned fields can be placed in approximate correspondence with the fields of three-dimensional continuum mechanics. We shall shortly use these correspondences to prescribe the inertia parameters.

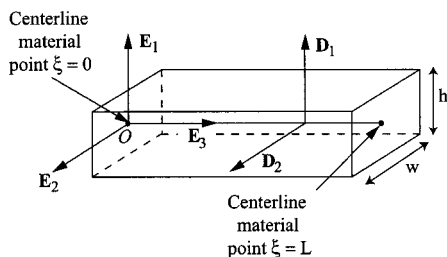
To close Eq. (1), it is necessary to specify constitutive relations for the force fields. Here, we assume that the rod is elastic with a strain-energy per unit length  $\lambda\psi$ . This energy is assumed to be a function of the following 12 independent strains:

$$\gamma_{ij} = \mathbf{d}_i \cdot \mathbf{d}_j - \mathbf{D}_i \cdot \mathbf{D}_j, \quad \kappa_{\alpha i} = \mathbf{d}'_\alpha \cdot \mathbf{d}_i - \mathbf{D}'_\alpha \cdot \mathbf{D}_i, \quad (2)$$

where  $\mathbf{d}_3 = \mathbf{r}'$ . That is,  $\lambda\psi = \lambda\hat{\psi}(\gamma_{ij}, \kappa_{\alpha k})$ . The resulting constitutive relations can be written in the form

$$\mathbf{n} = \lambda \frac{\partial \psi}{\partial \mathbf{r}'}, \quad \mathbf{k}^\alpha = \lambda \frac{\partial \psi}{\partial \mathbf{d}'_\alpha}, \quad \mathbf{m}^\alpha = \lambda \frac{\partial \psi}{\partial \mathbf{d}'_\alpha}. \quad (3)$$

Specifying  $\lambda$ ,  $y^\alpha$ ,  $y^{\alpha\beta}$ ,  $\psi$ ,  $\mathbf{f}$ , and  $\mathbf{l}^\alpha$ , and substituting (3) into (1), a set of nine scalar partial differential equations for the nine scalar fields  $\mathbf{r}$  and  $\mathbf{d}_\alpha$  are obtained.



**Fig. 2 The reference configuration of the rod-like body**

### 3 The Rod and Its Motions

In this paper, the reference state of the rod-like body is a prismatic, homogeneous parallelepiped whose cross sections are of height  $h$  and width  $w$ , and whose length is  $L$  (cf. Fig. 2). The steady motions of interest here are such that the deformed state of this body, when viewed by an observer corotating with an angular velocity  $\Omega \mathbf{E}_1$ , appears to be in a state of equilibrium. The cross sections and centerline of the steadily rotating body will be deformed. Specifically, the motion of the rod modeling the body is assumed to be of the form

$$\mathbf{r} = x \mathbf{P} \mathbf{D}_3, \quad \mathbf{d}_1 = d_1 \mathbf{P} \mathbf{D}_1, \quad \mathbf{d}_2 = d_2 \mathbf{P} \mathbf{D}_2, \quad (4)$$

where  $x = x(\xi)$ ,  $d_\alpha = d_\alpha(\xi)$ ,  $\mathbf{D}_i = \mathbf{E}_i$ , and the rotation tensor  $\mathbf{P}$  corresponds to a rotation about the  $\mathbf{E}_1$ -axis (cf. Fig. 1) with a constant (counterclockwise) speed  $\Omega$ .<sup>4</sup> It is straightforward to show that, for the motion (4),

$$\ddot{\mathbf{r}} = -\Omega^2 x \mathbf{P} \mathbf{E}_3, \quad \ddot{\mathbf{d}}_1 = \mathbf{0}, \quad \ddot{\mathbf{d}}_2 = -\Omega^2 d_2 \mathbf{P} \mathbf{E}_2. \quad (5)$$

Physically, the motion (4) is such that material planes which are normal to the centerline in the reference configuration of the body, remain normal in the present configuration. However, the width and height of these sections will have changed to  $wd_2$  and  $hd_1$ , respectively. To examine whether the motion (4) can be sustained in a particular rotating rod depends intimately on the reference geometry and material of the rod, and the applied forces acting on the rod. We now turn to specifying these quantities.

For the body of interest, we find, using the prescriptions discussed in Green and Naghdi [11], that

$$\lambda = \rho_0 h w, \quad \lambda y^{11} = \frac{\rho_0 h^3 w}{12}, \quad \lambda y^{22} = \frac{\rho_0 w^3 h}{12}, \quad (6)$$

while  $y^1$ ,  $y^2$ , and  $y^{12} = y^{21}$  are zero. Here, the constant mass density per unit volume of the rod-like body is  $\rho_0$ . While the rod is whirling, we assume that external body forces (such as gravity) and tractions on its lateral surface are absent. Consequently,  $\mathbf{f} = \mathbf{l}^1 = \mathbf{l}^2 = \mathbf{0}$ .

Pertaining to the material of the rod, we shall assume that the rod-like body is composed of a homogeneous isotropic elastic material. Following Green, Laws, and Naghdi [19], the strain-energy function of the rod is assumed to be a quadratic function of the strains:  $\lambda\psi = \lambda\psi_1$ . The function  $\psi_1$  is given by

$$\begin{aligned} 2\lambda\psi_1 &= k_1 \gamma_{11}^2 + k_2 \gamma_{22}^2 + k_3 \gamma_{33}^2 + k_7 \gamma_{11} \gamma_{22} + k_8 \gamma_{11} \gamma_{33} + k_9 \gamma_{22} \gamma_{33} \\ &+ \frac{k_4}{4} (\gamma_{12} + \gamma_{21})^2 + k_5 \gamma_{23}^2 + k_6 \gamma_{13}^2 + k_{10} \kappa_{11}^2 + k_{11} \kappa_{22}^2 \\ &+ k_{12} \kappa_{12}^2 + k_{13} \kappa_{21}^2 + k_{14} \kappa_{12} \kappa_{21} + k_{15} \kappa_{23}^2 + k_{16} \kappa_{13}^2 \\ &+ k_{17} \kappa_{11} \kappa_{22}. \end{aligned} \quad (7)$$

Assuming that the strains are infinitesimal, the constants  $k_1, \dots, k_{17}$  for various rods were determined in a series of works by comparing exact solutions of the rod theory to corresponding solutions from three-dimensional linear elasticity.<sup>5</sup> The resulting values of the constants are functions of Poisson's ratio  $\nu$ , Young's modulus  $E$ , and the geometric properties of the undeformed three-dimensional rod-like body that the rod theory is modeling. However, in contrast to these works, here the strains will be assumed finite. For future reference, we recall that Green and Naghdi [11] established that

$$k_1 = k_2 = k_3 = \frac{Ewh(1-\nu)}{4(1+\nu)(1-2\nu)},$$

<sup>4</sup>In particular,  $\mathbf{P} \mathbf{E}_1 = \mathbf{E}_1$ ,  $\mathbf{P} \mathbf{E}_2 = \cos(\theta) \mathbf{E}_2 + \sin(\theta) \mathbf{E}_3$ , and  $\mathbf{P} \mathbf{E}_3 = \cos(\theta) \mathbf{E}_3 - \sin(\theta) \mathbf{E}_2$ .

<sup>5</sup>This series of works was summarized and critiqued by O'Reilly [20].

$$k_7 = k_8 = k_9 = \frac{2\nu}{1-\nu} k_1,$$

$$k_{10} = \frac{Eh^3w}{24(1+\nu)}, \quad k_{11} = \frac{Ehw^3}{24(1+\nu)}, \quad k_{17} = 0. \quad (8)$$

The values of the remaining eight constants are not of interest here.

With the assistance of (4), (7), and (8), we find from (3) that

$$\mathbf{n} = 2k_3 \left( x'x' - 1 + \frac{\nu}{1-\nu} (\gamma_{11} + \gamma_{22}) \right) x' \mathbf{PE}_3,$$

$$\mathbf{k}^1 = 2k_3 \left( \gamma_{11} + \frac{\nu}{1-\nu} (\gamma_{22} + \gamma_{33}) \right) d_1 \mathbf{E}_1 + k_{10} \kappa_{11} d_1' \mathbf{E}_1,$$

$$\mathbf{k}^2 = 2k_3 \left( \gamma_{22} + \frac{\nu}{1-\nu} (\gamma_{11} + \gamma_{33}) \right) d_2 \mathbf{PE}_2 + k_{11} \kappa_{22} d_2' \mathbf{PE}_2,$$

$$\mathbf{m}^1 = k_{10} \kappa_{11} d_1 \mathbf{E}_1,$$

$$\mathbf{m}^2 = k_{11} \kappa_{22} d_2 \mathbf{PE}_2. \quad (9)$$

It is important to note that these expressions are only valid for the steady motions that we are considering.

#### 4 The Boundary Value Problem

For the rod of interest, the end  $\xi=0$  is assumed to be fixed while the end  $\xi=L$  is assumed to be free of tractions. In the model, the hub radius  $r$  is assumed to be zero. For this case, substituting (5), (6), and (9) into (1), we find that six of the nine equations are identically satisfied. The three remaining equations simplify to three scalar ordinary differential equations for  $x(\xi)$  and  $d_\alpha(\xi)$ .

Omitting details, the three equations can be expressed in a compact dimensionless form:

$$\frac{d^2 u}{ds^2} = - \frac{\frac{\nu}{1-\nu} \left( \frac{d\gamma_{11}}{ds} + \frac{d\gamma_{22}}{ds} \right) \left( \frac{du}{ds} + 1 \right) + 2\omega^2(u+s)}{3 \left( \frac{du}{ds} + 1 \right)^2 - 1 + \frac{\nu}{1-\nu} (\gamma_{11} + \gamma_{22})},$$

$$\frac{d^2 \gamma_{11}}{ds^2} = \frac{24}{1-2\nu} \left( \frac{L^2}{h^2} \right) ((1-\nu)\gamma_{11} + \nu(\gamma_{22} + \gamma_{33})),$$

$$\frac{d^2 \gamma_{22}}{ds^2} = \frac{24}{1-2\nu} \left( \frac{L^2}{w^2} \right) ((1-\nu)\gamma_{22} + \nu(\gamma_{11} + \gamma_{33})) - 4 \left( \frac{1-\nu}{1-2\nu} \right) \omega^2, \quad (10)$$

where

$$\gamma_{11} = d_1^2 - 1, \quad \gamma_{22} = d_2^2 - 1, \quad \gamma_{33} = \left( \frac{du}{ds} + 1 \right)^2 - 1. \quad (11)$$

In (10), the following dimensionless quantities were used:

$$s = \frac{\xi}{L}, \quad u = \frac{x-\xi}{L}, \quad \omega = \Omega \sqrt{\frac{L^2 \rho_0 (1-2\nu)(1+\nu)}{E(1-\nu)}}. \quad (12)$$

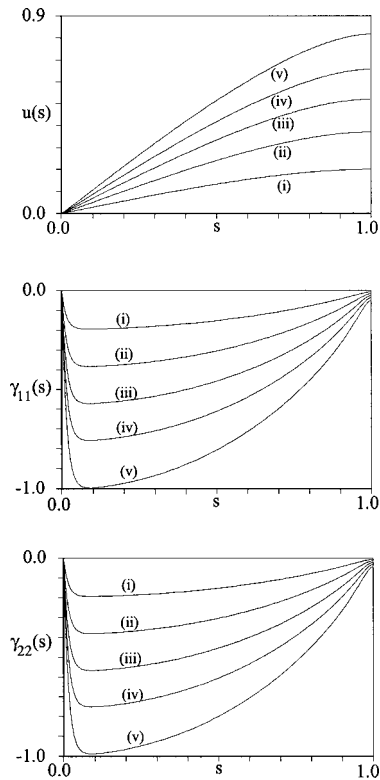
Here,  $Lu$  represents the axial displacement of the centerline of the rod.

Supplementing (10) are two sets of boundary conditions. For the fixed end,

$$u(s=0)=0, \quad \gamma_{11}(s=0)=0, \quad \gamma_{22}(s=0)=0. \quad (13)$$

At the free end,  $\mathbf{n} = \mathbf{m}^\alpha = \mathbf{0}$ . With the assistance of (9)<sub>1,4,5</sub>, we find that these conditions imply that

$$\gamma_{33}(s=1) + \frac{\nu}{1-\nu} (\gamma_{11}(s=1) + \gamma_{22}(s=1)) = 0,$$



**Fig. 3 The dimensionless axial displacement  $u(s)$  and lateral strains  $\gamma_{11}(s)$  and  $\gamma_{22}(s)$  for various values of  $\omega$ : (i),  $\omega^2=0.5$ ; (ii),  $\omega^2=1.0$ ; (iii),  $\omega^2=1.5$ ; (iv),  $\omega^2=2.0$ , and (v),  $\omega^2=2.64$ . For these results,  $\nu=0.3$  and a rod whose length is ten times its height and width was considered:  $h/L=w/L=0.1$ .**

$$\frac{d\gamma_{11}}{ds}(s=1)=0, \quad \frac{d\gamma_{22}}{ds}(s=1)=0. \quad (14)$$

In summary, determination of the steady motions has been reduced to finding solutions  $u(s)$ ,  $\gamma_{11}(s)$ , and  $\gamma_{22}(s)$  of (10) which, for a given  $\Omega$ , satisfy the boundary conditions (13) and (14).

When  $\gamma_{11} = -1$ ,  $d_1 = 0$  and the height of the rod-like body has shrunk to zero. Below this value of  $\gamma_{11}$ , a real-valued  $d_1$  does not exist. Related remarks apply to  $\gamma_{22}$ . For the axial displacement, when  $du/ds = -1$  for an interval of  $s$ , the centerline of the rod-like body in this interval has shrunk from a finite length to zero. Consequently, the following restrictions are used to distinguish physically realistic solutions:

$$\gamma_{11} > -1, \quad \gamma_{22} > -1, \quad \frac{du}{ds} > -1. \quad (15)$$

For the steady motions of interest in this paper, the first two criteria will be violated when  $\omega$  is sufficiently large.

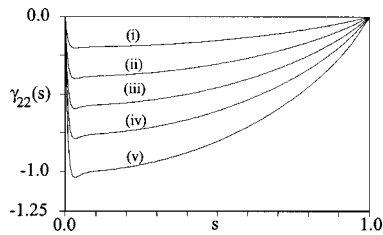
#### 5 The Steady Motions

We now turn to examining the steady motions predicted by the model discussed previously. Our results were obtained using AUTO97 and its continuation algorithm. To use this algorithm, it is necessary to have a solution to the boundary value problem. Fortunately, when  $\omega=0$ , the trivial solution exists:

$$u(s)=0, \quad \gamma_{11}(s)=\gamma_{22}(s)=0. \quad (16)$$

This solution allows us to use  $\omega$  as the bifurcation parameter for the continuation method. In essence, we are using  $\omega$  to construct a homotopy.

For the case of a square rod, solutions to the boundary value problem are presented in Fig. 3 for various values of  $\omega$ . One



**Fig. 4** The dimensionless lateral strain  $\gamma_{22}(s)$  for various values of  $\omega$ : (i),  $\omega^2=0.5$ ; (ii),  $\omega^2=1.0$ ; (iii),  $\omega^2=1.5$ ; (iv),  $\omega^2=2.0$ , and (v),  $\omega^2=2.64$ . For these results,  $\nu=0.3$  and a rectangular cross section was considered:  $h/L=0.1$  and  $w/L=0.05$ .

immediately notices from this figure that severe lateral strains ( $\gamma_{11}$ ,  $\gamma_{22}$ ,  $\kappa_{11}$ , and  $\kappa_{22}$ ) are encountered close to the fixed end. We shall refer to this region as the boundary layer. Its appearance in our numerical works is a consequence of (10) being a stiff system of differential equations. We note in particular that as  $\omega$  approaches a critical value  $\omega_{\text{crit}}$  the width of the cross sections near  $s=0.05$  tends to zero. For  $\nu=0.3$ , we found that  $\omega_{\text{crit}} \approx \sqrt{2.64}$ . When  $\omega$  is greater than  $\omega_{\text{crit}}$ , the solution of the boundary value problem is physically meaningless. It is also evident, from Fig. 3, that the lateral deformation is *not* a negligible effect.

The representative case of  $\gamma_{22}(s)$  for a rectangular rod whose height is twice its width is shown in Fig. 4. The behavior of  $\gamma_{11}(s)$  and  $u(s)$  for this rod are qualitatively similar to those shown in Fig. 3 and are not presented here. For this rod, we note that the lateral strain  $\gamma_{22}$  is larger than in the square case. Indeed, the steady motion predicted by the model when  $\omega^2=2.64$  is not a physically valid solution. Related remarks apply to a rectangular rod whose height is half its width. Indeed, the numerical results for  $\gamma_{11}(s)$  for this case are qualitatively similar to those presented for  $\gamma_{22}(s)$  in Fig. 4. In general, reducing  $h$  and  $w$  while keeping  $L$  constant has the effect of increasing the extreme values of the lateral strains and reducing the size of the boundary layer.

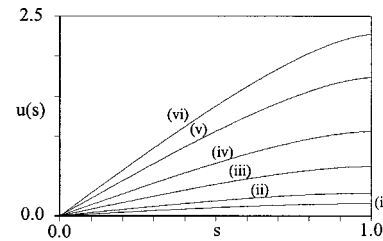
There are two consequences to our choice of strain-energy (7). First, as shown numerically, it implies that the boundary value problem will have a unique solution for each value of  $\omega$ . This uniqueness would probably be eliminated if the strain energy contained quartic terms in the strains. The second consequence of our constitutive selection is that it does not inhibit necking in the boundary layer. To see an instance of this effect, we substitute  $\gamma_{11} = \gamma_{22} = -1$  into (9) and note that this deformation (where the cross sections of the rod are deformed to have zero area) can be achieved with finite forces. Clearly, this is unacceptable. To eliminate this behavior, one could add terms to  $\lambda\psi_1$  to penalize such deformations. These terms would have the functional forms

$$\Psi(z) = \mu \log(1+z)^4, \quad (17)$$

where  $\mu$  is constant and  $z$  denotes any of the strains  $\gamma_{ij}$  and  $\kappa_{ak}$ . Because  $d\Psi/dz(z=0)=0$ , the addition of these terms does not entail a lengthy re-evaluation of  $k_1, \dots, k_{17}$ . However, they complicate the ordinary differential equations for the model significantly.<sup>6</sup>

## 6 The Uniaxial Model

As mentioned previously, the most commonly used model in the literature does not consider the Poisson effect. To illustrate how this uniaxial model relates to our developments, we now turn to presenting such a model. The first model we present is nonlinear and it is then specialized to the classical linear model. There are several methods for establishing the nonlinear model. The first method, which is discussed in the Appendix, models the rod-like body using a string theory. Alternatively, the model can be ob-



**Fig. 5** The dimensionless axial displacement  $u(s)$  predicted by the uniaxial model (18) for various values of  $\omega$ : (i),  $\omega^2=0.5$ ; (ii),  $\omega^2=1.0$ ; (iii),  $\omega^2=2.5$ ; (iv),  $\omega^2=5.0$ ; (v),  $\omega^2=10.0$ ; and (vi),  $\omega^2=15.0$ .

tained from the one discussed in Section 4, by setting  $\gamma_{11}, \gamma_{22}$ , and their derivatives equal to zero, and ignoring Eqs. (10)<sub>2,3</sub>.<sup>7</sup> The uniaxial models discussed in this section can also be considered as special cases of whirling elastic strings. For details on the extensive body of work in this area the reader is referred to Antman [1].

Using either of the aforementioned approaches, the nontrivial equation governing the axial displacement  $Lu$  is

$$\frac{d^2u}{ds^2} = -\frac{2\omega^2(u+s)}{3\left(\frac{du}{ds}+1\right)^2-1}, \quad (18)$$

which is solved subject to the boundary conditions, from (13) and (14),

$$u(s=0)=0, \quad \frac{du}{ds}(s=1)=0. \quad (19)$$

In (18),

$$\omega = \Omega \sqrt{\frac{L^2 \rho_0}{\bar{E}}}. \quad (20)$$

The modulus  $\bar{E}$  in uniaxial studies is often taken to be Young's modulus  $E$  (cf. Berdichevskii [23] and Cesnik and Hodges [24]). This specification ensures that the short wavelength limit of wave propagation coincides with the classical three-dimensional result of Chree [25]. Other authors identify  $\bar{E} = E(1-\nu)/((1+\nu)(1-2\nu))$  (cf. Cremer et al. [26]). It is important to note that the model represented by (18) is a geometrically exact linearly elastic model.

The ordinary differential Eq. (18) is integrable by quadrature. To see this, we use the obvious coordinate transformation  $x/L = u+s$  which renders (18) autonomous. Furthermore, an integral of motion  $I$  of the resulting equation is

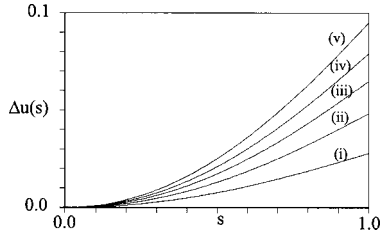
$$I = \frac{3}{4} \left( \frac{1}{L} \frac{dx}{ds} \right)^4 - \frac{1}{2} \left( \frac{1}{L} \frac{dx}{ds} \right)^2 + \omega^2 x^2. \quad (21)$$

Because the level sets of  $I$  for strictly positive  $x$  and  $dx/ds$  do not intersect, the solution to the boundary value problem for a specific  $\omega$  is unique. Using  $I, u(s)$  can be obtained from a quadrature. We remark that the resulting expression for  $u(s)$  would be useful in obtaining the equations governing the vibrations of the rod.

In Fig. 5, representative solutions for various values of  $\omega$  of (18) which satisfy (19) are shown. Based on our numerical investigations, we conjecture that solutions to this model exist for all values of  $\omega$ . In interpreting the results of Fig. 5, it is interesting to note that for a rod-like body composed of steel, where  $E$

<sup>6</sup>The interested reader is referred to Antman [1] and Antman and Carbone [21] for additional discussions on necking and growth conditions in nonlinearly elastic rods.

<sup>7</sup>More precisely, we are now considering a constrained rod theory where the directors are constrained to be constant vectors. In such a theory, the balance laws (1)<sub>2</sub> are identically satisfied by constraint responses. For details on the procedure used to obtain these responses the reader is referred to O'Reilly and Turcotte [22].



**Fig. 6 The dimensionless displacement  $\Delta u(s)$  for various values of  $\omega$ : (i),  $\omega^2=0.5$ ; (ii),  $\omega^2=1.0$ ; (iii),  $\omega^2=1.5$ ; (iv),  $\omega^2=2.0$ , and (v),  $\omega^2=2.64$ . For these results,  $\nu=0.3$  and a square cross section was considered:  $h/L=w/L=0.1$ .**

$=200$  GPa and  $\rho_0=7850$  kg/m<sup>3</sup>,  $\Omega \approx 48200(\omega/L)$  RPM (where  $L$  is in meters). Consequently, some of the solutions presented pertain to large rotational speeds.

To linearize the uniaxial model, the following approximation is used:  $1 \gg du/ds$ . As a result, (18) simplifies to

$$\frac{d^2 u}{ds^2} = -\omega^2(u+s). \quad (22)$$

This is the model investigated by Bhuta and Jones [5] and Brunelle [6]. These authors showed that the solution of this equation, subject to (19), is

$$u(s) = \frac{\sin(\omega s)}{\omega \cos(\omega)} - s. \quad (23)$$

In contrast to (18), they also noted that when  $\omega \rightarrow \pi/2$ ,  $u(1) \rightarrow \infty$ . However, it is also known (see [7]) that the approximation used to establish (23) becomes invalid long before  $\omega = \pi/2$ .

For completeness, we note that Hodges and Bless [7], in an attempt to correct the deficiencies in the linearized uniaxial model, proposed generalizations of (22). Their two generalizations are equivalent to assuming the following strain-energy functions:

$$\begin{aligned} \lambda \psi_s &= \frac{\tilde{E}wh}{2} \left( (\sqrt{\gamma_{33}-1}-1)^2 + \frac{2\beta}{3} (\sqrt{\gamma_{33}-1}-1)^3 \right), \\ \lambda \psi_s &= \frac{\tilde{E}wh}{2} (\log^2(\sqrt{\gamma_{33}-1})). \end{aligned} \quad (24)$$

For certain parameter regimes, the models developed and analyzed in [7] exhibit qualitatively similar results to those presented in Fig. 5.

## 7 Concluding Remarks

Comparing  $u(s)$  shown in Fig. 5 with the corresponding results in Fig. 3, the qualitative similarity of the axial displacement predicted by (10) and the uniaxial model (18) is apparent. It is clearly of interest to examine this issue in more detail. First, we observe that when Poisson's ratio  $\nu=0$ , then the equations governing the boundary value problem associated with (10) decouple. In particular, the axial displacement predicted by (10) and (18) are identical. In addition, for the boundary value problem of interest, Eqs. (10)<sub>2,3</sub> can be solved:

$$\begin{aligned} \gamma_{11}(s) &= 0, \\ \gamma_{22}(s) &= \frac{\omega^2}{6} \left( \frac{w^2}{L^2} \right) \left( 1 - \frac{\cosh\left(\frac{\sqrt{24}L}{w}(1-s)\right)}{\cosh\left(\frac{\sqrt{24}L}{w}\right)} \right). \end{aligned} \quad (25)$$

For nonzero values of  $\nu$ , it was not possible to perform an analytical comparison and, as shown in Fig. 6, we resorted to numerical methods. In this figure,  $\Delta u$  is the difference between the dis-

placement predicted by (10) and (18). Clearly, the model of Section 3 predicts a larger overall extension of the rod-like body compared to the uniaxial model. This difference increases from zero as Poisson's ratio  $\nu$  increases from zero to 0.5. In conclusion, the appropriateness of using uniaxial models to predict  $u(s)$  diminishes as  $\nu$  increases because the lateral deformations increasingly influence this displacement.

The model presented in Section 3 can also be modified to consider the effects of an added tip mass and rotor hub radius. In addition, it can be applied to the oft-studied problem of a rod clamped to a rotating ring (cf. Antman [1], Lakin and Nachman [27], and references therein). This problem is of particular interest because of the presence of buckling instabilities. To determine the vibrational equations for the steady motions of these problems, the procedure discussed in O'Reilly and Turcotte [8] could be followed. However, as we were unable to analytically integrate (10), and the analytically tractable (18) does not generally provide a good approximation to  $u(s)$ , any work on the vibrational response of the whirling rod will have to be performed using numerical methods.

## Acknowledgments

We are indebted to Prof. Andrew J. Szeri for pointing out the feasibility of using numerical continuation methods to solve boundary value problems. The helpful comments of two anonymous reviewers are also gratefully acknowledged.

## Appendix

**The Uniaxial Model and a String Theory.** As mentioned in Section 6, the uniaxial model can be established using a string theory. This theory is used to provide a one-dimensional model for a rod-like body, where the motion of the centerline is of sole importance. Models for the deformation of elastic bars and cables which are used in the literature are based on this theory.<sup>8</sup>

In a string theory, the body is modeled as a material curve which is embedded in Euclidean three-space. The material points of this curve are identified using the convected coordinate  $\xi$ . This coordinate is chosen to be the arc-length parameter of the curve in a fixed reference configuration. The position vector of a point of the material curve at time  $t$  is given by the vector-valued function  $\mathbf{r}(\xi, t)$  where  $\mathbf{r}$  is measured from a fixed origin. The deformation of the curve is measured using the strain

$$\gamma_{33} = \mathbf{r}' \cdot \mathbf{r}' - 1. \quad (A1)$$

This strain also provides the stretch  $\sqrt{1+\gamma_{33}}$  of the curve.

The local form of the balance of linear momentum for the string is (cf. (1)<sub>1</sub>)

$$\mathbf{n}' + \lambda \mathbf{f} = \lambda \ddot{\mathbf{r}}, \quad (A2)$$

where  $\mathbf{n}$  is the contact force,  $\mathbf{f}$  is the assigned force per unit mass, and  $\lambda = \lambda(\xi)$  is the mass per unit undeformed length. For an elastic string, one has the following constitutive relation for  $\mathbf{n}$  (cf. (3)<sub>1</sub>):

$$\mathbf{n} = \lambda \frac{\partial \psi_s}{\partial \mathbf{r}'}, \quad (A3)$$

where  $\lambda \psi_s$  is the strain energy. This energy is assumed to be a function of  $\gamma_{33}$ :

$$\mathbf{n} = 2\lambda \frac{\partial \psi_s}{\partial \gamma_{33}} \mathbf{r}'. \quad (A4)$$

The local form of the balance of angular momentum for the string is identically satisfied by this constitutive relation.

<sup>8</sup>Our developments here are specializations of the theory discussed in O'Reilly and Varadi [28]. Their work was based on the rod theory of Green and Naghdi.

In the development of the uniaxial model discussed in Section 6, the following prescriptions and assumptions were used:

$$\lambda = \rho_0 w h, \quad \mathbf{f} = \mathbf{0}, \quad \mathbf{r} = x \mathbf{P} \mathbf{E}_3, \quad \lambda \psi_s = \frac{\tilde{E} w h}{4} \gamma_{33}^2. \quad (\text{A5})$$

The strain-energy function used here implies that the string is nonlinearly elastic. Substituting (A5) into (A4) and (A1), and invoking the relations  $Lu = x - \xi$ , it is easy to see that the non-trivial differential equation which results is (18).

## References

- [1] Antman, S. S., 1995, *Nonlinear Problems of Elasticity*, Springer-Verlag, New York.
- [2] Leissa, A., 1981, "Vibrational Aspects of Rotating Turbomachinery Blades," *Appl. Mech. Rev.*, **34**, pp. 629–635.
- [3] Rao, J. S., 1987, "Turbomachinery Blade Vibrations," *Shock Vib. Dig.*, **19**, pp. 3–10.
- [4] Wright, A. D., Smith, C. E., Thresher, R. W., and Wang, J. L. C., 1982, "Vibration Modes of Centrifugally Stiffened Beams," *ASME J. Appl. Mech.*, **49**, pp. 197–202.
- [5] Bhuta, P.-G., and Jones, J. P., 1963, "On Axial Vibrations of a Whirling Bar," *J. Acoust. Soc. Am.*, **35**, pp. 217–221.
- [6] Brunelle, E. J., 1971, "Stress Redistribution and Instability of Rotating Beams and Disks," *AIAA J.*, **9**, pp. 758–759.
- [7] Hodges, D. H., and Bless, R. R., 1994, "Axial Instability of Rotating Rods Revisited," *Int. J. Non-Linear Mech.*, **29**, pp. 879–887.
- [8] O'Reilly, O. M., and Turcotte, J. S., 1997, "On the free vibration of a whirling rod," *Proceedings of DETC'97: 1997 ASME Design Engineering Technical Conferences*, Paper Number DETC97VIB4072.
- [9] Green, A. E., and Laws, N., 1966, "A General Theory of Rods," *Proc. R. Soc. London, Ser. A*, **A293**, pp. 145–155.
- [10] Naghdi, P. M., 1982, "Finite Deformation of Elastic Rods and Shells," *Proceedings of the IUTAM Symposium on Finite Elasticity*, D. E. Carlson and R. T. Shield, eds., Martinus Nijhoff, The Hague, pp. 47–104.
- [11] Green, A. E., and Naghdi, P. M., 1979, "On Thermal Effects in the Theory of Rods," *Int. J. Solids Struct.*, **15**, pp. 829–853.
- [12] Naghdi, P. M., and Rubin, M. B., 1989, "On the Significance of Normal Cross-Sectional Extension in Beam Theory With Application to Contact Problems," *Int. J. Solids Struct.*, **25**, pp. 249–265.
- [13] Nordenholz, T. R., and O'Reilly, O. M., 1997, "On Steady Motions of an Elastic Rod With Application to Contact Problems," *Int. J. Solids Struct.*, **34**, pp. 1123–1143; 3211–3212.
- [14] Krishnaswamy, S., and Batra, R. C., 1998, "On Extensional Vibration Modes of Elastic Rods of Finite Length Which Include the Effect of Lateral Deformation," *J. Sound Vib.*, **215**, pp. 577–586.
- [15] O'Reilly, O. M., 2001, "On Coupled Longitudinal and Lateral Vibrations of Elastic Rods," *J. Sound Vib.*, to appear.
- [16] Doedel, E. J., Champneys, A. R., Fairgrieve, T. F., Kuznetsov, Y. A., Sandstede, B., and Wang, X., 1997, "AUTO97: Continuation and bifurcation software for ordinary differential equations (with HomCont)," Department of Computer Science, Concordia University, Montreal, Canada.
- [17] Doedel, E. J., 1997, "Nonlinear Numerics," *J. Franklin Inst.*, **334B**, pp. 1049–1073.
- [18] Seydel, R., 1988, *From Equilibrium to Chaos: Practical Bifurcation and Stability Analysis*, Elsevier, New York.
- [19] Green, A. E., Laws, N., and Naghdi, P. M., 1967, "A Linear Theory of Straight Elastic Rods," *Arch. Ration. Mech. Anal.*, **25**, pp. 285–298.
- [20] O'Reilly, O. M., 1998, "On Constitutive Relations for Elastic Rods," *Int. J. Solids Struct.*, **35**, pp. 1009–1024.
- [21] Antman, S. S., and Carbone, E. R., 1977, "Shear and Necking Instabilities in Nonlinear Elasticity," *J. Elast.*, **7**, pp. 125–151.
- [22] O'Reilly, O. M., and Turcotte, J. S., 1997, "Elastic Rods With Moderate Rotation," *J. Elast.*, **48**, pp. 193–216.
- [23] Berdichevskii, V. L., 1981, "On the energy of an elastic rod," *J. Appl. Math. Mech.*, **45**, 518–529.
- [24] Cesnik, C. E. S., and Hodges, D. H., 1993, "Variational-asymptotical analysis of initially curved and twisted composite beams," *Appl. Mech. Rev.*, **46(11/2)**, S211–S220.
- [25] Chree, C., 1889, "On Longitudinal Vibrations," *Quart. J. Pure Appl. Math.*, **23**, 317–342.
- [26] Cremer, L., Heckl, M., and Ungar E. E., 1988, *Structure-Borne Sound: Structural Vibrations and Sound Radiation at Audio Frequencies*, 2nd Ed., Springer-Verlag, New York.
- [27] Lakin, D. A., and Nachman, A., 1979, "Vibration and Buckling of Rotating Flexible Rods at Transitional Parameter Values," *J. Eng. Math.*, **13**, pp. 339–346.
- [28] O'Reilly, O. M., and Varadi, P. C., 1999, "A Treatment of Shocks in One-Dimensional Thermomechanical Media," *Continuum Mech. Thermodyn.*, **11**, pp. 339–352.



# Size Effects in a Slowly Sheared Granular Media

**S. J. Antony**

Lecturer

e-mail: S.J.Antony@leeds.ac.uk

**M. Ghadiri**

Professor

e-mail: M.Ghadiri@leeds.ac.uk

Department of Chemical Engineering,  
School of Process,  
Environmental and Materials Engineering,  
University of Leeds,  
Leeds LS2 9JT, UK

*In this paper, we analyze the nature of stress distribution experienced by large particles in a dense granular media subjected to slow shearing, using the distinct element method. The particles were generated in a three-dimensional cuboidal periodic cell in which a large solid spherical particle was submerged ("submerged particle") at the center of a bed of monodispersed spherical particles. The granular systems with different size ratio (i.e., the ratio of the diameter of submerged particle to that of the surrounding monodispersed particles) were subjected to quasi-static shearing under constant mean stress condition. The evolution of stress distribution in the submerged particle during shearing was carefully tracked down and presented here. The nature of stress distribution is bifurcated into two components, viz., (i) hydrostatic and (ii) deviatoric components. It has been shown that, for size ratio greater than c.a. 10, the nature of stress distribution in the submerged particle is hydrostatically dominant (increases the 'fluidity'). For smaller size ratios, the nature of stress distribution in the submerged particle is dominantly deviatoric. [DOI: 10.1115/1.1387443]*

## 1 Introduction

In recent years, granular materials have received an increasing attention due to their technologically challenging behavior in diverse applications in fields such as civil, mechanical, and process engineering. This has been facilitated greatly by the rapid growth of computer power, which has enabled an insight to be gained of the complex and often mysterious behavior of granular materials by numerical simulations. In this paper, we report on the results of our investigations on the size effects of particles in granular system subjected to slow shear deformation. The phenomenon under study is relevant to dispersion, milling, grinding, and other industrial processes. The vibration induced size segregation problem, also known as "brazil-nut effect" has been the subject of several investigations ([1]). When a container having larger particles embedded in smaller granular particles is vibrated, for example vertically, the bigger particles tend to move toward the top of the container. Conflicting reasons have been attached to this phenomena ([2]). Nevertheless, two-dimensional studies have indicated that ([3]) there exists a threshold size ratio (diameter ratio of large particle to the surrounding mono-dispersed particle), above which, the movement of larger particle increases. For smaller size ratio (less than about three), no ascent of the large particle (intruder) was observed; for size ratio 5.3, the intruder undergoes an intermittent ascent; and for size ratio greater than c.a. 10, the intruder ascends continuously and hence the fluidity (movement) of the large particles becomes higher. However, we have not quite understood about how does an increase in particle size mysteriously favors such an easy movement.

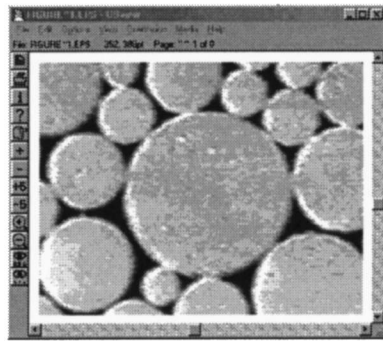
Recent studies on the size effects in compacted beds based on experiments and micromechanical modeling have shown interesting characteristics. Bonnenfant et al. [4] have studied the effect of presence of hard inclusions (glass) in a polymethylmethacrylate (PMMA) matrix on compaction in a triaxial cell. Their experimental and analytical studies have shown no influence of large inclusions on the global stiffness for the size ratio of the inclusion equal to 2 as considered by them. Tsoungui et al. [5] have adopted

two-dimensional molecular dynamics simulations (with elastic interactions between grains) to study the crushing mechanism of grains inside a granular material under oedometric compression. They have also carried out experiments on the two-dimensional packing of moulding plaster disks under oedometric compression and the video-captured images of their experiments is presented in Fig. 1. They observed the geometric evolution of interparticle contacts inside the packing during the crushing process. The experimental and numerical observations have shown a saturation regime, where large grains, surrounded by smaller grains experience less deviatoric stress and more hydrostatic stress. Despite an increase of external pressure, the large grain fracture was impossible because of this hydrostatic pressure (Fig. 1(b)). However, they have not quantified at what size scale the hydrostatic effect of the grains would become dominant.

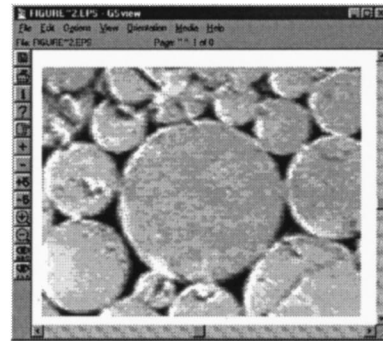
Gundepudi et al. [6] have investigated analytically the state of three-dimensional stresses in a single sphere for some selected cases of loading conditions. They have observed that as the number of contact points increases, the maximum tensile stress in the sphere, in general, decreases and the state of stress inside the sphere approaches hydrostatic compression. Based on this, they carried out some tests on glass and alumina spheres and correlated the observations with an analytical study. Recently McDowell and Bolton [7] have studied the micromechanical behavior of crushable granular soils and have presented interesting arguments on the breakage process. At first, it may be anticipated that the larger particles are most likely to fracture. However, the smaller particles are likely to have fewest contacts and there are two opposing effects on the particle survival; size and coordination number. If the effect of the coordination number dominates over that of the particle size in the evolution of stresses in the aggregate, then the smallest particles would always fracture, although they have not explicitly attributed this effect to the hydrostatic nature of the larger particles.

From the literature presented above, we can identify that the size effects in granular media have a significant contribution to the evolution of stress distribution. As far as we are aware, a clear description of the nature of three-dimensional evolution of stress distribution in a slowly sheared granular media containing one or more larger particles is not available in the literature. In this paper, we analyze the nature of stress distribution for large size particles in granular media subjected to slow shearing using the discrete element method. The particles were generated in a periodic cell having a system of particles in which a large solid spherical particle is submerged ("submerged particle") at the center of a bed 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 by the ASME Applied Mechanics Division, April 18, 2000; final revision, January 8, 2001. Associate Editor: D. Kouris. Discussion on the paper should be addressed to the 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.



(a)



(b)

**Fig. 1 (a) Enlargement of region of granular material showing the hydrostatic effect on a large grain surrounded with small grains, (b) despite the increase of external pressure, the large grain fracture becomes difficult because of this hydrostatic pressure ([5])**

monodispersed spherical particles. The granular systems with different size ratio (i.e., the ratio of the diameter of submerged particle “ $D$ ” to that of the surrounding monodispersed particles “ $d$ ”) were subjected to quasi-static shearing under constant mean stress condition. The hydrostatic and deviatoric stress states of the submerged particle have been evaluated during shearing. The critical size ratio at which the nature of stress distribution in a submerged particle becomes hydrostatically dominant has been identified and presented below.

## 2 Simulations

The simulations were carried out using discrete element method which was originally developed by Cundall and Strack [8]. The interactions between contiguous particles are modeled as a dynamic process and the time evolution of the particles is advanced using an explicit finite difference scheme. The interactions between the neighboring particles are modeled by algorithms based on theoretical contact mechanics provided by Thornton and Yin [9] and Thornton [10]. The time-step used in the simulation is based on the minimum particle size and the Rayleigh wave speed ([11]). For detailed information about the numerical methodology, the readers could refer to Cundall and Strack [8]. The advantage of applying the discrete element method to granular materials is its ability to give more information about what happens inside the system. For example, investigations based on discrete element method simulations have provided insights into the evolution of normal contact force distribution in granular media under quasi-static shearing ([12,13]). This enhances our understanding of the physics of the granular media. In the present situation, we would first identify the contact forces acting on the submerged particle during shearing. Having known the boundary forces, we would address the total stress tensor of the submerged particle and later bifurcated into the hydrostatic and deviatoric components and analyzed.

The simulations were carried out in a three-dimensional cuboidal periodic cell. A large spherical particle was generated at the

center of the periodic cell and was surrounded by 5000 randomly generated mono-sized spherical particles (Fig. 2) with diameter 0.01 mm. All the particles were given the following properties: Young’s modulus,  $E=70$  GPa (hard), Poisson’s ratio  $\nu=0.3$ , coefficient of interparticle friction  $\mu=0.3$  or 0.01, and interface energy  $\Gamma=0.6$  J/m<sup>2</sup>. After the particles were initially generated, the system was isotropically compressed until a mean stress  $p=1$  kPa was obtained using a servo-control algorithm of the following form ([13]):

$$\dot{\epsilon} = \dot{\epsilon} + g(p_d - p_o). \quad (1)$$

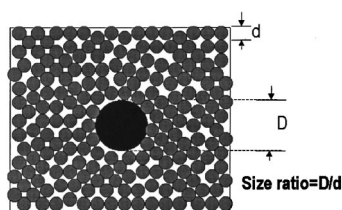
In the above equation,  $\dot{\epsilon}$  is the strain rate,  $p_d$  is the desired isotropic stress and  $p_o$  is the obtained isotropic stress at a particular stage. An initial strain rate of  $10^{-5} \text{ s}^{-1}$  was specified which was progressively modified according to (1) using a value for the gain parameter ( $g$ ) calculated from

$$g = (\dot{\epsilon} / (p_d - p_o))_{\text{initial}} \quad (2)$$

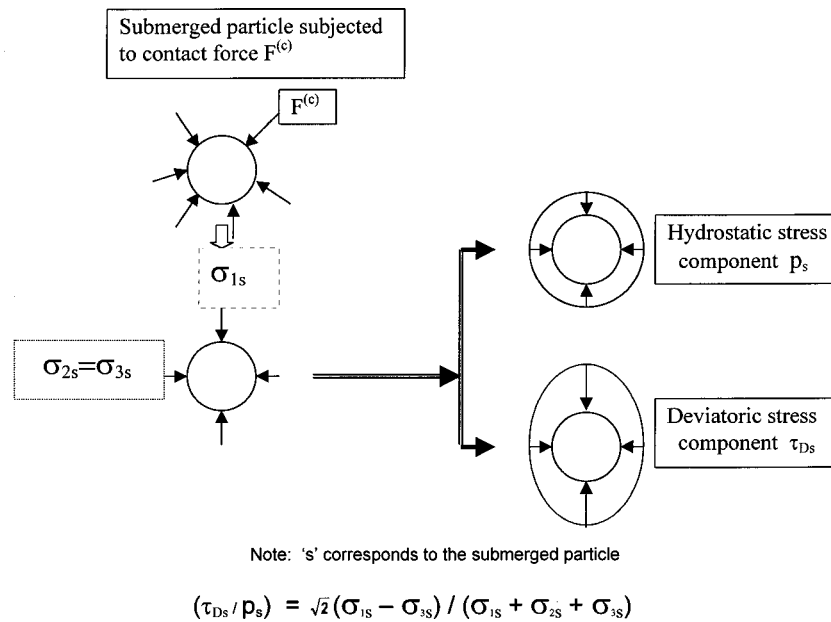
The Eqs. (1) and (2) ensure that the strain rate decreases as the stress difference decreases and that the strain rate tends to zero as the calculated value of the isotropic stress approaches the desired value, thereby bringing the system to equilibrium. In order to get a stable system at the desired stress level, calculation cycles were continued until the solid fraction and coordination number had attained constant values. This procedure was used to progressively raise the isotropic stress to create a sample with an isotropic stress level of 100 kPa. At the end of the isotropic compression, the microstructure of the samples was isotropic. At this stage, the solid fraction and mechanical coordination number (average number of load-bearing contacts) of the samples considered in this study were  $0.650 \pm 0.017$  and  $5.83 \pm 0.26$ , respectively. For shearing, a strain rate of  $10^{-5} \text{ s}^{-1}$  was employed in the simulations. The samples were subjected to the axisymmetric compression test ( $\sigma_1 > \sigma_2 = \sigma_3$ ) ([12,13]). During shearing, the mean stress  $p = (\sigma_1 + \sigma_2 + \sigma_3)/3$  was maintained constant at 100 kPa using the servo-control algorithm.

## 3 Results and Discussion

As mentioned earlier, the contact forces ( $F^{(c)}$ ) acting on the submerged particle and hence the stress tensor of the submerged particle are first calculated. Then, the principal components of the stress tensor for the submerged particle is calculated, followed by the bifurcation of the stress tensor into two components viz., (i) hydrostatic ( $p_s$ ) and (ii) deviatoric ( $\tau_{Ds}$ ) components. This process is schematically presented in Fig. 3. The evolution of these stress components in the submerged particle during shearing is presented here.

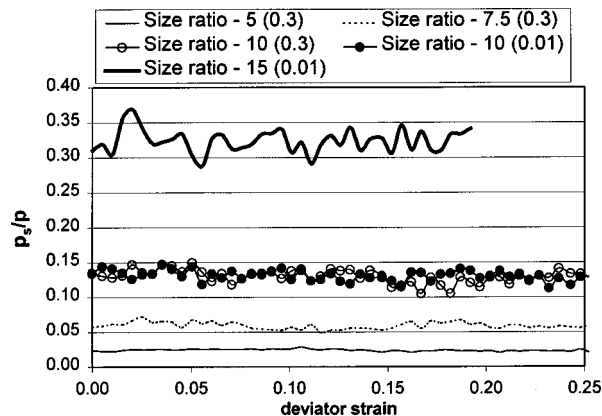


**Fig. 2 Schematic diagram showing the sectional view passing through the center of the periodic cell with submerged particle**



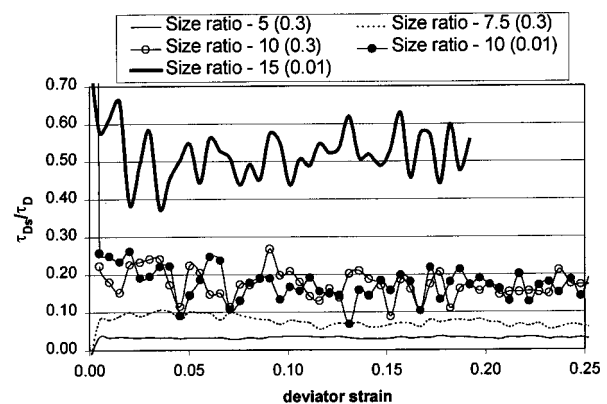
**Fig. 3 Illustrative diagram showing the state of stress and its bifurcation in the submerged particle**

Figures 4 and 5 show the variation of the hydrostatic and deviatoric stress components, respectively, for the submerged particle during shearing (deviator strain =  $\epsilon_1 - \epsilon_3$ ). In these graphs, the hydrostatic and deviatoric stress components have been normalized to that of the entire granular system ( $p$  and  $\tau_D$ , respectively). From these figures it can be observed that, as the size ratio increases, the components of both the hydrostatic and deviatoric contribution to the submerged particle increases. In these figures, the numbers indicated within brackets in the legend correspond to the value of interparticle friction of granular system. It shall be noted that, the kind of computations carried out here require substantial amount of computing power and hence we limited our investigation to the case of submerged particle with size ratio 10 to see the effect of interparticle friction. It can be seen that, for the values of coefficient of friction considered here, the behavior has not changed substantially. Hence, we carried out our analysis further for the case of a granular system with size ratio 15 with the coefficient of friction as 0.01 only. The reason that we selected to study the effect of interface friction with the system having size ratio as 10 would be evident later.

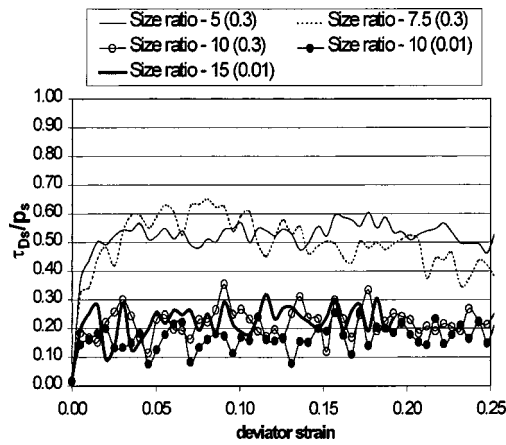


**Fig. 4 Variation of the hydrostatic stress component of submerged particle normalized to that of the entire granular system during shearing**

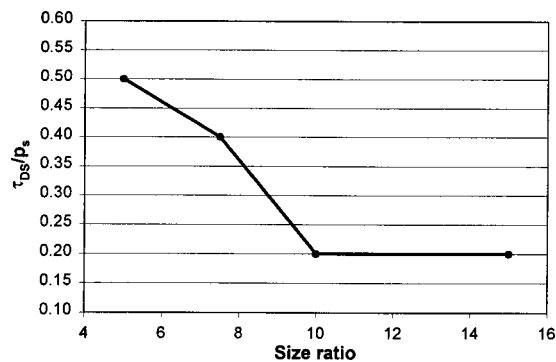
The most striking feature of the nature of stress distribution in the submerged particle during shearing is presented in Fig. 6. Figure 6 shows the variation of the ratio of deviatoric stress component to the hydrostatic stress component of submerged particle during shearing. It shall be noted that, as the size ratio of the submerged particle increases, the values of  $\tau_{Ds}/p_s$  decreases, thereby, showing a growing dominance of the hydrostatic component of the submerged particle for an increase in size ratio. We plot the values  $\tau_{Ds}/p_s$  corresponding to the steady state (saturation regime; deviator strain c.a. 0.25) and presented in Figure 7. From Figs. 6 and 7 it can be concluded that, in a dense granular packing subjected to slow shearing, the nature of stress distribution in particles at steady state becomes increasingly hydrostatic as the particle size ratio is increased. It is fully dominated by the hydrostatic stress component for particles with a size ratio equal to or greater than c.a. 10. Hence it is reasonable to expect that, for particles with size ratio equal to or greater than c.a. 10 in a dense system, the fluidity of the particles would increase. At the critical size ratio, the submerged particle attained a value of  $\tau_{Ds}/p_s$  equal to c.a. 0.2 at saturation regime.



**Fig. 5 Variation of the deviatoric stress component of submerged particle normalized to that of the entire granular system during shearing**



**Fig. 6 Variation of the ratio of deviatoric stress component to the hydrostatic stress component of submerged particle during shearing**



**Fig. 7 Variation of the ratio of deviatoric stress component to the hydrostatic stress component with size ratio of submerged particle at steady state**

Although not directly investigated, it would be interesting to comment about a correlation of the influence of critical size ratio in an analogical problem of aggregated suspensions in shear flow. In rheological problems, theoretical analysis based on continuum theories for aggregated concentrated suspensions under slow shear flow is applicable if the size of suspension particles is an order higher than that of fluid particles. Although it is customary to treat the size of suspension an order higher than that of fluid particle, to the best of our knowledge, no precise reason is available in the literature to justify this. The present study could, in a way, provide a justification of keeping the critical size ratio as 10, as at this size ratio, the nature of stress distribution in the particle is dominantly hydrostatic.

#### 4 Concluding Remarks

The evolution of stress distribution in a large particle submerged in a dense granular system (hard) has been investigated

using discrete element method simulations. The particles were sheared in a periodic cell, subjected to axisymmetric compression test under constant mean stress condition. Results show that as the size of the submerged particle increases, the number of contacts would grow and contact forces at each contact contributes to the stress tensor. This contribution has been bifurcated into hydrostatic and deviatoric stress components. As the size ratio increases, the hydrostatic stress component of submerged particle increases. For size ratio greater than c.a. 10, the nature of stress distribution in the submerged particle at saturation regime is essentially hydrostatic ( $\tau_{Ds}/p_s = \text{c.a. } 0.2$ ). The ongoing investigations on the related features of this study such as the fabric nature of contacts, macroscopic stresses, and influence of packing is reported elsewhere ([14]). It is to be said that the present simulations were carried out with relatively small number of particles due to the limitations in computer power available. However, further support from the researchers is necessary to confirm the findings reported in this paper with simulations having large number of particles, thus allowing statistically more accurate calculations.

#### Acknowledgments

This work has been supported by EPSRC and ICI Strategic Technology Group Technology Ltd., Wilton, UK (Grant No. GR/M33907). The support from Prof. R. Oliver and Dr. D. Sutton, ICI, is gratefully acknowledged. The authors thank Prof. J. C. Charmet for providing Fig. 1 and Dr. C. Thornton for making useful comments on the manuscript.

#### References

- [1] Huntley, J. M., 1998, "Fluidization, Segregation and Stress Propagation in Granular Materials," *Philos. Trans. R. Soc. London, Ser. A*, **356**, pp. 2569–2590.
- [2] Ehrichs, E. E., Flint, J. K., Jager, H. M., Knight, J. B., Nagel, S. R., Karczmar, G. S., and Kuperman, V. Y., 1998, "Convection in Vertically Vibrated Granular Materials," *Philos. Trans. R. Soc. London, Ser. A*, **356**, pp. 2561–2567.
- [3] Duran, J., Mazozi, T., Clement, E., and Rajchenbach, J., 1994, "Size Segregation in a Two Dimensional Sample: Convection and Arching Effects," *Phys. Rev. E*, **50**, No. 6, pp. 5138–5141.
- [4] Bonnenfant, D., Mazerolle, F., and Suquet, P., 1998, "Compaction of Powders Containing Hard Inclusions: Experiments and Micromechanical Modelling," *Mech. Mater.*, **29**, pp. 93–109.
- [5] Tsoungui, O., Vallet, D., and Charmet, J. C., 1999, "Numerical Model of Crushing of Grains Inside Two-Dimensional Granular Materials," *Powder Technol.*, **105**, pp. 190–198.
- [6] Gundepudi, M. K., Sankar, B. V., Mecholsky, J. J., and Clupper, D. C., 1997, "Stress Analysis of Brittle Spheres Under Multiaxial Loading," *Powder Technol.*, **94**, pp. 153–161.
- [7] McDowell, G. R., and Bolton, M. D., 1998, "On the Micromechanics of Crushable Aggregates," *Geotechnique*, **48**, No. 5, pp. 667–679.
- [8] Cundall, P. A., and Strack, O. D. L., 1979, "A Discrete Numerical Model for Granular Assemblies," *Geotechnique*, **29**, pp. 47–65.
- [9] Thornton, C., and Yin, K. K., 1991, "Impact of Elastic Spheres With and Without Adhesion," *Powder Technol.*, **65**, pp. 153–166.
- [10] Thornton, C., 1997, "Coefficient of Restitution for Collinear Collisions of Elastic-Perfectly Plastic Spheres," *ASME J. Appl. Mech.*, **64**, pp. 383–386.
- [11] Thornton, C., and Randall, C. W., 1988, *Micromechanics of Granular Materials*, M. Satake, and J. T. Jenkins, eds., Elsevier, Amsterdam, pp. 133–142.
- [12] Antony, S. J., 2001, "Evolution of Force Distribution in Three Dimensional Granular Media," *Phys. Rev. E*, **63**, Paper No. 011302.
- [13] Thornton, C., and Antony, S. J., 1998, "Quasi-static Deformation of Particulate Media," *Philos. Trans. R. Soc. London, Ser. A*, **356**, pp. 2763–2782.
- [14] Antony, S. J., and Ghadiri, M., 2001, "Shear Resistance of Granular Media Containing Large Inclusions: DEM Simulations," *First MIT Conference on Computational Fluid and Solid Mechanics*, Paper No. 439, Elsevier, Amsterdam, in press.



# A New Method for Nonlinear Two-Point Boundary Value Problems in Solid Mechanics

L. S. Ramachandra<sup>1</sup>

e-mail: lsr@civil.iitkgp.ernet.in

D. Roy

Department of Civil Engineering,  
Indian Institute of Technology,  
Kharagpur 721 302, India

*A local and conditional linearization of vector fields, referred to as locally transversal linearization (LTL), is developed for accurately solving nonlinear and/or nonintegrable boundary value problems governed by ordinary differential equations. The locally linearized vector field is such that solution manifolds of the linearized equation transversally intersect those of the nonlinear BVP at a set of chosen points along the axis of the only independent variable. Within the framework of the LTL method, a BVP is treated as a constrained dynamical system, which in turn is posed as an initial value problem. (IVP) In the process, the LTL method replaces the discretized solution of a given system of nonlinear ODEs by that of a system of coupled nonlinear algebraic equations in terms of certain unknown solution parameters at these chosen points. A higher order version of the LTL method, with improved path sensitivity, is also considered wherein the dimension of the linearized equation needs to be increased. Finally, the procedure is used to determine post-buckling equilibrium paths of a geometrically nonlinear column with and without imperfections. Moreover, deflections of a tip-loaded nonlinear cantilever beam are also obtained. Comparisons with exact solutions, whenever available, and other approximate solutions demonstrate the remarkable accuracy of the proposed LTL method.*

[DOI: 10.1115/1.1387444]

## 1 Introduction

There exist many numerical and semi-analytical techniques for solving nonlinear boundary value problems. For instance, different versions of Newton's method ([1–3]) happen to be very popular among the researchers. Recently a new iterative method ([4]) has been proposed based on Pade approximations for numerical analysis of nonlinear problems. In a companion paper ([5]) new computational methods have been proposed to improve the iterative procedure based on the Newton-Raphson procedure. These techniques, based on the first-order perturbation technique, allow to define an adaptive step strategy and to improve the trial solution for the first iteration at each step. The power series method ([6]) and the double Fourier series ([7]) method have also been employed in the solution of large deflection problems of circular and rectangular plates, respectively. The perturbation technique ([8,9]) is a very popular tool among the researchers working in the broad area of buckling and post-buckling of structures. In an interesting paper ([10]) asymptotic-numerical methods, based on perturbation techniques, are proposed to solve nonlinear boundary value problems. The large deflection problem of circular plates has been solved by the dynamic relaxation method ([11,12]). Budiansky [13] and Thurston [14] have analyzed the nonlinear differential equations of the thin shallow spherical shells by converting the pair of differential equations into a pair of integral equations which are then solved adopting numerical techniques. Kai-yuan Yeh et al. [15,16] have proposed an analytical solution to the von Karman's equations of a circular plate under a concentrated load. The finite element method coupled with iterative techniques has been extensively used to solve geometrically and materially nonlinear problems ([17–19]). Pollandt [20] has solved

von Karman's plate equations using a boundary element method along with radial basis functions. Recently He [21] has proposed a new perturbation method coupled with the homotopy technique for the solution of nonlinear problems. In contrast to the traditional perturbation methods, the proposed method does not require a small parameter in the equation. However, even though the method is apparently capable of handling strong nonlinearity, it is limited by the constraint that the basic trajectory-type (such as periodic or exponentially increasing or decreasing) for the nonlinear problem has to be similar to that of the associated linear problem, which corresponds to the perturbation parameter  $p=0$ . In another recent development, Gao [22] has developed a general duality principle to search for the extremas (along with their stability types) in nonconvex energy expressions, used for post-buckling analysis of thick beams. There is nevertheless no precise and widely applicable method available to date to solve different kinds of nonlinear boundary value problems with continuous or piecewise continuous vector fields. It is believed that the present method, that attempts to "exactly" solve the nonlinear ordinary differential equations at a countable set of points, should help us better understand the behavior of typically nonlinear systems in structural mechanics and the impact of nonlinearity on these system responses.

It is usually not possible to replace the nonlinear governing differential equations in terms of some linearized equations even over small step sizes. The basic reason behind this is that tangent spaces in nonlinear differential equations are themselves functions of the dependent variables. A way out of this problem is, however, found in the proposed locally transversal linearization (LTL) method, wherein solution manifolds of linearized equations are made to intersect transversally those of the given nonlinear equations at a set of points along the independent axis where the solution vectors need to be determined. The most important feature of this method is that it attempts to precisely satisfy the governing nonlinear ordinary differential equations of the physical problem at these chosen set of points along the axis of the independent variable. The LTL-based solution may, however, fall off considerably from the actual path in between any two such successive points of intersections. The LTL method finally reduces a given

<sup>1</sup>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, July 26, 2000; final revision, May 17, 2001. Editor: M. Ortiz. Discussion on the paper should be addressed to the 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.



set of nonlinear ordinary differential equations to a collection of coupled sets of nonlinear algebraic equations in terms of the unknown solution vectors at each of the chosen points along the independent axis. In the present study, the principle of LTL is applied to nonlinear boundary value problems, governing static deflections of beams, by treating them as conditionally dynamical systems. In other words, a boundary value problem is looked upon as a constrained initial value problem, wherein one of the boundaries (say, the left boundary) of the domain is treated as the initial point to start integrating the ordinary differential equations. Towards this, a precise knowledge of all the state variables at the left boundary is needed. However, to begin with, only a few of these conditions are known at the left boundary while the rest are specified at the right boundary of the domain. In this study, the unknown initial conditions at the left boundary are treated as unknown variables to be determined such that the resulting solution satisfies prescribed boundary conditions at the right boundary. Note that an usual dynamical problem, modeled as an initial value problem, is nonanticipative in that all the initial conditions are precisely known at the starting time, say  $t = t_i$ . Thus the solution vector at time  $t = t_{i+1}$  may be calculated without a prior knowledge of the solution at  $t = t_{i+2}$ . On the other hand, a boundary value problem, treated as a constrained initial value problem, is obviously anticipative and thus, adopting the LTL procedure, equations at all the transversal intersections must be written down before any solution can be obtained. This implies that the nonlinear algebraic equations at different intersections points are all coupled in this case. In the process, the conditionally constructed LTL solution automatically satisfies the conditions at the right boundary of the domain.

A limited numerical work has been carried out to demonstrate the capabilities of the proposed LTL technique. First, deflections and slopes of a tip loaded cantilever beam are obtained. Next, the post-buckling equilibrium paths of a geometrically nonlinear column with and without imperfections are obtained with various boundary conditions. Comparisons are made with exact solutions, whenever available, and other approximate solutions and in the process, high numerical accuracy of the proposed method is readily brought out. A higher order version of the LTL method with still improved accuracy is also constructed, wherein the dimension of the linearized equation is suitably increased for a higher path sensitivity.

## 2 The Methodology

The basic LTL methodology, as adapted for solving nonlinear boundary value problems, is described in this section. For the present study, only the subclass of boundary value problems governed by the following system of ordinary differential equations is considered:

$$\frac{dw}{ds} = A(s)w + Q(w, s) + f(s) = V(w, s) \quad (1)$$

where  $w = \{^1w, ^2w, \dots, ^nw\}^T \in R^n$ ,  $A(s)$  is an  $n \times n$  state-independent coefficient matrix associated with the linear terms,  $f(s): D \subset R \rightarrow R^n$  is the external (nonparametric) force vector,  $Q(w, s): D \times R^n \rightarrow R^n$  is that part of the vector field which is nonlinear in  $s$ , and  $V(w, s)$  stands for the entire vector field. Let the given beam length be divided into  $r$  segments by nodes ( $r+1$  nodes) and the nodal coordinates (along  $s$ ) be strictly ordered such that  $0 = s_0 < s_1 < s_2 < \dots < s_i < \dots < s_r = L$  and  $h_i = s_i - s_{i-1}$  where  $i \in Z^+$ . Note that the only independent variable in such systems is denoted as  $s \in R$  and the solution domain  $D = [s_0, s_r]$  is a closed and compact subset of  $R$ . The boundary conditions are nontrivially defined on both the boundary points of  $D$ . In particular, for an  $n$ -dimensional boundary value problem,  $m$  ( $0 < m < n$ ) boundary conditions,  $^jw(s_0) \triangleq ^jw_0$ ,  $j = j_1, j_2, \dots, j_m$ , are known at the left end  $s_0$  and the remaining  $(n-m)$  of them, i.e.,  $^kw(s_r) \triangleq ^kw_r$ ,  $k = k_1, k_2, \dots, k_{n-m}$ , are known at the right end

$s_r$ . To solve the above nonlinear ordinary differential equation within the framework of LTL, one needs to treat the boundary value problem as a constrained initial value problem, wherein the left boundary  $s_0$  of  $D$  is treated as the initial point to start integrating the ordinary differential equations. Now, towards integrating these governing system of ordinary differential equations, a precise knowledge of all the  $n$ -independent state variables at  $s = s_0$  is needed. However, only  $m$  ( $0 < m < n$ ) of these conditions are known at  $s = s_0$  and the remaining  $(n-m)$  are only provided at the right boundary  $s_r$  of  $D$ . Thus the remaining  $(n-m)$  unspecified initial conditions,  $^lw_0$ ,  $l = l_1, l_2, \dots, l_{n-m}$ , enter the LTL system as unknown variables to be so determined as to satisfy the far-end boundary conditions  $^kw_r$ ,  $k = k_1, k_2, \dots, k_{n-m}$ . It is now needed to replace the nonlinear boundary value problem as described by Eq. (1) by a set of  $r$  conditionally linearized system of ordinary differential equations wherein the solution manifold of the  $i$ th linear system of ordinary differential equations transversally intersects the original equilibrium path at  $s = s_i$ . Moreover the  $i$ th linear system satisfies  $n$  initial conditions at  $s = s_{i-1}$  (not all of which are known a priori) and should preferable (but not necessarily) represent the original solution  $w(s)$ , the  $i$ th length segment  $D_i = (s_{i-1}, s_i]$ , where  $\cup_i D_i = D$  and  $\cap_i D_i = \phi$ , i.e., the null set. Let  $\bar{w}^{(i)}(s) = \{^1\bar{w}^{(i)}(s) \ ^2\bar{w}^{(i)}(s) \ \dots \ ^n\bar{w}^{(i)}(s)\}^T$  denote the vector solution flow of the  $i$ th linearized system of ordinary differential equations. It should be apparent that these linearized equations may at best be only conditionally linear, being conditioned upon their exactly satisfying the given nonlinear differential equation at a discrete number of points,  $S = \{s_i | i = 1, 2, 3, \dots\}$ , but not necessarily elsewhere. Since there are an uncountably infinite distinct linear ordinary differential equations transversal to a given (nonlinear) ordinary differential equation the procedures to derive the conditionally linear system of ordinary differential equations are consequently nonunique and uncountably many ([23–25]). Here a convenient and easily adaptable methodology of linearization using LTL, henceforth labeled as the LTL-1 scheme, is first described.

For a complete clarity of defining the intersection point and ease of implementation, it is desirable to derive the  $i$ th locally linearized system such that it is also  $n$ -dimensional and is preferably obtainable from the given nonlinear system with the simplest and least alterations. Towards this, the  $i$ th linearized ordinary differential equations are constructed by recasting Eq. (1) over  $D_i$  as

$$\frac{d\bar{w}^{(i)}}{ds} = A(s_i)\bar{w}^{(i)} + B(w_i, s_i)\bar{w}^{(i)} + f(s) = \bar{V}(\bar{w}, s). \quad (2)$$

Note that elements of the arbitrarily chosen matrix  $B(w_i, s_i)$  are functions of the still unknown solution vector  $w_i \triangleq w(s_i)$  and thus Eq. (2) is clearly conditionally linear with constant coefficients provided that the vector  $w_i$  is precisely known. Since Eq. (2) is required to satisfy Eq. (1) at the left end of the domain segment  $D_i$ , the initial condition vector to Eq. (2) is  $\bar{w}(s_{i-1}) \triangleq \bar{w}_{i-1} = w_{i-1}$ . At this stage it will be useful to construct the locally linearized variational equation associated with Eq. (1) based at the point  $w_i \in M$ , where  $M$  is the compact solution manifold with a local  $R^n$  structure, as follows:

$$\frac{dy}{ds} = D_{w_i}V(w, s)y = A(s)y + D_{w_i}Q(w, s)y \quad (3)$$

where “ $D$ ” stands for the vector derivative or Jacobian operator. It is observed that  $D_{w_i}V(w, s)$  is the tangent map based at the point  $w_i \in M$  with the usual Riemann structure of inner product norms (see [26]). On the other hand, since Eq. (2) is conditionally linear, the corresponding variational equation at  $w_i$  is obtainable by using the same vector field without the external forcing term. Thus the variational equation is

$$\frac{d\bar{y}}{ds} = D_{w_i}\bar{V}(\bar{w}, s) = A(s_i)\bar{y} + B(w_i, s_i)\bar{y}. \quad (4)$$

Here  $D_{w_i}\bar{V}$  is the tangent space of the linearized system at  $w_i \in M$ . In fact, the unknown solution  $w_i$ , being a point of intersection of the flows  $w(s)$  and  $\bar{w}^{(i)}(s)$ , belongs to  $M \cap \bar{M}^{(i)}$ , where  $\bar{M}^{(i)}$  is the solution manifold generated by  $\bar{w}^{(i)}(s)$ . This is a condition that can be viewed as a constraint on the nonlinear response. Consequently, a suitable constraint equation needs to be arrived at. Towards this, the solution of the linear system of Eq. (2) is explicitly written down as

$$\bar{w}^{(i)}(s) = \psi(w_i, s, s_{i-1}) \left( w_{i-1} + \int_{s_{i-1}}^{s \leq s_i} \psi^{-1}(w_i, t, s_{i-1}) f(s) dt \right) \quad (5)$$

where  $\psi(w_i, s, s_{i-1})$ , is the (locally constructed) fundamental solution matrix. While the first term on the right-hand side of Eq. (5) represents the complementary solution, the second one, involving integration, stands for the particular integral due to the external forcing function,  $f(s)$ . Equation (5) may now be differentiated once to yield

$$\frac{d\bar{w}^{(i)}(s)}{ds} = \frac{d\psi}{ds} w_{i-1} + \frac{d\psi}{ds} \int_{s_{i-1}}^s \psi^{-1}(\eta) f(\eta) d\eta + I f(s) \quad (6)$$

where  $I$  is the identity map in dimension  $n$ . The constraint condition, i.e.,  $w_i \in M \cap \bar{M}^{(i)}$ , may be considered equivalent to the identity

$$\bar{w}_i = w_i. \quad (7)$$

In order to satisfy the above identity,  $\bar{w}(s)$  and  $d\bar{w}(s)/ds$  from Eqs. (5) and (6) are substituted in Eq. (1) for  $\bar{w}(s)$  and  $d\bar{w}(s)/ds$ , respectively, at  $s = s_i$ . This leads to the following  $n$  algebraic nonlinear equations for the unknown vector  $w_i$ .

$$\mu_i(w_i, s_i, s_{i-1}) = 0 \quad (8a)$$

where the vector nonlinear function  $\mu_i$  is given by

$$\mu_i = \frac{d\psi(w_i, s_i, s_{i-1})}{ds} w_{i-1} + \frac{d\psi(w_i, s_i, s_{i-1})}{ds} \int_{s_{i-1}}^{s_i} \psi^{-1}(w_i, \eta, s_{i-1}) f(\eta) d\eta - A(s_i) \bar{w}(s_i) - Q(w_i, s_i). \quad (8b)$$

Similarly one can construct a system of  $(r-1)n$  algebraic equations at intermediate nodal stations  $i=1, 2, \dots, r-1$ . All these algebraic equations, however, contain the unknown boundary condition vector  ${}^l w_0$ ,  $l=l_1, l_2, \dots, l_{n-m}$  at  $s=s_0$ . Hence  $n-m$  additional equations are necessary to solve these set of nonlinear equations. These additional equations are readily obtainable by constructing the  $r$ th LTL system at  $s=s_r$  followed by using the specified boundary conditions at the right end of the boundary, i.e.,  ${}^k w_r$ ,  $k=k_1, k_2, \dots, k_{n-m}$ . Moreover, this  $r$ th LTL system may also be used to construct  $m$  nonlinear algebraic equations to account for  $m$  unknown elements,  ${}^p w_r^{(r)}$ ,  $p=p_1, p_2, \dots, p_m$ , at  $s=s_r$ . Solving all these  $(r-1)n + n - m + m = nr$  coupled and nonlinear algebraic equations together, one can, in general, obtain all the unknown elements of the state vector at all the stations simultaneously. Depending on the specific form of the vector field associated with the given boundary value problem, the number of such coupled algebraic equations to be solved within the LTL framework may, however, be suitably reduced. A few such cases have been discussed in Section 4. Attention is now focused on the analytical expressions for the two tangent maps,  $T_{w_i}V$  and  $T_{w_i}\bar{V}$  given, respectively, by Eqs. (3) and (4). It is clear that  $T_{w_i}V$  is transversal to  $T_{w_i}\bar{V}$  for almost all  $w_i$  except at a countable number of points satisfying the following  $n$  nonlinear coupled algebraic equations:

$$D_w Q(w_i, s_i) = B(w_i, s_i) \quad (9a)$$

where only real roots,  $w_i$ , of the above equation are of interest. Consider, as a simple example, a one-dimensional ( $n=1$ ) case with  $Q(w, s) = w^3$  and  $B(w_i, s_i) = w_i^2$ . In this case, Eq. (9a) reduces to

$$2w_i^2 = 0. \quad (9b)$$

In other words,  $w_i=0$  is the only point on the real line,  $R$ , where the original and reconstructed tangent spaces,  $T_{w_i}V$  and  $T_{w_i}\bar{V}$ , fail to be transversal. Thus if the solutions,  $w_i$ , of Eqs. (8) and (9) happen to coincide, then the conditionally linear flow  $\bar{\phi}_s$  does not transversally intersect the nonlinear flow  $\phi_s$  at  $s_i$ , thereby rendering the LTL method ineffective. *An obvious way out is to change  $s_i$  such that  $w_i$  changes away from its singular value.*

In the procedure just presented the values of  $\bar{w}(s)$  and  $d\bar{w}(s)/ds$  from Eqs. (5) and (6) are substituted in Eq. (1) for  $w(s)$  and  $d\bar{w}(s)/ds$  to explicitly satisfy the identity  $\bar{w}(s_i) = w(s_i)$  at  $s_i$ . However, one can avoid this-step provided that the conditionally constant coefficient matrix  $B(w_i, s_i)$  the vector field of LTL Eq. (2) is chosen such that

$$[B(w_i, s_i)]\{w_i\} = \{Q(w_i, s_i)\}. \quad (10)$$

In case such a decomposition, as in Eq. (10), is possible, the LTL Eq. (2) becomes identical in form with that of Eq. (1) at  $s = s_i$ . In this case, one straightaway obtains the transcendental and nonlinear algebraic equations in the unknown vector  $w_i$  via the linearized solution (5) as

$$w_i = \bar{w}_i = \psi(w_i, s_i, s_{i-1}) \left( w_{i-1} + \int_{s_{i-1}}^{s_i} \psi^{-1}(w_i, \eta, s_{i-1}) f(\eta) d\eta \right). \quad (11)$$

Similarly one can construct the nonlinear algebraic equations at all the  $(r-1)$  nodal stations (i.e., leaving out the two boundary nodes). As in the previous case, all these nonlinear algebraic equations are coupled with the unknown boundary vector  ${}^l w_0$ ,  $l=l_1, l_2, \dots, l_{n-m}$  at the left boundary. The additional  $(n-m)$  equations, to account for the unknown vector at the left boundary, are made available, as before, by constructing the  $r$ th LTL system (in terms of  $\bar{w}^{(r)}$ ) and using the known boundary condition vector  ${}^k w_r$ ,  $k=k_1, k_2, \dots, k_{n-m}$  at the right boundary. As before, this  $r$ th LTL system may also be used to further construct  $m$  more algebraic equations to take into account the unknown vector  ${}^p w_r^{(r)}$ ,  $p=p_1, p_2, \dots, p_m$ , at the right boundary.

### 3 Higher Level Locally Transversal Linearization (LTL) Systems

It is of interest to derive other forms of LTL schemes with capabilities to remain close to the original path as followed by the nonlinear system with the given boundary conditions, provided that the chosen step-size is sufficiently small. The basic form of LTL, presented in the previous section, only ensures transversal intersections at the two end-points of a chosen interval and not a consistent closeness of the paths in between. It is thus obvious that there is a scope to improve the basic LTL procedure, henceforth called the zeroth-level LTL procedure, to improve upon its path sensitivity. This may be done by suitably differentiating the zeroth-level LTL-based ordinary differential equations thereby constructing higher level LTL systems. Thus consider the case of a nonlinear boundary value problem where the associated zeroth level  $i$ th LTL system is given by Eq. (2). The  $i$ th higher level LTL equation ( $2n$ -dimensional) may be constructed by differentiating the original boundary value problem, given by Eq. (1), once with respect to  $s$  and suitably freezing parts of the resulting vector field at  $s = s_i$ . This finally leads to

$$\frac{d^2 \bar{w}^{(i)}}{ds^2} = A(s_i) \frac{d\bar{w}^{(i)}}{ds} + \left[ \frac{dA(s)}{ds} + \frac{\partial B(w_i, s)}{\partial s} \right] \bigg|_{s=s_i} + \Lambda \left( \frac{dw_i}{ds} \right) \bar{w}^{(i)} + \frac{df(s)}{ds} \quad (12)$$

where the  $n \times n$  matrix  $\Lambda(dw_i/ds)$  has the following typical elements:

$$\Lambda_{kl} = \sum_{q=1}^n \frac{\partial B_{kl}}{\partial (q w)} \left( \frac{d^q w}{ds} \right) \quad (13)$$

subject to the following  $n$  boundary conditions and another  $n$  intermediate constraints:

$$\begin{aligned} q \bar{w}^{(i)}(s_{i-1}) &\triangleq q \bar{w}_{i-1}^{(i-1)} = q w_{i-1}, \quad q = 1, 2, \dots, n \\ \frac{d^q \bar{w}_{i-1}^{(i)}}{ds} &= \frac{d^q w_{i-1}}{ds} = A(s_{i-1}) q w_{i-1} + Q(q w_{i-1}, s_{i-1}) \\ &+ f(s_{i-1}), \quad q = 1, 2, \dots, n. \end{aligned} \quad (14)$$

While the first  $n$  initial (boundary) conditions (14) are the same as those used for the zeroth-level LTL system, the second constitutes a set of  $n$  constraint conditions at  $s = s_{i-1}$ . It may also be noted that these constraint conditions involving derivatives at  $s = s_{i-1}$  are implicitly defined in terms of the unknown vector  $w_{i-1}$ , which is the desired state vector at  $s = s_{i-1}$ . Satisfaction of these constraint conditions therefore implies that in the first-level LTL procedure, the  $i$ th LTL solution path not only emerges from  $s = s_{i-1}$  with the same values of the state vector  $w_{i-1}$  as followed by the original solution path, but also with the same derivatives  $dw_{i-1}/ds$ . The same equality is also ensured at  $s = s_i$ , wherein a transversal intersection now occurs between the  $2n$ -dimensional linearized and nonlinear solution manifolds (obtained by once differentiating the given  $n$ -dimensional system of ordinary differential equations). This obviously enhances path sensitivity of the first-level LTL system, which was the objective of constructing such systems. Proceeding in the same way, even higher levels of LTL systems may readily be conceived.

#### 4 Illustrative Examples

In this section, a few beam elastica problems are studied using the new LTL methods. In addition, nonlinear beam buckling problems in the presence of initial imperfections are also studied. In the first problem, deflections and slopes of a tip-loaded cantilever beam are calculated. When a concentrated load “ $Q$ ” is applied at the tip of the cantilever, it is possible to obtain exact solutions in the form elliptic integrals [27]. The cantilever beam along with the coordinate axes is shown in Fig. 1(a). The distance “ $s$ ” is measured along the length of the beam and “ $x$ ” denotes the projection of “ $s$ ” along the undeformed longitudinal axis of the beam. The exact equilibrium equation for the problem, written in terms of the slope,  $\theta$ , of the deflection curve, is

$$EI \frac{d\theta}{ds} = (L-x)Q \quad (15)$$

where  $d\theta/ds$  denotes the exact curvature of the beam. Differentiating the equation with respect to “ $s$ ” and noting that  $dx/ds = \cos \theta$ , one gets

$$\frac{d^2 \theta}{ds^2} + \lambda^2 \cos \theta = 0 \quad (16)$$

where  $\lambda^2 = P/EI$ . Equation (16) is the governing nonlinear differential equation for the problem and the associated boundary conditions are

$$\theta(0) = \theta_0 = 0, \quad \frac{d\theta(L)}{ds} = \theta'_r = 0. \quad (17)$$

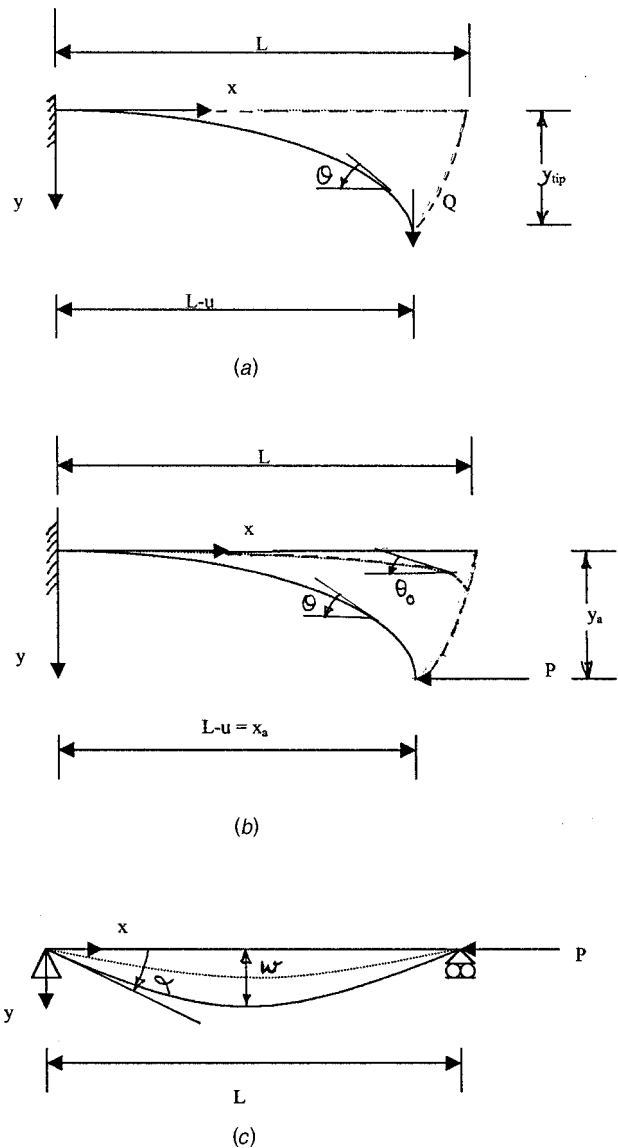


Fig. 1 (a) Tip-loaded cantilever beam showing the coordinate axes (b) Cantilever beam with initial imperfection (c) Simply supported beam showing coordinate axes and initial curvature

For solving the boundary value problem (16), the following zeroth-level LTL equation over the  $i$ th subdomain  $D_i$  and with initial conditions  $\{\bar{\theta}_{i-1}^{(i-1)}, d\bar{\theta}_{i-1}^{(i-1)}/ds = \bar{\theta}'_{i-1}^{(i-1)}\}$  is employed:

$$\frac{d^2 \bar{\theta}^{(i)}}{ds^2} = \beta_i \quad (18)$$

where  $\theta_i = \theta(s_i)$ . The complete solution of the conditionally linear ordinary differential equation (18) is easily written as

$$\bar{\theta}^{(i)}(s) = \frac{\beta_i s^2}{2} + C_1 s + C_2. \quad (19)$$

The arbitrary constants of integration  $C_1$  and  $C_2$  may be found via initial conditions:

$$\begin{aligned} \bar{\theta}_{i-1} &= \theta_{i-1} \quad \text{if } i > 1 \\ &= 0 \quad \text{if } i = 1 \end{aligned} \quad (20)$$

$$\bar{\theta}'_{i-1} = \theta'_{i-1}. \quad (21)$$

After obtaining the expression for  $\bar{\theta}(s)$ , it is simple to derive analytical expressions for  $\bar{\theta}''(s)$  in terms of the unknown solution parameters  $\beta_i$ . Substituting  $\bar{\theta}^{(i)}(s)$  and  $\bar{\theta}^{(i)''}(s)$  in the original Eq. (16) at  $s=s_i$ , a nonlinear algebraic equation in  $\beta_i$  is obtained. In a similar way one may continue obtaining nonlinear equations for the conditionally constant parameters  $\beta_j$  ( $j=1,2,\dots,r$ ) at all the intersection points. This results in “ $r$ ” nonlinear algebraic equations in  $r+1$  unknown parameters  $\beta_j$  ( $j=1,2,\dots,r$ ) and  $\theta'_0$ . However, it may be observed that for the present problem  $\theta'_r=0$  is known at the right end i.e.,  $s=L$ . Thus, in the conditional initial value problem approach, the unknowns  $\theta'_0$  is to be so determined as to result in a constrained dynamical trajectory with the curvature vanishing at  $s=L$ . This constraint condition leads to the required additional equation via the analytical expression available for  $\bar{\theta}^{(r)}(s)$ .

Solving the above nonlinear algebraic equation one obtains the value of  $\theta$  at all the chosen points along the  $s$ -axis. The Cartesian coordinates of the deflected curve at any point  $i+1$  is evaluated from the relation

$$y_{i+1}=y_i+\int_{s_i}^{s_{i+1}} \sin \theta ds \quad (22a)$$

$$x_{i+1}=x_i+\int_{s_i}^{s_{i+1}} \cos \theta ds. \quad (22b)$$

The above integrations are accomplished via the two-point Gaussian quadrature rule.

The above problem may also be solved using a first-level LTL method to obtain even better results. Here the linearized equation over the  $i$ th interval takes the form

$$\frac{d^3 \bar{\theta}^{(i)}}{ds^3} = \gamma_i. \quad (23)$$

Instead of substituting  $\bar{\theta}^{(i)}$ , and  $\bar{\theta}^{(i)''}$  in the original nonlinear equation for obtaining the nonlinear algebraic equations, one can alternatively set up the required nonlinear algebraic equation (without a need of substitution) by choosing  $\beta_i = \lambda^2 \cos(\theta_i)$ . In such a case, it is clear that the vector fields of the original nonlinear differential and linearized equations become identical at  $s=s_i$ . Hence one can set up the algebraic equation at  $s=s_i$  using

$$\bar{\theta}(s=s_i) = \theta(s=s_i). \quad (24)$$

Thus solutions of Eqs. (16) and (18) would instantaneously match at  $s=s_i$  if one could find a real root  $\theta_i$  of the following transcendental equation:

$$\theta_i - \bar{\theta}(s_i, \beta_i(\theta_i)) = 0 \quad \text{with} \quad \beta_i = \lambda^2 \cos(\theta_i). \quad (25)$$

The above procedure may easily be extended to construct higher order LTL systems. For instance, the solution parameter  $\gamma_i$  in the first-level LTL Eq. (23) may be chosen as  $\gamma_i = -\lambda^2 \sin(\theta_i)\theta'_i$ . The boundary conditions are

$$\begin{aligned} \bar{\theta}_{i-1} &= \theta_{i-1} \quad \text{if } i > 1 \\ &= 0 \quad \text{if } i = 0 \end{aligned} \quad (26a)$$

$$\begin{aligned} \frac{d^2 \bar{\theta}_{i-1}}{ds^2} &= -\lambda^2 \cos(\theta_{i-1}) \quad \text{if } i > 1 \\ &= -\lambda^2 \quad \text{if } i = 0. \end{aligned} \quad (26b)$$

Finally, the following pair of nonlinear algebraic equations in terms of  $\theta_i$  and  $d\theta_i/ds$  may be arrived at

$$\theta_i - \bar{\theta}\left(s_i, \gamma_i\left(\theta_i, \frac{d\theta_i}{ds}\right)\right) = 0$$

$$\frac{d\theta_i}{ds} - \frac{d\bar{\theta}}{ds}\left(s_i, \gamma_i\left(\theta_i, \frac{d\theta_i}{ds}\right)\right) = 0. \quad (27)$$

The second illustrative example is on the nonlinear buckling of a cantilever beam. The classical elastica problem of cantilever beam buckling is based on the exact expression for curvature and solutions of the resulting nonlinear differential equations may be readily obtained in terms of elliptic integrals ([28]). In this case, the governing nonlinear differential equation may be written as

$$\frac{d^2 \theta}{ds^2} + \frac{P}{EI} \sin \theta = 0. \quad (28)$$

The associated boundary conditions are the same as in the previous example. For the present problem, the zeroth-level LTL equation over the  $i$ th subdomain  $D_i$  and with associated initial conditions  $\{\bar{\theta}_{i-1}^{(i-1)}, d\bar{\theta}_{i-1}^{(i-1)}/ds = \bar{\theta}'_{i-1}^{(i-1)}\}$  may be written in many ways. Here two different ways of linearizations are tried. In the first method, the following form for the linearized ordinary differential equation is tried out:

$$\frac{d^2 \bar{\theta}}{ds^2} = -\frac{P}{EI} \sin \bar{\theta}_i = \beta_i. \quad (29)$$

One could also linearize Eq. (28) as

$$\frac{d^2 \bar{\theta}}{ds^2} + \left(\frac{P}{EI} \frac{\sin \theta_i}{\theta_i}\right) \bar{\theta} = 0 \quad (30a)$$

or

$$\frac{d^2 \bar{\theta}}{ds^2} + \lambda^2 \bar{\theta} = 0 \quad (30b)$$

where  $\lambda^2 = ((P/EI)(\sin \theta_i/\theta_i))$  and  $\theta_i = \theta(s=s_i)$ . In both these cases, substitution of the linearized solution into the nonlinear ordinary differential equation is avoided. In the first method, solution of the conditionally linear Eq. (29) is the same as given in Eq. (19). In the second method, the two eigenvalues needed to construct the solution of Eq. (30b) are given by

$$\lambda_{1,2} = \pm \sqrt{\frac{P}{EI} \frac{\sin \theta_i}{\theta_i}}. \quad (31)$$

Depending on the value of the argument within the square root two possible cases arise. In the first case, let  $\lambda^2 > 0$ . This implies that the eigenvalues are complex conjugate and are given by

$$\lambda_{1,2} = \pm j\beta, \quad \beta = \sqrt{\frac{P}{EI} \frac{\sin \theta_i}{\theta_i}}, \quad j = \sqrt{-1}. \quad (32)$$

The solutions of the conditionally linear Eq. (30b) may be written as

$$\bar{\theta}(s) = C_1 \sin \lambda s + C_2 \cos \lambda s \quad (33)$$

where the constants of integration are evaluated from the boundary conditions. In the second case one has  $\lambda^2 < 0$ . Thus the eigenvalues are real and are given by the Eq. (31). The solution  $\bar{\theta}(s)$  takes the form

$$\bar{\theta}(s) = C_1 \exp(\beta s) + C_2 \exp(-\beta s). \quad (34)$$

The rest of the steps for implementing the LTL method are precisely the same as already outlined previously. As a third problem, buckling of a cantilever beam with initial curvature is considered. The equilibrium equation for the initially curved cantilever column may be written as



$$EI \left( \frac{d^2 \theta}{ds^2} - \frac{d^2 \theta_0}{ds^2} \right) + P \sin \theta = 0 \quad (35a)$$

or

$$\frac{d^2 \theta}{ds^2} + \frac{P}{EI} \sin \theta = \frac{d^2 \theta_0}{ds^2}. \quad (35b)$$

The notations are shown in Fig. 1(b). The tangent angles of the deformed and undeformed column are  $\theta$  and  $\theta_0$ , respectively. The initial shape of the column is assumed to be of the form

$$\theta_0 = A_0 \left( 1 - \cos \left( \frac{\pi s}{2L} \right) \right) \quad (36)$$

where  $A_0$  is the amplitude of tangent angle of the undeformed column. The linearized equation over the interval  $D_i$  may be written as

$$\frac{d^2 \bar{\theta}}{ds^2} + \left( \frac{P}{EI} \frac{\sin \theta_i}{\theta_i} \right) \bar{\theta} = \theta''_0 \quad (37a)$$

or

$$\frac{d^2 \bar{\theta}}{ds^2} + \lambda^2 \bar{\theta} = A_0 \left( \frac{\pi s}{2L} \right)^2 \cos \left( \frac{\pi s}{2L} \right) \quad (37b)$$

The functional form of the complete solution for the above ordinary differential equation may be written as

$$\lambda^2 > 0, \quad \lambda_{1,2} = \pm j\beta, \quad \beta = \sqrt{\frac{P}{EI} \frac{\sin \theta_i}{\theta_i}}, \quad j = \sqrt{-1},$$

$$\bar{\theta}(s) = A \sin(\beta s) + B \cos(\beta s) + p(s) \quad (38)$$

and when

$$\lambda^2 < 0, \quad \lambda_{1,2} = \pm \beta, \quad \beta = \sqrt{\frac{P}{EI} \frac{\sin \theta_i}{\theta_i}}$$

$$\bar{\theta}(s) = C_1 \exp(\beta s) + C_2 \exp(-\beta s) + p(s). \quad (39)$$

In the above expression  $p(s)$  denotes the particular solution and rest of the terms on the right-hand side constitutes the complementary solution of Eq. (37b). The particular solution is given by

$$p(s) = \left( \frac{A_0 \left( \frac{\pi}{2L} \right)^2}{\left( -\frac{\pi}{2L} \right)^2 + \lambda^2} \right) \cos \left( \frac{\pi s}{2L} \right). \quad (40)$$

The simply supported beam-buckling problem is solved as the next example to further demonstrate the application of the proposed LTL method. The governing equilibrium equations and the associated boundary conditions are

$$\frac{d^2 \theta}{ds^2} + \frac{P}{EI} \sin \theta = 0 \quad (41a)$$

$$\theta'(0) = 0 \quad \text{and} \quad \theta'(L) = 0. \quad (41b)$$

The above nonlinear differential equation is solved exactly in the same manner as explained in the case of cantilever buckling problem except that the boundary conditions are different in the present case. The last illustrative example considered here is that of the nonlinear buckling of an imperfect simply supported column. The initial shape is expressed in the present case as

$$\theta_0 = A_0 \cos \left( \frac{\pi s}{L} \right). \quad (42)$$

The rest of the procedure is identical to the one already explained for the case of cantilever beam problem with initial curvature.

## 5 Numerical Results

The remarkable accuracy and potential of the proposed method to solve nonlinear boundary value problems is demonstrated in this section via a few numerical results corresponding to the well-known nonlinear beam problems as outlined in the previous section. In Table 1, deflections and slopes at the free end of a tip-loaded cantilever beam are given. It is clear from the table that the results compare well with the exact result ([27]). It may be observed here that the LTL-2 (first level) method gives better results than the LTL-1 (zeroth-level) method. In Fig. 2, the deflected shape of the cantilever beam is shown for various values of concentrated load. The difference between the calculated values as obtained from LTL-1 and LTL-2 and the exact results can hardly be represented graphically. They indeed compare very well. In Fig. 3, the tip deflection of a cantilever beam under compressive in-plane load as obtained from LTL method is compared with the exact result ([28]). The results obtained via LTL-1 and LTL-2 methods are shown in Table 2. In Table 3 the results obtained by using the LTL-1 scheme and different forms of the LTL-based ordinary differential equations (see Eqs. (29) and (30)) are presented along with the exact results. Between the two LTL-1 methods it is seen that LTL-1 (2) works better. This is obvious as it takes into account the frequency information of the system while linearization. It may be observed from the above tables and figures that LTL procedure works accurately even for very high nonlinearity. Since the error between the LTL-1 method and the exact results are very small, all further calculations are carried out with LTL-1 method alone. All the further results are obtained by setting

**Table 1 Results for a tip-loaded cantilever beam**

$\frac{QL^2}{EI}$	Zeroth-Order LTL 80 segments			First-Order LTL 40 segments			Exact		
	$y_{\text{tip}}/L$	$\theta$ (radians)	$\frac{(L-u)}{L}$	$y_{\text{tip}}/L$	$\theta$ (radians)	$\frac{(L-u)}{L}$	$y_{\text{tip}}/L$	$\theta$ (radians)	$\frac{(L-u)}{L}$
1.0	0.3015	0.4611	0.9436	0.3017	0.4614	0.9436	0.3017	0.4613	0.9436
2.0	0.4928	0.7807	0.8398	0.4934	0.7818	0.8394	0.4935	0.7817	0.8394
3.0	0.6022	0.9843	0.7466	0.6031	0.9861	0.7456	0.6033	0.9860	0.7456
4.0	0.6687	1.1194	0.6725	0.6698	1.1211	0.6710	0.6696	1.1212	0.6711
5.0	0.7124	1.2132	0.6142	0.7136	1.2153	0.6123	0.7138	1.2154	0.6124
6.0	0.7431	1.2813	0.5676	0.7444	1.2840	0.5654	0.7446	1.2837	0.5654
7.0	0.7659	1.3324	0.5295	0.7672	1.3350	0.5270	0.7674	1.3350	0.5271
8.0	0.7835	1.3721	0.4978	0.7848	1.3749	0.4951	0.7850	1.3744	0.4952
9.0	0.7976	1.4032	0.4651	0.7989	1.4058	0.4681	0.7991	1.4055	0.4682
10.0	0.8091	1.4279	0.4479	0.8104	1.4302	0.4449	0.8106	1.4303	0.4450

$y_{\text{tip}}$  = deflection at the tip,  $\theta$  = slope at the tip



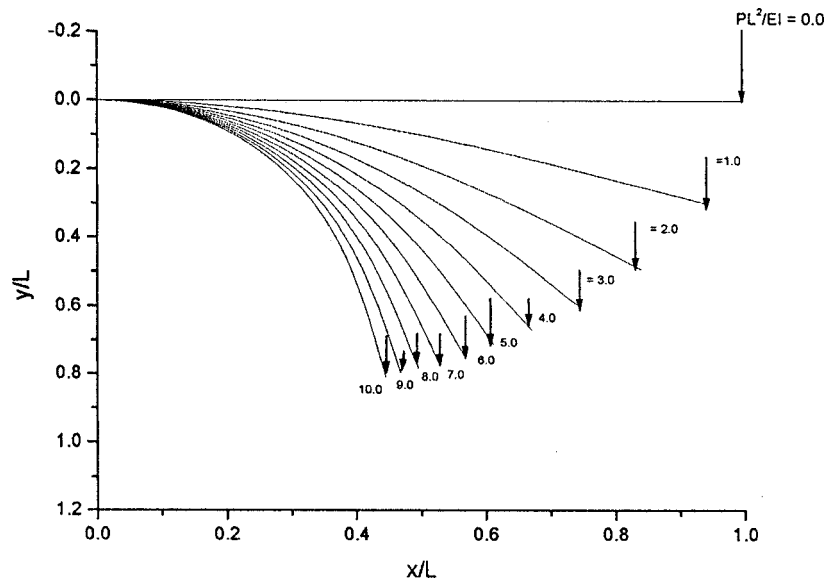


Fig. 2 Load displacement curves for a tip-loaded cantilever beam

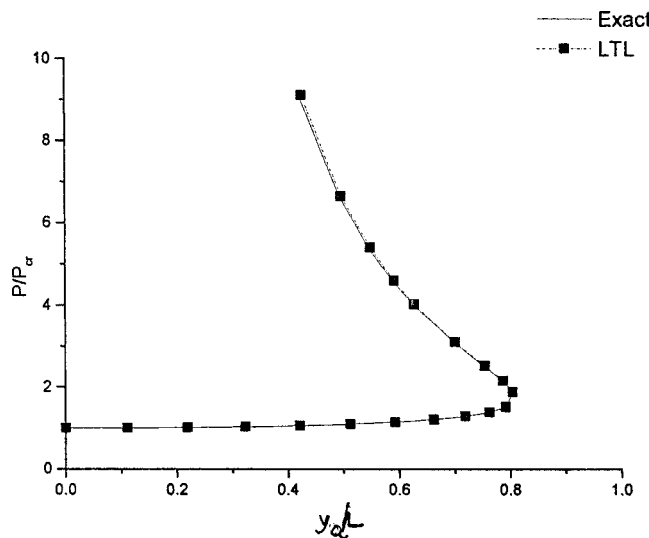


Fig. 3 Cantilever buckling:  $P/P_{cr}$  versus normalized tip deflection

up the nonlinear algebraic equations without resubstituting  $\bar{\theta}(s)$  and  $\bar{\theta}''(s)$  in the original nonlinear equations. In Fig. 4, the shape of the cantilever column under different loading intensities are shown. All these shapes have been computed using the LTL technique. In the next figure (Fig. 5) the variation of axial strain is plotted against  $P/P_{cr}$  for the case of imperfect cantilever buckling for different amplitudes of imperfections. Next (Fig. 6), the nonlinear buckling problem of a simply supported beam is solved using the LTL method. In Fig. 7, the  $P/P_{cr}$  versus central deflection plot is compared with that of the exact solution. From the figure it may be observed that the LTL method works very well for the nonlinear boundary value problems. The shape of the simply supported column under different intensities of load is shown in Fig. 8. The LTL method is promising as it can predict any shape without difficulty. The results are also shown in Table 4 along with the exact results. In Fig. 9 the axial strains are plotted against the  $P/P_{cr}$  for the case of imperfect simply supported beam. As the imperfection amplitude increases, the column deflects more for the same given load. The  $P/P_{cr}$  versus central deflection is shown in Fig. 10.

Table 2 Comparison of zeroth and first-order LTL methods in the case of nonlinear buckling of cantilever beam

$\frac{P}{P_{cr}}$	Zeroth-Order LTL (LTL-1)			First-Order LTL (LTL-2)			Exact		
	$\theta$ (radian)	$x_a$	$y_a$	$\theta$ (radian)	$x_a$	$y_a$	$\theta$ (radian)	$x_a$	$y_a$
1.0	0.0	1.0	0.0	0.0	1.0	0.0	0.0	1.0	0.0
1.0038	0.1846	0.9915	0.1172	0.1741	0.9923	0.1105	0.1745	0.9924	0.1108
1.0153	0.3918	0.9619	0.2456	0.3485	0.9698	0.2191	0.3491	0.9698	0.2194
1.0351	0.5522	0.9249	0.3407	0.5233	0.9325	0.3237	0.5235	0.9325	0.3239
1.0636	0.7179	0.8744	0.4332	0.6970	0.8816	0.4216	0.6981	0.8812	0.4222
1.1021	0.8885	0.8104	0.5206	0.8728	0.8170	0.5126	0.8727	0.8170	0.5125
1.1518	1.0586	0.7351	0.5988	1.0473	0.7410	0.5933	1.0472	0.7410	0.5932
1.2147	1.2307	0.6495	0.6663	1.2218	0.6547	0.6626	1.2217	0.6546	0.6626
1.2939	1.4028	0.5550	0.7219	1.3962	0.5594	0.7195	1.3963	0.5593	0.7195
1.3932	1.5756	0.4534	0.7642	1.5710	0.4570	0.7627	1.5710	0.4569	0.7627
1.5184	1.7489	0.3462	0.7923	1.7452	0.3490	0.7915	1.7453	0.3489	0.7915

$\theta$ =slope at the free end of the beam

$x_a = (L-u)/L$ , projected length of the beam (see Fig. 1(b))

$y_a$ =Nondimensionalized deflection at the tip of the cantilever beam

$$P_{cr} = \frac{\pi^2 EI}{4L^2}$$

**Table 3 Nonlinear buckling of cantilever beam: comparison of results obtained from two different ways of linearizations**

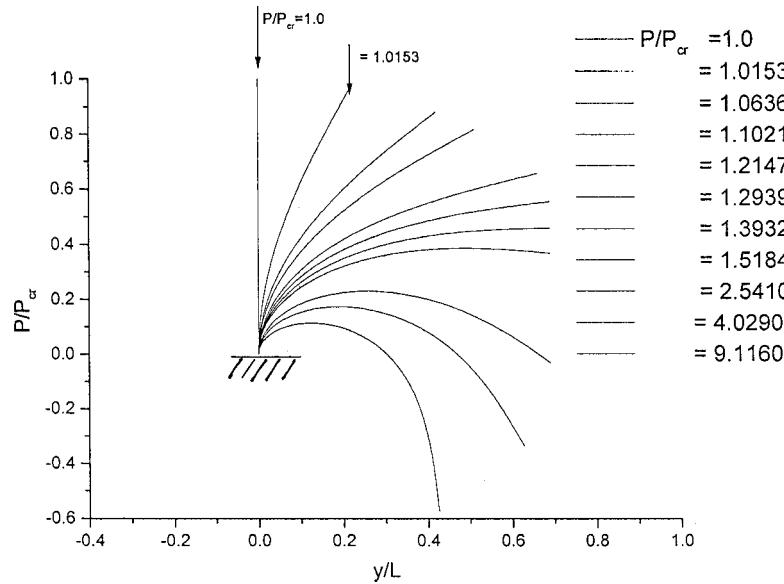
$\frac{P}{P_{cr}}$	LTL-1(1) Eq. (29)			LTL-1(2) Eq. (30)			Exact		
	$\theta$ (radian)	$x_a$	$y_a$	$\theta$ (radian)	$x_a$	$y_a$	$\theta$ (radian)	$x_a$	$y_a$
1.0153	0.3918	0.9619	0.2456	0.3477	0.9700	0.2186	0.3491	0.9698	0.2194
1.0351	0.5522	0.9249	0.3407	0.5220	0.9328	0.3230	0.5235	0.9325	0.3239
1.0636	0.7179	0.8744	0.4332	0.6952	0.8822	0.4206	0.6981	0.8812	0.4222
1.1021	0.8885	0.8104	0.5206	0.8705	0.8179	0.5115	0.8727	0.8170	0.5125
1.1518	1.0586	0.7351	0.5988	1.0454	0.7423	0.5920	1.0472	0.7410	0.5932
1.2147	1.2307	0.6495	0.6663	1.2180	0.6564	0.6614	1.2217	0.6546	0.6626
1.2939	1.4028	0.5550	0.7219	1.3932	0.5616	0.7184	1.3963	0.5593	0.7195
1.3932	1.5756	0.4534	0.7642	1.5674	0.4595	0.7618	1.5710	0.4569	0.7627
1.5184	1.7489	0.3462	0.7923	1.7413	0.3520	0.7909	1.7453	0.3489	0.7915
2.541	2.4314	-0.0982	0.7535	2.4382	-0.1026	0.7519	2.4435	-0.1070	0.7500
4.029	2.7871	-0.3362	0.6274	2.7874	-0.3361	0.6274	2.7925	-0.3400	0.6250
9.116	3.0685	-0.5726	0.4249	3.0693	-0.5732	0.4254	3.0718	-0.5770	0.4210

$\theta$ =slope at the free end of the beam

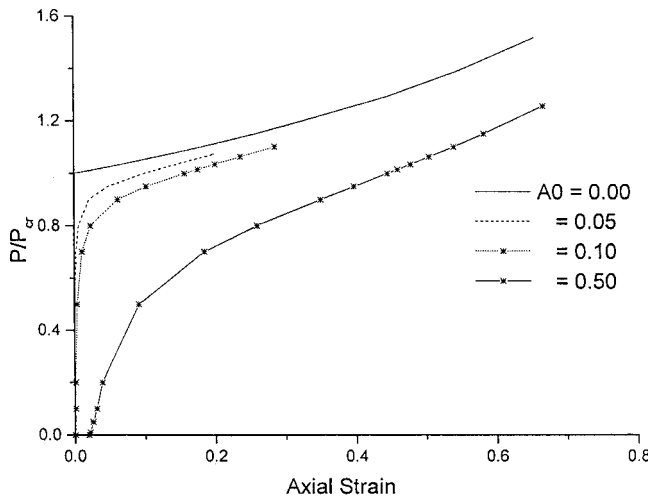
$x_a=(L-u)/L$ , projected length of the beam (see Fig. 1(b))

$y_a$ =Nondimensional deflection at the tip of the cantilever beam

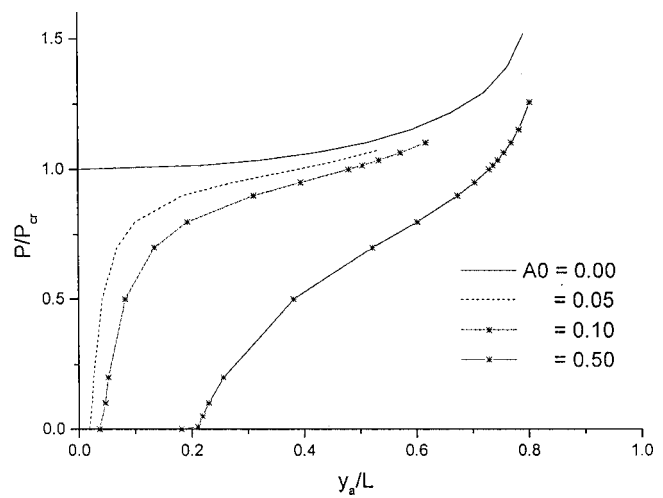
$$P_{cr} = \frac{\pi^2 EI}{4L^2}$$



**Fig. 4 Deflected shape of the cantilever column under different levels of axial load**



**Fig. 5 Imperfect cantilever buckling:  $P/P_{cr}$  versus axial strain for different amplitudes of imperfection**



**Fig. 6 Imperfect cantilever buckling:  $P/P_{cr}$  versus tip deflection for different amplitudes of imperfection**

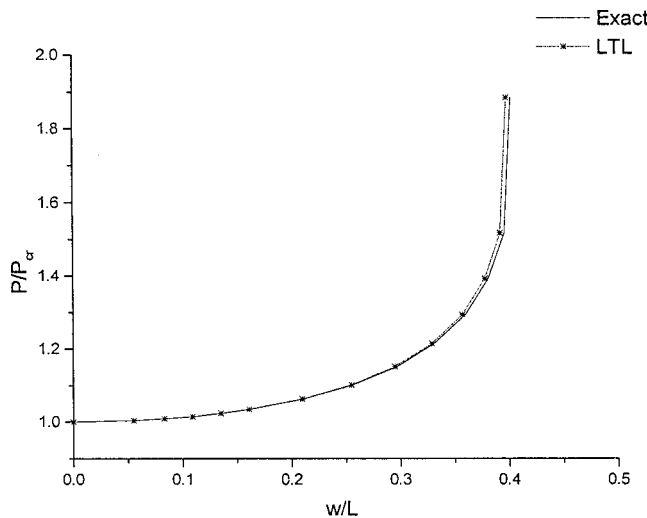


Fig. 7 Nonlinear buckling of simply supported beam:  $P/P_{cr}$  versus central deflection plot

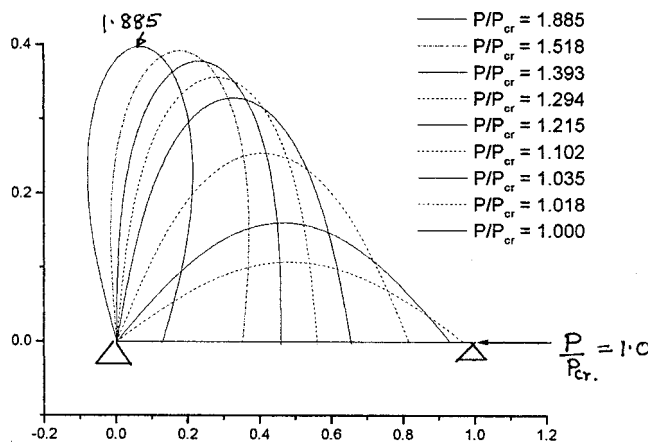


Fig. 8 Deflected shape of the simply supported column under different levels of axial loads as predicted by the LTL

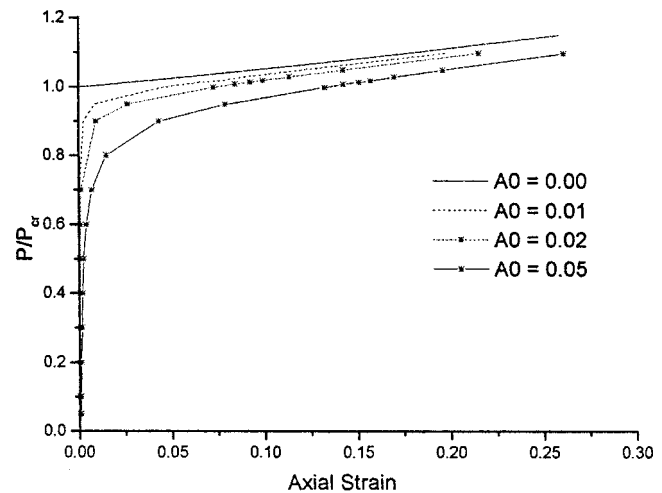


Fig. 9 Imperfect simply supported column buckling:  $P/P_{cr}$  versus axial strain for different amplitudes of imperfection

Initially a convergence study was carried out to decide on the number of segments,  $r$ , required in the LTL-1 procedure to accurately predict the deflection profile of the column and the post-buckling load. Details of these results are not shown for the sake of brevity. It has, however, been observed from the study that the converged results are obtained using 40 segments in most cases. However, when the beam shapes are severely distorted under very high curvature nonlinearity, 40 divisions are not enough to obtain converged results. In such cases, more number of divisions (about 120) are necessary to obtain the converged results. In the present work, 120 divisions are taken in all the numerical study.

As mentioned earlier, one has to solve all the nonlinear algebraic equations simultaneously in order to obtain the equilibrium path of the structure. In the present study, a globally convergent nonlinear equation solver based on line searches and backtracking along the Newton directions (see [29]) has been adopted for solving the system of the nonlinear algebraic equations.

## 6 Discussion and Conclusions

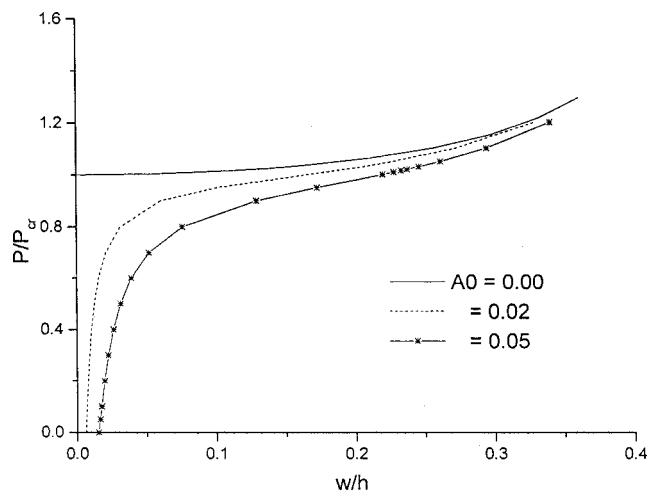
A novel and remarkably accurate local linearization methodology, called the locally transversal linearization (LTL), is proposed in the present study for solving a class of two-point nonlinear boundary value problems of relevance in structural mechanics. Given a system of nonlinear ordinary differential equations, the LTL method derives a set of conditionally linear ordinary differ-

Table 4 Support slopes and center deflections of a simply supported beam under axial compressive load

$\frac{P}{P_{cr}}$	Zeroth-Order LTL Results		First-Order LTL Results		Exact Results	
	$\alpha$ (radian)	$\frac{w}{L}$	$\alpha$ (radian)	$\frac{w}{L}$	$\alpha$ (radian)	$\frac{w}{L}$
1.004	0.179	0.056	0.178	0.055	0.175	0.055
1.024	0.435	0.134	0.435	0.135	0.436	0.136
1.064	0.700	0.209	0.700	0.210	0.698	0.211
1.102	0.873	0.255	0.873	0.255	0.872	0.256
1.152	1.049	0.294	1.049	0.295	1.048	0.297
1.215	1.224	0.328	1.223	0.329	1.222	0.331
1.294	1.399	0.356	1.398	0.358	1.396	0.360
1.393	1.575	0.377	1.573	0.379	1.570	0.381
1.518	1.752	0.391	1.749	0.392	1.746	0.396
1.885	2.107	0.396	2.104	0.399	2.094	0.402

Note:  $P_{cr} = \frac{\pi^2 EI}{L^2}$ ;  $w$  = center deflection; and  $L$  = length of the column

$\alpha$  = slope at the support



**Fig. 10 Imperfect simply supported beam buckling:  $P/P_{cr}$  versus central deflection for different amplitudes of imperfection**

ential equations, each having its validity over a chosen step-size and satisfying the nonlinear system at the two boundary points of the interval. In the present LTL approach, a boundary value problem is essentially treated as a constrained dynamical system or an initial value problem wherein the left boundary is treated as the initial point to start integrating the ordinary differential equations. To be precise, with the given step-size and prescribed boundary conditions at the left boundary (not all of which are known a priori at this boundary to start integration), the additional boundary conditions needed to convert the problem into an initial value problem are treated as unknown variables. Such an initial value problem is constrained in that these unknown initial conditions at the left boundary have to be so determined that the constructed solution satisfies the prescribed (boundary) conditions at the right boundary. This constrained and nonlinear dynamical system is in turn replaced, following the LTL philosophy, by a set of conditionally linearized ordinary differential equations. It is necessary to construct as many systems of such linearized ordinary differential equations as the number of segments into which the entire solution domain is discretized. Moreover the  $i$ th linearized system of ordinary differential equations is derived in such a way that the linearized solution manifold transversally intersects the (locally defined) nonlinear solution manifold at the  $i$ th discretization point (node) in the solution domain. The conditional linearity of these derived dynamical systems stems from the fact that the desired solution vectors enter these equations as unknown parameters. The postulated condition of transversal intersection between the original and linearized trajectories leads to a system of nonlinear algebraic equations in terms of these unknown solution vectors.

It may be noted that posing a boundary value problem as a constrained dynamical system within the LTL framework results in a coupled system of nonlinear algebraic equations in terms of the unknown solution parameters. Thus, unlike the case of an initial value problem (see, for instance, [23,24]), there is an added computational cost for treating boundary value problems with such a scheme. The enhanced cost may, however, be offset to a great extent by choosing a higher step size along the axis of the independent variable. Taking, for instance, a nonlinear beam post-buckling problem, there is enough numerical evidence to suggest that using as high a step-size as one-fourth of the length of the undeformed beam also leads to reasonably correct results when nonlinearity is small. Another distinctive advantage of the proposed method is its applicability to patently nonintegrable problems with weak and strong nonlinearities alike. Moreover, unlike in the case of homotopy-based techniques, the user of the LTL

technique is not burdened with the task of having to assume some form for the nonlinear equilibrium path. In fact, irrespective of the complexities of the equilibrium paths, the principle of LTL is able to predict them with reasonable (and often with very high) accuracy. Efforts are presently on to apply the same principle for boundary value problems governed by nonlinear partial differential equations.

## Acknowledgments

The authors wish to thank the Council for Scientific and Industrial Research (CSIR) in India for its financial support.

## References

- [1] Thurston, G. A., 1963, "A Numerical Solution of the Nonlinear Equations for Axisymmetric Bending of Shallow Spherical Shells," *ASME J. Appl. Mech.*, **83**, pp. 557–568.
- [2] Thurston, G. A., 1969, "Continuation of Newton's Method Through Bifurcation Points," *ASME J. Appl. Mech.*, **30**, pp. 425–430.
- [3] Kalaba, R., 1963, "Some Aspects of Quasilinearization," *Nonlinear Differential Equations and Nonlinear Mechanics*, Academic Press New York, pp. 135–146.
- [4] Damil, N., Potier-Ferry, M., Najah, A., Chari, R., and Lahmam, H., 1999, "An Iterative Method Based Upon Padé Approximants," *Int. J. Numer. Methods Eng.*, **45**, pp. 701–708.
- [5] Cao, H.-L., and Potier-Ferry, M., 1999, "An Improved Iterative Method for Large Strain Viscoplastic Problems," *Int. J. Numer. Methods Eng.*, **44**, pp. 155–176.
- [6] Way, S., 1934, "Bending of Circular Plates With Large Deflections," *ASME J. Appl. Mech.*, **56**, pp. 627–636.
- [7] Levy, S., 1942, "Bending of Rectangular Plates with Large Deflection," *NACA T.N. No. 846, NACA Rep. No. 737*.
- [8] Archer, R. R., 1957, "Stability Limits for a Clamped Spherical Shell Segment Under Uniform Pressure," *Q. Appl. Math.*, **15**, pp. 355–366.
- [9] Schimdt, R., and Da Deppo, D., 1974, "A New Approach to the Analysis of Shells, Plates and Membranes With Finite Deflections," *Int. J. Non-Linear Mech.*, **9**, pp. 409–419.
- [10] Vannucci, P., Cochelin, B., Damil, N., and Potier-Ferry, M., 1998, "An Asymptotic-Numerical Method to Compute Bifurcating Branches," *Int. J. Numer. Methods Eng.*, **41**, pp. 1365–1389.
- [11] Murthy, S. D. N., and Sherbourne, A. N., 1974, "Nonlinear Bending of Elastic Plates of Variable Profile," *J. Eng. Mech.*, **100**, No. EM2, pp. 251–265.
- [12] Turvey, G. J., 1978, "Large Deflection of Tapered Annular Plates by Dynamic Relaxation," *J. Eng. Mech.*, **104**, No. EM2, pp. 351–366.
- [13] Budiansky, B., 1959, "Buckling of Clamped Shallow Spherical Shells," *Proc. IUTAM Symp. on the Theory of Thin Elastic Shells*, W. T. Koiter, ed., North-Holland, Delft, Netherlands, pp. 64–94.
- [14] Thurston, G. A., 1961, "A Numerical Solution of the Nonlinear Equations for Axisymmetric Bending of Shallow Spherical Shells," *ASME J. Appl. Mech.*, **28**, pp. 557–562.
- [15] Kai-yuan, Yeh, Zhang, Xiao-jing, and Zhou, You-he, 1989, "An Analytical Formula of the Exact Solution to the von Karman Equations of a Circular Plate Under a Concentrated Load," *Int. J. Non-Linear Mech.*, **24**, pp. 551–560.
- [16] Kai-yuan, Yeh, Zhang, Xiao-jing, and Wang, Xin-zhi, 1990, "On Some Properties and Calculation of the Exact Solution to von Karman Equations of Circular Plates Under a Concentrated Load," *Int. J. Non-Linear Mech.*, **25**, pp. 17–31.
- [17] Hughes, T. J. R., and Liu, W. K., 1981, "Nonlinear Finite Element Analysis of Shells-II, Two-Dimensional Shells," *Comput. Methods Appl. Mech. Eng.*, **27**, pp. 167–181.
- [18] Reddy, J. N., and Singh, I. R., 1981, "Large Deflections and Large-Amplitude Free Vibrations of Straight and Curved Beams," *Int. J. Numer. Methods Eng.*, **17**, pp. 829–852.
- [19] Noor, A. K., and Peters, J. M., 1980, "Nonlinear Analysis via Global-Local Mixed Finite Element Approach," *Int. J. Numer. Methods Eng.*, **15**, pp. 1363–1380.
- [20] Pollandt, R., 1997, "Solving Nonlinear Differential Equations of Mechanics With the Boundary Element Method and Radial Basis Functions," *Int. J. Numer. Methods Eng.*, **40**, pp. 61–73.
- [21] He, Ji-Huan, 2000, "A Coupling Method of a Homotopy Technique and a Perturbation Technique for Non-linear Problems," *Int. J. Non-Linear Mech.*, **35**, pp. 37–43.
- [22] Gao, D. Y., 2000, "Finite Deformation Beam Models and Triality Theory in Dynamical Post-buckling Analysis," *Int. J. Non-Linear Mech.*, **35**, pp. 103–131.
- [23] Roy, D., and Ramachandra, L. S., 2001, "A Generalized Local Linearization Principle for Non-linear Dynamical Systems," *J. Sound Vib.*, **241**, pp. 653–679.
- [24] Roy, D., and Ramachandra, L. S., 2001, "A Semi-Analytical Locally Transversal Linearization Method for Non-linear Dynamical Systems," *Int. J. Numer. Methods Eng.*, **51**, pp. 203–224.

- [25] Ramachandra, L. S., and Roy, D., 2001, "An Accurate Semi-Analytical Technique for a Class of Non-linear Boundary Value Problems in Structural Mechanics," J. Eng. Mech., submitted for publication.
- [26] Chillingworth, D. R. J., 1976, *Differential Topology With a View to Applications*, Pitman, London.
- [27] Mattiasson, K., 1981, "Numerical Results From Large Deflection Beam and Frame Problems Analyzed by Means of Elliptic Integrals," Int. J. Numer. Methods Eng., **16**, pp. 145–153.
- [28] Timoshenko, S. P., and Gere, J. M., 1963, *Theory of Elastic Stability*, McGraw-Hill, New York.
- [29] Ortega, J. M., and Rheinboldt, W. C., 1970, *Iterative Solutions of Non-linear Equations in Several Variables*, Academic Press, New York.



## D. T. Corr<sup>1</sup>

e-mail: corr@cae.wisc.edu  
Departments of Orthopedic Surgery  
and Mechanical Engineering,  
University of Wisconsin,  
G5/332 Clinical Sciences Center,  
600 Highland Avenue,  
Madison, WI 53792-3228

## M. J. Starr

Department of Engineering Physics,  
University of Wisconsin,  
538 Engineering Research Building,  
1500 Engineering Drive,  
Madison, WI 53706

## R. Vanderby, Jr.

Mem. ASME  
Departments of Orthopedic Surgery,  
Mechanical Engineering  
and Biomedical Engineering,  
University of Wisconsin,  
G5/332 Clinical Sciences Center,  
600 Highland Avenue,  
Madison, WI 53792-3228

## T. M. Best

Departments of Family Medicine  
and Orthopedic Surgery,  
University of Wisconsin Medical School,  
621 Science Drive,  
Madison, WI 53711

# A Nonlinear Generalized Maxwell Fluid Model for Viscoelastic Materials

*A nonlinear Maxwell fluid model consisting of a linear dashpot in series with a parallel arrangement of a linear spring and a second-order nonlinear spring, was developed. This configuration provides the flexibility necessary to describe both the stiffening and the softening responses of some viscoelastic materials. A noteworthy feature of the model is that under constant rate displacement, the force equation can be solved in closed form, thereby providing a continuous, exact general solution. [DOI: 10.1115/1.1388615]*

## Introduction

The existing techniques for modeling stiffening, nonlinear viscoelastic materials have a variety of limitations. Quasi-linear viscoelasticity (QLV) and other integral techniques require integration of time-dependent material functions as modeling parameters, resulting in an increase of computational complexity and often a lack of physical understanding of the model's coefficients. Numerical solutions provide approximations to the exact solution, and are susceptible to instabilities. Spring-and-dashpot phenomenological models describe the force-time relationship with constant-value modeling coefficients. However, without the inclusion of additional nonconstant elements, these models are incapable of producing nonlinear elastic responses ([1,2]).

In this study, a constant parameter spring-and-dashpot model was developed to capture the strain stiffening response of the low load region of some viscoelastic materials. The resulting model is capable of capturing both strain stiffening and material softening. Additionally, for constant rate elongations, a closed-form solution was found for the resulting Riccati differential equation.

## Nonlinear Model Development

A nonlinear Maxwell fluid spring-and-dashpot model was developed to describe the complex nonlinear behavior of some vis-

coelastic materials. The standard Maxwell fluid model was chosen as the foundation to capture the softening behavior that is typically present outside the low-load region. A nonlinear second-order spring ( $k_2$ ) was added in parallel to the linear ( $k_1$ ) spring (Fig. 1). This combination allows the second-order spring to capture the nonlinear stiffening response of the initial low-load region while the Maxwell fluid portion can model the linearity and eventual softening of the material. While each of the two springs has its own unique contribution to the overall stiffness, it is easiest to view the parallel combination as a black-box equivalent spring. Thus, the equivalent spring stiffness is comprised of both linear and quadratic responses. With this approach, the physical restrictions on each spring can be suspended, as long as the resulting equivalent spring does not violate physical limitations. For example, a negative linear spring constant, while not physically reasonable, in combination with a positive quadratic spring results in an equivalent stiffness between first and second order.

## Model Derivations

The major drawback of adding a nonlinear element to the model is the loss of linearity in the differential equation that describes the system. Finding a closed-form solution to such equations can be quite daunting. The current model was derived without any simplifications, for a constant rate of displacement,  $\alpha$ , to the end result of a closed-form solution. By first imposing equilibrium it can be seen that

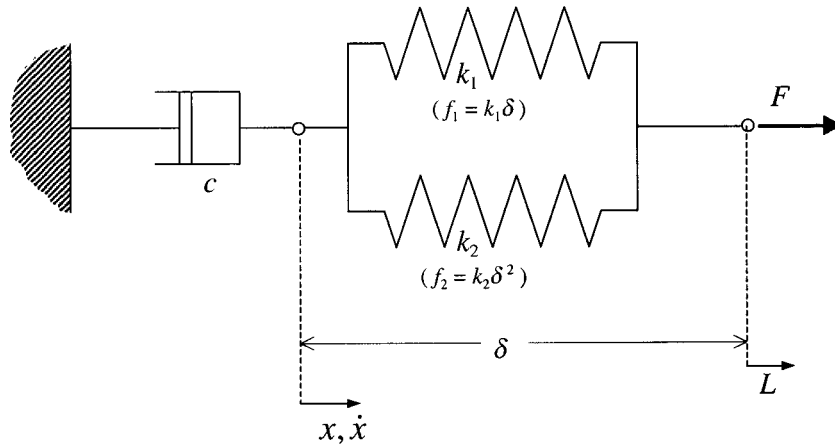
$$c\dot{x} = k_1\delta + k_2\delta^2 \quad (1)$$

where  $\delta$  and  $x$  are the displacement of the springs and dashpot from the rest position, respectively.

When subjected to a displacement controlled stretch at a constant rate,  $\alpha$ , the length equations can be written as

<sup>1</sup>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, September 20, 2000; final revision, April 26, 2001. Associate Editor: D. A. Siginer. Discussion on the paper should be addressed to the 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.



**Fig. 1 The proposed nonlinear model, consisting of a parallel arrangement of a linear spring ( $k_1$ ) and a second-order spring ( $k_2$ ), in series with a linear dashpot ( $c$ ).**

$$L = L_o + \alpha(t - t_o) \quad (2a)$$

$$L = L_o + x + \delta \quad (2b)$$

in which the subscript  $o$  denotes the start of elongation. Thus, a test starting at time  $t_o = 0$  produces the following relation,  $\alpha t = x + \delta$ . Solving for  $\delta$  and substituting into (1) yields

$$c\dot{x} = k_1(\alpha t - x) + k_2(\alpha t - x)^2. \quad (3)$$

Noting that the dashpot displacement at the onset of elongation is zero-valued,  $x(0) = 0$ , and allowing  $\mu = k_1/c$ , and  $\xi = k_2/c$ , the nonlinear Eq. (3) can be rewritten in the Riccati differential equation form

$$\dot{x} = a(t)x^2 + b(t)x + c(t) \quad (4)$$

where

$$a(t) = \xi$$

$$b(t) = -(\mu + 2\xi\alpha t)$$

$$c(t) = \alpha t(\mu + \xi\alpha t).$$

### Determining a Solution

The aforementioned Riccati Eq. (4) is not readily solvable by any nonnumerical techniques, and to the best knowledge of the authors, has not previously been solved in closed form. In order to solve this nonlinear nonhomogeneous first-order differential equation, a particular solution is first needed. By picking a solution form of  $x_p = At + B$ ,  $\dot{x}_p = A$ , (4) becomes

$$A = \xi(At + B)^2 - (\mu + 2\xi\alpha t)(At + B) + \alpha t(\mu + \xi\alpha t). \quad (5)$$

Matching orders of  $t$  yields

$$O(1): A = \xi B^2 - \mu B \quad (i)$$

$$O(t): 0 = 2\xi AB - \mu A - 2\xi\alpha B + \mu\alpha \quad (ii)$$

$$O(t^2): 0 = \xi A^2 - 2\xi\alpha A + \xi\alpha^2. \quad (iii)$$

Solving the above equations for  $A$  and  $B$  gives

$$(iii): A = \alpha$$

$$(ii): 0 = 0 \text{ identically satisfied}$$

$$(i): B = \frac{\mu \pm \sqrt{\mu^2 + 4\xi\alpha}}{2\xi}.$$

Thus,  $x_p = \alpha t + B$  is a particular solution to (4).

In order to determine the general solution, the reduction of order was implemented to gain the complementary solution. Allowing the complementary solution to be represented as  $u(t)$ , the general solution takes on the unique form (6), as guaranteed

$$x(t) = x_p(t) + u(t) = \alpha t + B + u(t) \quad (6a)$$

$$\dot{x}(t) = \alpha + \dot{u}(t) \quad (6b)$$

by the fundamental existence and uniqueness theorem. Substituting into (4) yields

$$\alpha + \dot{u} = \xi B^2 + 2\xi B u + \xi u^2 - \mu B - \mu u. \quad (7)$$

This, in turn simplifies to a Bernoulli equation by employing (i) and (iii),

$$\dot{u} = (2\xi B - \mu)u + \xi u^2. \quad (8)$$

This result is a nonlinear first order ordinary differential equation. Since the coefficients are constant, it is separable. Letting  $\phi = 2\xi B - \mu$ , (8) can be separated and rewritten as

$$\frac{du}{u(\phi + \xi u)} = dt. \quad (9)$$

Expanding with partial fractions gives

$$\frac{1}{\phi} \left[ \frac{1}{u} - \frac{\xi}{(\phi + \xi u)} \right] du = dt, \quad (10)$$

which when integrated and simplified gives

$$u = \frac{\phi K e^{\phi t}}{1 - K \xi e^{\phi t}}. \quad (11)$$

Applying the initial condition,  $x(0) = 0$ , to (6a),  $u(0) = u_o = -B$ , and solving for  $K$  yields

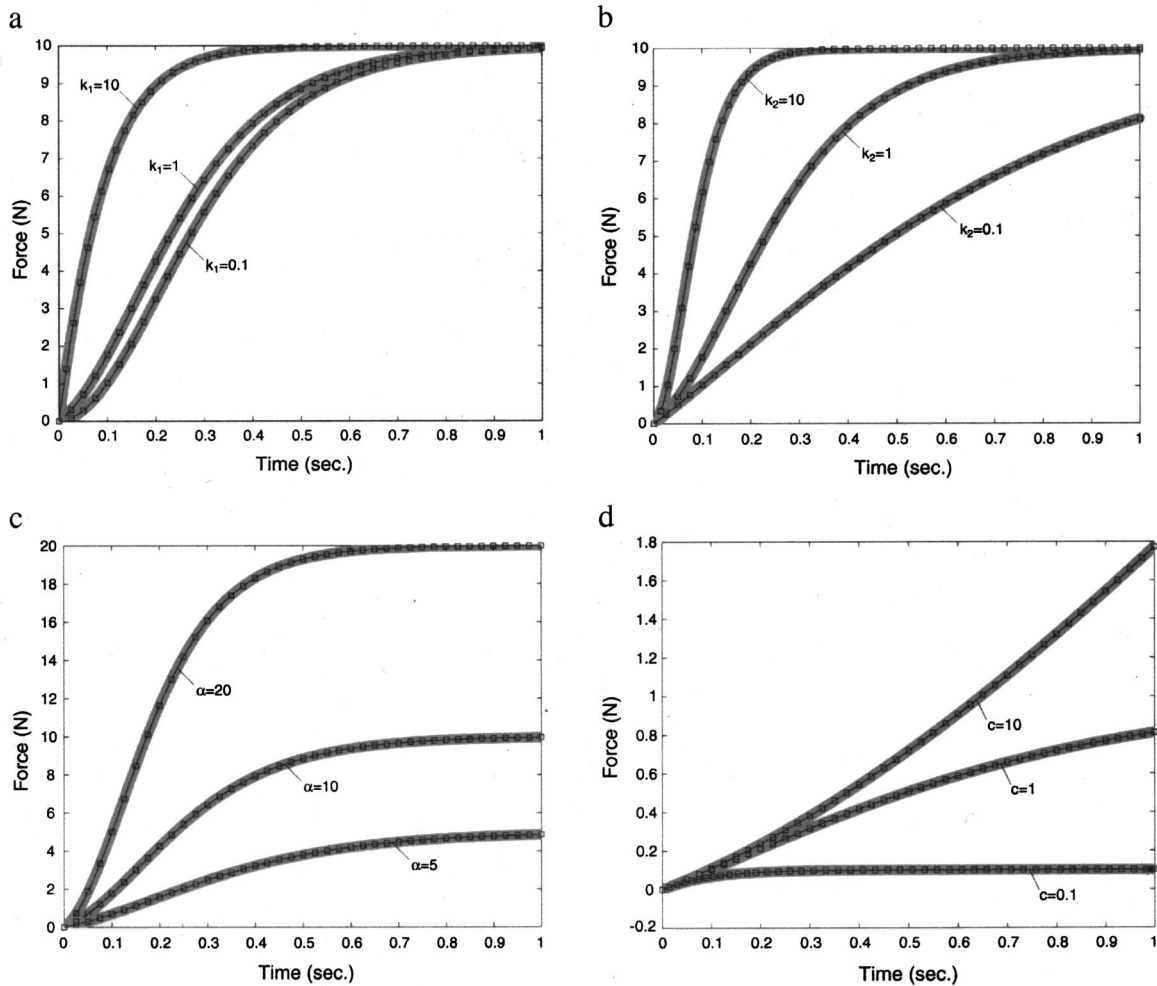
$$K = \frac{B}{B\xi - \phi}. \quad (12)$$

The complementary solution becomes

$$u(t) = \frac{\phi B e^{\phi t}}{B\xi(1 - e^{\phi t}) - \phi}. \quad (13)$$

Thus, the unique general closed-form solution is

$$x(t) = \alpha t + B + \frac{\phi B e^{\phi t}}{B\xi(1 - e^{\phi t}) - \phi} \quad (14)$$



**Fig. 2** Parameter sensitivity illustrated via force-time plots of numerical solutions (hollow squares), positive root (thick gray line), and negative root closed-form solutions (black line). (a) Three decades of linear spring stiffness,  $k_1 = 0.1, 1, 10$  N/cm<sup>2</sup> ( $\alpha = 10$  cm/s,  $k_2 = 1$  N/cm<sup>2</sup>,  $c = 1$  N·s/cm). As  $k_1$  increases the slope increases and the nonlinearity of the low load region becomes less prominent. Variation of the linear stiffness ( $k_1$ ) has no effect on the peak load value, however, increasing  $k_1$  shortens the time needed to reach it. (b) Three decades of second-order spring stiffness,  $k_2 = 0.1, 1, 10$  N/cm<sup>2</sup> ( $\alpha = 10$  cm/s,  $k_1 = 1$  N/cm,  $c = 1$  N·s/cm). As  $k_2$  increases the slope increases and the nonlinearity of the low load region becomes more prominent. Variation of  $k_2$  has no effect on the maximum load value, however, increasing  $k_2$  shortens the time needed to reach the peak load. (c) Varying displacement rates,  $\alpha = 5, 10, 20$  cm/s ( $k_1 = 1$  N/cm,  $k_2 = 1$  N/cm<sup>2</sup>,  $c = 1$  N·s/cm). Note that the stiffness increases with increasing  $\alpha$ , and the load approaches a value of  $c \cdot \alpha$  as time evolves. (d) Three decades of dashpot values,  $c = 0.1, 1, 10$  N·s/cm ( $\alpha = 10$  cm/s,  $k_1 = 1$  N/cm,  $k_2 = 1$  N/cm<sup>2</sup>). An increase in the viscous damping,  $c$ , increases the value of the peak load as well as the time required to reach maximum force.

where

$$B = \frac{\mu \pm \sqrt{\mu^2 + 4\xi\alpha}}{2\xi}$$

$$\mu = \frac{k_1}{c}, \quad \xi = \frac{k_2}{c}, \quad \phi = 2\xi B - \mu$$

$\alpha = \text{constant rate of displacement.}$

Therefore, the resulting velocity equation is

$$\dot{x} = \alpha + \frac{\phi^2 B e^{\phi t} (B\xi - \phi)}{B^2 \xi^2 e^{2\phi t} - 2B\xi e^{\phi t} (B\xi - \phi) + (B\xi - \phi)^2}. \quad (15)$$

When examined via limit analysis, (15) obtains

$$\lim_{t \rightarrow \infty} (\dot{x}) = \alpha. \quad (16)$$

This result makes physical sense since at long time periods the velocity of the dashpot will approach the controlled displacement rate.

### Higher-Order Models

To expand the range of order in the equivalent spring, and thereby increase the model's flexibility, additional spring combinations were explored. Higher-order models were created by adding either a third-order, or a third and a fourth-order spring, each in parallel, to the linear and second-order springs, resulting in equivalent springs capable of describing responses with cubic and quartic components, respectively.

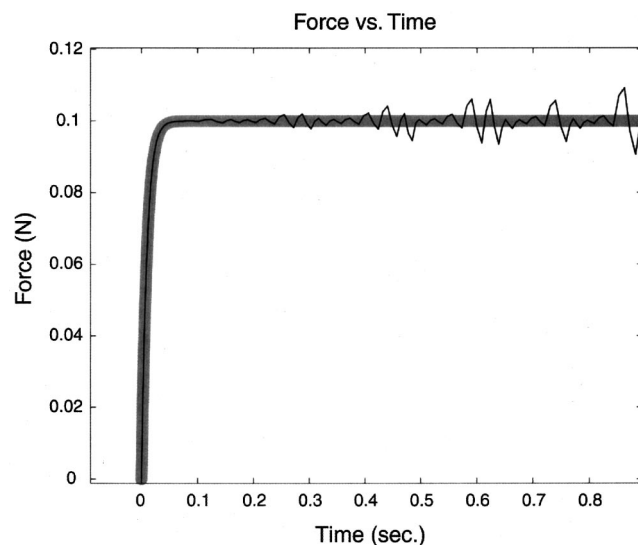
An exploration of the associated differential equations shows that if a general polynomial function is assumed, and closed-form solutions are sought, then only nonlinear representations up to and including degree four are allowed by the mathematical approaches taken herein. A deeper exploration shows that no closed-form so-

lutions for general polynomials above degree two are admissible. The second-degree model admits an exact solution because the complementary equation is separable, and the dependent variable,  $u$ , can be solved explicitly as a function of time. However, for the third-degree model, the complementary solution is again separable, but no explicit solution for  $u(t)$  can be found. Further, an expansion solution can be found, but the expansions must be truncated after the second degree so that an explicit representation of  $u(t)$  can be found. This is due to the fact that closed-form solutions for the roots of polynomial equations can only be found, in general, up to the fourth degree. Using this expression of  $u(t)$ , an approximate solution for general polynomials of degree three can be found. The fourth-degree model yields a separable equation, but is not generally integrable. The same holds true for all general higher degree models.

### Model Behavior

The closed-form solution of the model has two solutions, corresponding to the positive and negative roots of  $B$  (14). A numerical solution was obtained within MATLAB using a medium order method to solve nonstiff ordinary differential equations. Equation (4) did not exhibit rapidly decaying solutions, and thus was not considered stiff. The positive root, negative root, and numerical solution force-time traces were plotted for a wide range of  $k_1$ ,  $k_2$ , and  $c$  values, demonstrating the model's ability to reproduce initial nonlinear stiffening, followed by material softening (Fig. 2).

The sensitivity of the force-time relationship to given parameters was studied by varying one coefficient while holding the other three constant. As linear spring values ( $k_1$ ) increased, the nonlinearity of the low-load region became less prominent, the slope increased, and the model approached a standard Maxwell fluid response (Fig. 2(a)). Similarly, an increase in the second-order spring ( $k_2$ ) increased the slope, but made the nonlinearities more prominent (Fig. 2(b)). Increasing either spring constant reduced the time needed to attain peak load, but had no effect on its value. As the displacement rate ( $\alpha$ ) increased, both stiffness and peak load increased (Fig. 2(c)). The viscous damping ( $c$ ) directly affected the maximum value and time period of the load-time response (Fig. 2(d)).



**Fig. 3** Force-time plot for  $\alpha=10$ ,  $k_1=1$  N/cm,  $k_2=0.5$  N/cm<sup>2</sup>,  $c=0.001$  N-s/cm. Note the very close agreement between the numerical solution (thin black line) and the positive root (thick gray line) closed-form solutions until  $t>0.1$  sec, at which the numerical solution becomes unstable.

**Table 1** Numerical solution instability ratios

$\alpha/c$	$k_1/c$	$k_2/c$
$>300$	1	1
1	$>34.6$	1
1	1	$>300$

The numerical solution showed great agreement with both closed-form solutions; however, there were some extreme cases in which the numerical solution became unstable (Fig. 3). Stability limits were found for each parameter by setting all parameters equal to  $c$ , and increasing a particular parameter until instabilities began to arise (Table 1).

### Discussion

This particular type of model was chosen due to its relative simplicity and computational clarity. As noted by Lockett [1], "The main value of spring-dashpot models is in providing a relatively simple qualitative picture of the manner in which viscoelastic materials behave, and this is only feasible when the models contain only a few elements." Thus, our three-element model facilitates an improved understanding of material behavior, and how it is affected by the individual model components.

The positive aspects of model simplicity do not come without a cost. In order for a differential equation description of materials to function over a wide range of time or frequency, it is usually required to have higher order equations with many terms ([3]). Attempts to increase the order of our model's equivalent spring, and therefore increase model flexibility, resulted in unattainable closed-form solutions. Thus, this model sacrifices a degree of flexibility for an increase in the understanding and clarity of the material response, as do the classic models of Maxwell and Kelvin (Voigt).

Time and frequency flexibility is typically achieved through the use of integral equations. While integral equations greatly improve model flexibility, they do not clearly establish the material's response in terms of an equivalent stiffness, its order, and the strength of viscous damping. These properties may be available; however, they are not as clearly identified, and their contribution is hard to separate from the overall response.

The closed-form solution of our model was determined for constant-rate displacement-control testing. The model is fully capable of handling rates of displacement that are functions of time; however, the process taken to obtain the closed-form solution will not provide a particular solution to the resulting differential equation. Therefore, it appears that numerical solutions remain the only option for the treatment of nonconstant displacement rates.

### Conclusion

A new phenomenologic model, while a fluid model, captures both the nonlinear stiffening and the softening response exhibited by a variety of viscoelastic materials. The relative simplicity of this model helps isolate the contributions of the viscous damping, linear, and second-order stiffness. Moreover, the closed-form solution to the resulting Riccati differential equation provides a previously unavailable exact solution.

### Acknowledgments

The authors would like to thank the financial support of The Whitaker Foundation [TMB], as well as University of Wisconsin Surgical Associates.

### References

- [1] Lockett, F. J., 1972, *Nonlinear Viscoelastic Solids*, Academic Press, London.
- [2] Forcinito, M., Epstein, M., and Herzog, W., 1998, "Can a Rheological Muscle Model Predict Force Depression/Enhancement?" *J. Biomech.*, **31**, pp. 1093–1099.
- [3] Lakes, R. S., 1999, *Viscoelastic Solids*, CRC Press, Boca Raton, FL.



# Structural Modification for the Assignment of Zeros Using Measured Receptances

J. E. Mottershead

Alexander Elder Professor of Applied Mechanics,  
Department of Engineering,  
The University of Liverpool,  
Liverpool L69 3GH, UK

*In many engineering applications it is desirable to assign the zeros of point or cross receptances to particular frequencies. This means that at the chosen coordinate no vibration will be experienced at those frequencies. It is shown how such an objective can be achieved by means of passive stiffness, damping, and mass modifications to the structure. The frequency responses of a subsidiary eigenvalue problem, the eigenvalue problem of the zeros, are determined in order that the methods of inverse structural modification may be applied. The technique requires only a small number of measurements from the structure in its unmodified condition. Numerical and physical experiments are carried out to illustrate the application of the method. [DOI: 10.1115/1.1388616]*

## Introduction

There are two structural modification problems: forward and inverse. The *forward structural modification problem* is to determine the eigenvalues and eigenvectors of a system with a known modification to its stiffness and mass terms. This procedure, sometimes known as re-analysis, is common in vibration optimization and in finite element model updating. Usually a close approximation to the eigenvalues of the modified system is acceptable. Baldwin and Hutton [1] reviewed a number of methods including Rayleigh quotient, sensitivity, and perturbation techniques. Ram [2] determined the eigenvalues of damped subsystems with known connections using transfer function or spectral and modal data from the separate subsystems.

The *inverse structural modification problem* is to determine the modification (the added stiffness, damping, or mass) that will bring about a desired change in the eigenvalues and eigenvectors of a vibrating system. The inverse problem is generally more demanding than the forward problem and can be applied to measured vibration data to determine a modification without the need for a finite element model. The modification may be passive or active. Weissenburger [3] and Pomazal and Snyder [4] used unit-rank modifications to relocate a single natural frequency. The other natural frequencies were affected by the modification but not controlled. Mottershead and Lallement [5] used the same method to bring about the cancellation of a pole with a zero thereby creating a vibration node. Bucher and Braun [6,7] developed an exact method for the assignment of vibration mode shapes from spectra and modes, including measured left eigenvectors. The left eigenvectors were determined from noisy measurements using regularization by L-curves. Tsuei and Yee [8] described a method for shifting natural frequencies by using measured frequency response data. Li et al. [9] used a similar method to create a global pole-zero cancellation (cancellation of a natural frequency with an antiresonance) which left another, uncanceled, pole at the same frequency. Mottershead et al. [10] showed how the nodes of normal modes (and natural frequencies) could be assigned using measured receptances. The modifications were determined from the null space of a matrix containing the measured vibration data.

Until now inverse structural modification has not been applied

to adjust the zeros (antiresonances) of measured receptances. But this is a desirable objective, with many engineering applications, because the zeros define those frequencies at which vibrations disappear. The dynamic vibration absorber, a good description of which can be found in Inman [11], is really a device for the assignment of point-receptance zeros. In the case of the classical undamped vibration absorber the zeros lie on the imaginary axis of the complex eigenvalue plane and when the absorber includes a damper the zeros generally become complex. Ram and Elhay [12] determined the parameters of a multi-degree-of-freedom undamped absorber secondary system by solving an inverse eigenvalue problem. In a recent survey paper Sun et al. [13] cited some 90 papers on passive and active tuned vibration absorbers.

The zeros of different point and cross receptances generally occur at different frequencies, whereas the poles are unchanged. They can be determined by solving a "subsidiary" eigenvalue problem. The subsidiary problem is symmetric for point-receptance zeros, and asymmetric for cross-receptance zeros (Mottershead [14]). In numerical studies, without damping, the subsidiary matrices are formed from the stiffness and mass ( $\mathbf{K}$ ,  $\mathbf{M}$ ) matrices of the system by deleting a single row and column. When the row and column have the same index the resulting matrix system will be symmetric and its eigenvalues will be the zeros of a point receptance. They will interlace the eigenvalues of the poles. When the deleted row and column have different indices the resulting matrices will not be symmetric, interlacing rules will not apply, and the eigenvalues (cross-receptance zeros) may become complex and/or defective (Mottershead [15]).

In this paper it is shown how the eigenvectors and frequency responses of the subsidiary eigenvalue problem of the zeros can be obtained from receptance measurements from the whole system. The method can be applied to point and cross-receptance zeros and thereby allows existing techniques to be applied to assign zeros by structural modification. In the cross-receptance case it is necessary to take four measurements at the frequency to which a zero will be assigned. For point receptances only three measurements are needed. The inverse method elaborated in the following sections differs completely from the well-known problem of assigning the poles and zeros of a linear control system (Kailath [16]). It has the advantage (over the dynamic vibration absorber) that zeros can be assigned to both point and cross receptances, and whereas an absorber must be attached at the same coordinate as the assigned point-receptance zero, the modification can be applied at a different coordinate. The method is illustrated with numerical and physical examples.

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, October 4, 2000; final revision, May 17, 2001. Associate Editor: A. A. Ferri. Discussion on the paper should be addressed to the 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.



## Theory

We write the system stiffness, damping, and mass matrices,  $\mathbf{K} = \mathbf{K}^T \geq 0$  (or  $> 0$ ),  $\mathbf{C} = \mathbf{C}^T \geq 0$  and  $\mathbf{M} = \mathbf{M}^T > 0$ , in the form

$$\begin{bmatrix} k_{pq} & \mathbf{k}_p^T \\ \mathbf{k}_q & \mathbf{K}_{pq} \end{bmatrix} \in \mathbb{R}^{n \times n}, \quad (1)$$

$$\begin{bmatrix} c_{pq} & \mathbf{c}_p^T \\ \mathbf{c}_q & \mathbf{C}_{pq} \end{bmatrix} \in \mathbb{R}^{n \times n}, \quad (2)$$

$$\begin{bmatrix} m_{pq} & \mathbf{m}_p^T \\ \mathbf{m}_q & \mathbf{M}_{pq} \end{bmatrix} \in \mathbb{R}^{n \times n}. \quad (3)$$

$\mathbf{K}_{pq}$  is the matrix formed from  $\mathbf{K}$  by deleting the  $p$ th row and  $q$ th column,  $k_{pq}$  is the  $pq$ th term of  $\mathbf{K}$ ,  $\mathbf{k}_p^T$  is the  $p$ th row of  $\mathbf{K}$  (except for  $k_{pq}$ ) and  $\mathbf{k}_q$  is the  $q$ th column of  $\mathbf{K}$  (except for  $k_{pq}$ ). Similar definitions apply to the terms in the partitioned damping and mass matrices.

The matrices  $\mathbf{K}_{pq}, \mathbf{C}_{pq}, \mathbf{M}_{pq}$  define the subsidiary system

$$(\mathbf{K}_{pq} + \bar{s}_i \mathbf{C}_{pq} + \bar{s}_i^2 \mathbf{M}_{pq}) \psi_i = 0, \quad i = 1, \dots, n, \quad (4)$$

where the eigenvalues  $\bar{s}_i$  determine the frequencies of the zeros of the frequency responses  $h_{pq} = h_{qp}$ . The physical meaning of the eigenvectors,  $\psi_i$ , and the subsidiary frequency responses,

$$\mathbf{H}_{pq}(s) = (\mathbf{K}_{pq} + s \mathbf{C}_{pq} + s^2 \mathbf{M}_{pq})^{-1}, \quad s = i\omega, \quad (5)$$

is obscure but can be used to assign the zeros by structural modification. When  $p = q$  the subsidiary system represents a point receptance and when  $p \neq q$  a cross receptance.

## Eigenvectors of the Subsidiary System

We define the frequency responses  $\mathbf{H}(s)$ ,  $s = i\omega$ , of the complete system by the relationship

$$\mathbf{B}(s) \mathbf{H}(s) = \mathbf{I}_{n \times n}, \quad (6)$$

where

$$\mathbf{B}(s) = (\mathbf{K} + s \mathbf{C} + s^2 \mathbf{M}). \quad (7)$$

When the  $q$ th column of  $\mathbf{H}(\bar{s}_i)$  is considered (at the  $i$ th zero of the frequency response  $h_{pq}$ ), then

$$\mathbf{B}_p(\bar{s}_i) \begin{bmatrix} h_{1q} \\ h_{2q} \\ \vdots \\ h_{p-1,q} \\ h_{p+1,q} \\ \vdots \\ h_{nq} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \leftarrow q\text{th entry} \quad (8)$$

where  $\mathbf{B}_p(\bar{s}_i) \in \mathbb{R}^{n \times (n-1)}$  is the same as  $\mathbf{B}(\bar{s}_i)$  except that the  $p$ th column is omitted. When the  $q$ th row as well as the  $p$ th column is taken out we find that

$$\mathbf{B}_{qp}(\bar{s}_i) \begin{bmatrix} h_{1q} \\ h_{2q} \\ \vdots \\ h_{p-1,q} \\ h_{p+1,q} \\ \vdots \\ h_{nq} \end{bmatrix} = 0, \quad (9)$$

so that  $\mathbf{h}_q = (h_{1q} h_{2q} \dots h_{p-1,q} h_{p+1,q} \dots h_{nq})^T$  is an eigenvector of  $\mathbf{B}_{qp}(\bar{s}_i)$ . A similar analysis (when  $p \neq q$ ) shows that  $\mathbf{h}_p$  is a left eigenvector of  $\mathbf{B}_{qp}(\bar{s}_i)$ . If  $p = q$  then  $\mathbf{B}_{pp}(s)$  is symmetric so that its eigenvalues have nonpositive real parts.

## Frequency Responses of the Subsidiary System

The frequency response equation for the complete system can be expressed as

$$\mathbf{H}(s) \mathbf{f}(s) = \mathbf{x}(s), \quad s = i\omega, \quad (10)$$

and in the case of a zero of  $h_{pq}(s)$  the force applied at coordinate  $q$  gives rise to zero displacement at coordinate  $p$ . We write Eq. (10) in full (and with rows and columns rearranged) as

$$\begin{bmatrix} h_{pq} & h_{p1} & h_{p2} & \dots & h_{pn} \\ h_{1q} & h_{11} & h_{12} & & \\ h_{2q} & h_{21} & h_{22} & & \\ \vdots & & & \ddots & \\ h_{nq} & & & & h_{nn} \end{bmatrix} \begin{Bmatrix} f_q^* \\ f_1 \\ f_2 \\ \vdots \\ f_n \end{Bmatrix} = \begin{Bmatrix} x_p = 0 \\ x_1 \\ x_2 \\ \vdots \\ x_n \end{Bmatrix}. \quad (11)$$

The first row of Eq. (11) leads to the expression

$$f_q^* = -\frac{1}{h_{pq}} (h_{p1} h_{p2} \dots h_{p,q-1} h_{p,q+1} \dots h_{pn}) \begin{Bmatrix} f_1 \\ f_2 \\ \vdots \\ f_{q-1} \\ f_{q+1} \\ \vdots \\ f_n \end{Bmatrix}, \quad (12)$$

$$\text{or } \mathbf{f}(s) = \mathbf{A}(s) \tilde{\mathbf{f}}(s), \quad (13)$$

where

$$\mathbf{f}(s) = (f_q^* f_1 f_2 \dots f_{q-1} f_{q+1} \dots f_n)^T, \quad (14)$$

$$\tilde{\mathbf{f}}(s) = (f_1 f_2 \dots f_{q-1} f_{q+1} \dots f_n)^T, \quad (15)$$

$$\mathbf{A} = \begin{bmatrix} -\frac{1}{h_{pq}} (h_{p1} h_{p2} \dots h_{p,q-1} h_{p,q+1} \dots h_{pn}) \\ \mathbf{I}_{(n-1) \times (n-1)} \end{bmatrix}. \quad (16)$$

$f_q^*(s)$  is the force at coordinate  $q$  that maintains  $x_p = 0$  across the frequency range of the measured data from the complete system. Equations (12) and (13) have a unique solution except at the frequency of a zero  $h_{pq}(\bar{s}_i) = 0$ . In that case  $\tilde{\mathbf{f}}(\bar{s}_i) = 0$  simultaneously and there exists an infinity of solutions to Eqs. (12) and (13). In practice this hardly ever arises and would only do so when by coincidence a discrete frequency of the digital frequency response data was at exactly the same frequency as a zero of  $h_{pq}$ .

A similar analysis to the above, but based on the force  $f_p^*$  to produce  $x_q = 0$ , yields

$$\bar{\mathbf{A}} = \begin{bmatrix} -\frac{1}{h_{qp}} (h_{q1} h_{q2} \dots h_{q,p-1} h_{q,p+1} \dots h_{qn}) \\ \mathbf{I}_{(n-1) \times (n-1)} \end{bmatrix}. \quad (17)$$

Now we combine Eqs. (10) and (13) and premultiply by  $\bar{\mathbf{A}}^T(s)$  to obtain,

$$\mathbf{H}_{pq}(s) \tilde{\mathbf{f}}(s) = \tilde{\mathbf{x}}(s), \quad (18)$$

where

$$\mathbf{H}_{pq}(s) = \bar{\mathbf{A}}^T(s) \mathbf{H}(s) \mathbf{A}(s) \in \mathbb{C}^{(n-1) \times (n-1)}, \quad (19)$$

and

$$\tilde{\mathbf{x}}(s) = \bar{\mathbf{A}}^T(s) \mathbf{x}(s) \in \mathbb{C}^{(n-1)}, \quad (20)$$

or

$$\tilde{\mathbf{x}}(s) = (x_1 x_2 \dots x_{p-1} x_{p+1} \dots x_n)^T, \quad (21)$$

because of the constraint  $x_p = 0$ .

The frequency response  $\tilde{h}_{it}$  is now defined to be the term from  $\mathbf{H}_{pq}(s)$ ,  $s=i\omega$ , at the coordinate  $t$  of the original system (and  $t \neq p, q$ ).  $\tilde{h}_{it}$  will not necessarily lie on the diagonal of  $\mathbf{H}_{pq}(s)$ , but can in any case be shown, from Eq. (19), to be given by

$$\tilde{h}_{it} = h_{it} - (h_{pt}h_{qt})/h_{pq}. \quad (22)$$

This subsidiary-system frequency response is used in the following section to assign a zero in  $h_{pq}$  by a modification at coordinate  $t$ . It should be noted that it is only necessary to measure four frequency responses  $h_{it}$ ,  $h_{pt}$ ,  $h_{qt}$ , and  $h_{pq}$  in order to obtain  $\tilde{h}_{it}$ . In the case when  $p=q$  Eq. (22) can be simplified to give

$$\tilde{h}_{it} = h_{it} - h_{pt}^2/h_{pp}, \quad (23)$$

so that it is only necessary to measure three frequency responses.

### Assignment of Zeros

In this section we review the method presented by Pomazal and Snyder [4] and applied subsequently by Zhang and Lallement [17] and Mottershead and Lallement [5]. The eigenvalue equation of the system with a modification  $b_f$  at the  $f$ th coordinate can be written as

$$(\mathbf{K} + s_r \mathbf{C} + s_r^2 \mathbf{M} + b_f \mathbf{e}_f \mathbf{e}_f^T) \boldsymbol{\varphi}_r = \mathbf{0}, \quad r = 1, \dots, n, \quad (24)$$

where

$$h_f = s_r^2 m_f + s_r c_f + k_f, \quad (25)$$

$s_r$  and  $\boldsymbol{\varphi}_r$  are the  $r$ th eigenvalue and eigenvector of the modified system and  $\mathbf{e}_f$  is the  $f$ th column of  $\mathbf{I}_{n \times n}$ . When Eq. (24) is pre-multiplied by  $\mathbf{H}(s_r) = (\mathbf{K} + s_r \mathbf{C} + s_r^2 \mathbf{M})^{-1}$  it is found that

$$(\mathbf{I} + b_f \mathbf{H}(s_r) \mathbf{e}_f \mathbf{e}_f^T) \boldsymbol{\varphi}_r = \mathbf{0}, \quad (26)$$

the  $f$ th term of which is given by

$$(1 + \mathbf{e}_f^T b_f \mathbf{H}(s_r) \mathbf{e}_f) \boldsymbol{\varphi}_r = 0. \quad (27)$$

Since  $\mathbf{e}_f^T \boldsymbol{\varphi}_r \neq 0$  it can be seen that

$$-\frac{1}{b_f} = h_{ff}(s_r). \quad (28)$$

For frequency responses, when  $s=i\omega$ , the modifications can be expressed as

$$k_f - \omega_r^2 m_f = -\frac{\cos \angle h_{ff}(\omega_r)}{|h_{ff}(\omega_r)|}, \quad (29)$$

and

$$\omega_r c_f = \frac{\sin \angle h_{ff}(\omega_r)}{|h_{ff}(\omega_r)|}. \quad (30)$$

Clearly a passive modification ( $k_f$  or  $m_f$ ) can always be made depending upon the sign of the right-hand side of Eq. (29). For a stable system having unmodified eigenvalues with nonpositive real parts the assignment of a pole to the imaginary axis would require a negative damping modification. Singh and Ram [18] considered the assignment of point-receptance zeros by passive modifications. They established mutually exclusive conditions for the assignment of a positive mass or stiffness-modification for a constrained mass-spring system. Their paper contains the same conclusion on the need to add negative damping for vibration absorption in a damped system, but by a different argument.

In the case of assigning a zero to  $h_{pq}$  by a modification at coordinate  $t$  it is only necessary to replace  $h_{ff}(s_r)$  with  $\tilde{h}_{it}(\tilde{s}_i)$  in Eq. (28), where  $\tilde{s}_i$  now denotes the zero to be assigned (not a zero of the unmodified system). By combining Eqs. (22) and (28), when  $p \neq q$ , we obtain

$$-\frac{1}{b_t} = h_{it} - (h_{pt}h_{qt})/h_{pq}, \quad (31)$$

and when  $p=q$ ,

$$-\frac{1}{b_t} = h_{it} - h_{pt}^2/h_{pp}. \quad (32)$$

The choice of modification coordinate has been considered by previous authors (Pomazal and Snyder [4], Zhang and Lallement [17], and Mottershead and Lallement [5]). In particular if the  $i$ th eigenvector (or left eigenvector)  $\mathbf{h}_p(\tilde{s}_i)$  (or  $\mathbf{h}_q(\tilde{s}_i)$ ) contain a zero entry at coordinate  $t$  then the zero  $\tilde{s}_i$  will be unchanged by the modification. In order to reduce the magnitude of the modification  $b_t$  to acceptable levels it is usual to choose a  $t$  corresponding to relatively a large entry in  $h_p(\tilde{s}_i)$  (and  $h_q(\tilde{s}_i)$ ).

### Numerical Example

The six-degree-of-freedom system shown in Fig. 1 is considered where the value of all the stiffnesses and masses is unity.

**Case 1—Undamped System.** The “measured” cross receptance  $h_{45}$  of the unmodified system is shown in Fig. 2 and it is required to assign a zero at 1.5 rad/s by a modification at coordinate 3. The eigenvalues (poles and zeros) of the system are given in Table 1. A modification in the form of a grounded spring  $k^* = 0.4605$  relocates the zeros to frequencies shown in Table 2 and the receptance of the modified system, together with the subsidiary-system receptance at the modification coordinate, is illustrated in Fig. 3. It is clear that the poles of the subsidiary system correspond to the zeros of  $h_{45}$ . A pole-zero cancellation occurs at 1 rad/s in the subsidiary system.

**Case 2—Added Damping.** Damping can be added to the system by a further modification to reduce the “sharpness” of the two poles on either side of the zero assigned in case 1. Of course the damping does not have to be applied at the same coordinate as the stiffness modification  $k^*$ . Figure 4 shows the effect of a grounded dashpot,  $c^* = 0.2$ , at coordinate 3. In Fig. 5 the same dashpot is applied at coordinate 4. The effect of the two modifications on the 3rd and 4th poles is similar (the 2nd pole merges with the zero at 1 rad/s when damping is added) but when the index of the damping coordinate is either  $p$  or  $q$  (4 or 5 in this particular example) the zeros are unaffected and remain “sharp.” This effect can be seen in Fig. 5. The damper may be applied between two coordinates instead of being grounded. The case of a connection between coordinates 2 and 3 is shown in Fig. 6 and coordinates 4 and 5 in Fig. 7.

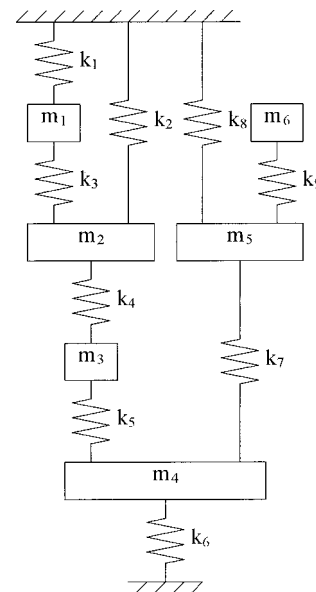


Fig. 1 Six-degree-of-freedom spring-mass system

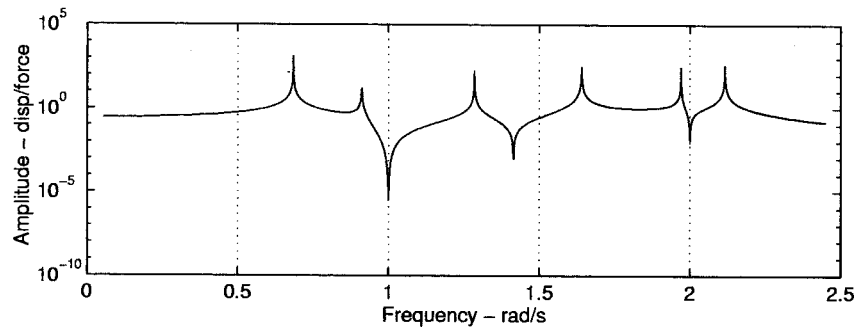


Fig. 2 Undamped receptance  $h_{45}$

Table 1 Poles and zeros of  $h_{45}$

Poles (rad/s)	Zeros (rad/s)
0.684	1.000
0.911	1.000
1.286	1.414
1.640	2.000
1.970	$\infty$
2.117	

Table 2 Zeros of  $h_{45}$  for the modified system

Zeros (rad/s)
1.000
1.058
1.500
2.227
$\infty$

**Case 3—Damped System.** Proportional damping  $C=0.05 K$  is applied and a modification  $b^*=0.4646-i 0.2584$  at coordinate 3 is determined for the assignment of a zero at 1.5 rad/s. It is undesirable to make a negative-damping modification because it would need to be applied actively and could destabilize the system. Therefore a stiffness modification  $k^*=0.4646$  is applied without damping. The receptances before and after the modification are shown in Figs. 8 and 9. It is clear from Fig. 9 that the stiffness modification results in a complex zero with an imaginary part close to 1.5 rad/s.

### Experimental Example

A physical experiment was carried out using a steel beam in free-free and clamped-free (cantilever) configurations. The length of the beam was 1.6 m with a rectangular cross section having the nominal dimensions of 2 cm breadth by 1 cm depth. Modal tests were carried out on transverse vibrations in the flexible direction.

**Free-Free Beam Experiments.** In the first case a zero was assigned at 500 rad/s to a point receptance at a location 1.2 m (coordinate 13) from the end of the beam where an added mass was applied (coordinate 1). The point receptance of the unmodi-

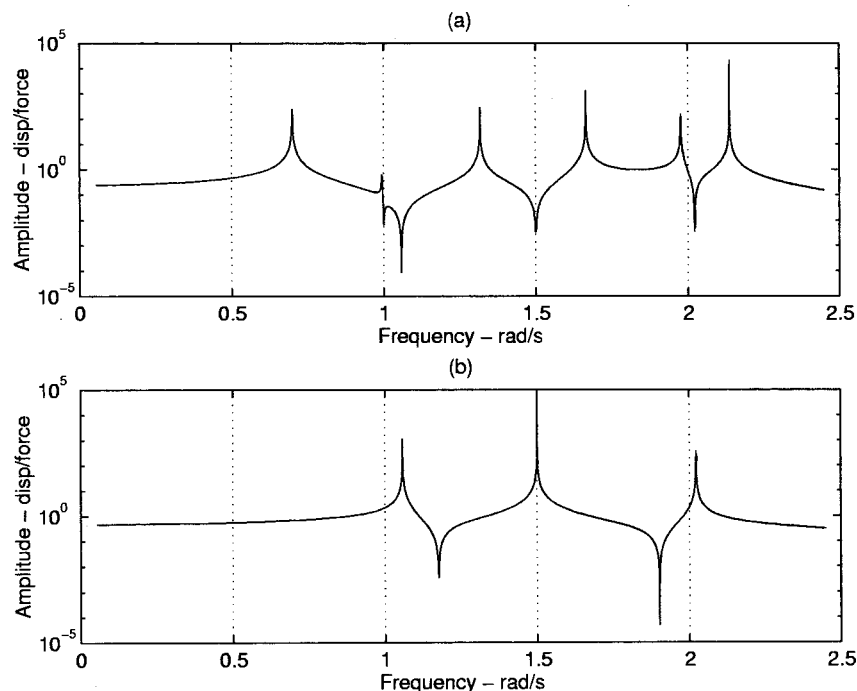


Fig. 3 (a) Undamped receptance with an assigned zero at 1.5 rad/s; (b) undamped receptance for the subsidiary system of the zeros of  $h_{45}$

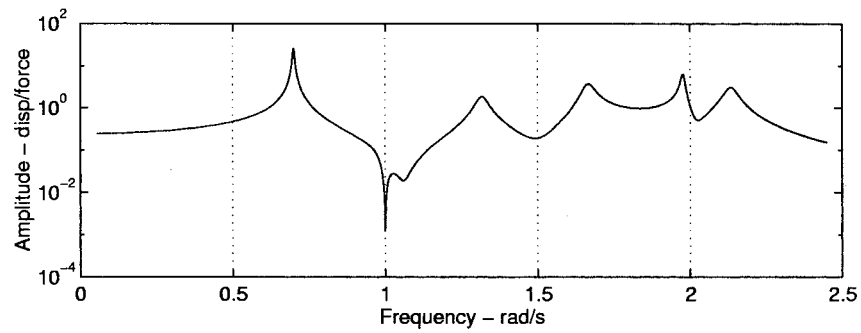


Fig. 4 Damped receptance—grounded dashpot at coordinate 3

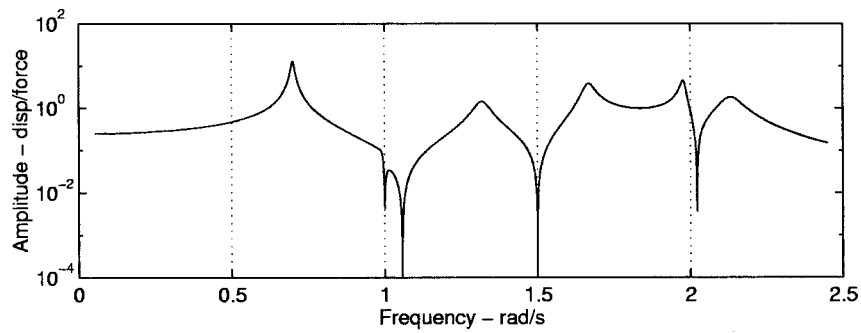


Fig. 5 Damped receptance—grounded dashpot at coordinate 4

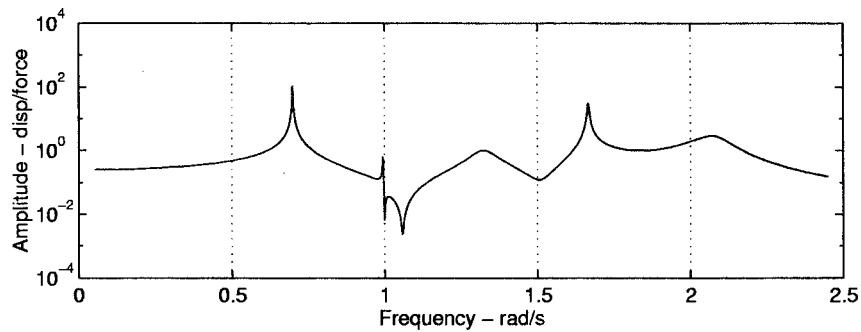


Fig. 6 Damped receptance-dashpot connected between coordinates 2 and 3

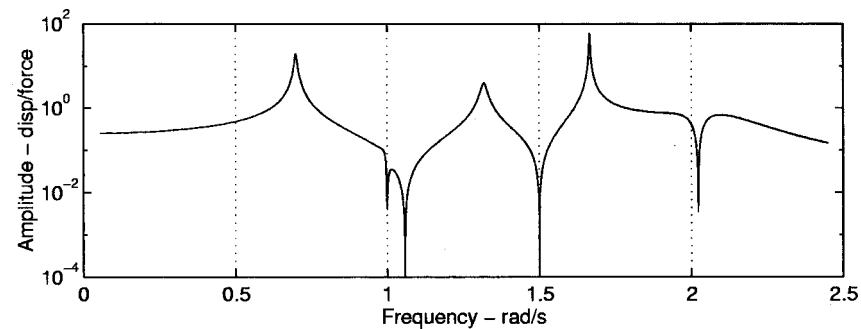


Fig. 7 Damped receptance-dashpot connected between coordinates 4 and 5

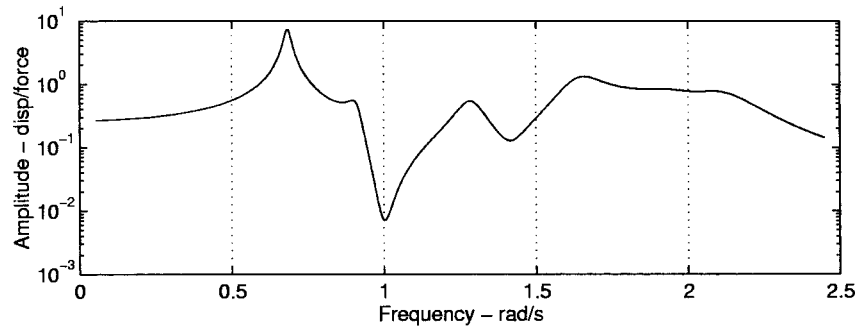


Fig. 8 Receptance  $h_{45}$  of the damped system

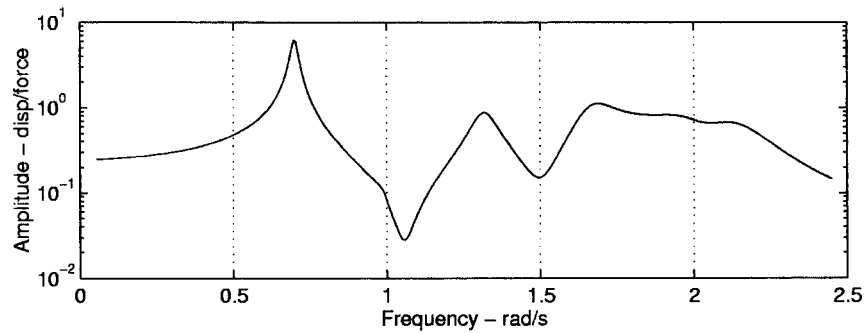


Fig. 9 Receptance of the damped system following modification by a grounded spring at coordinate 3

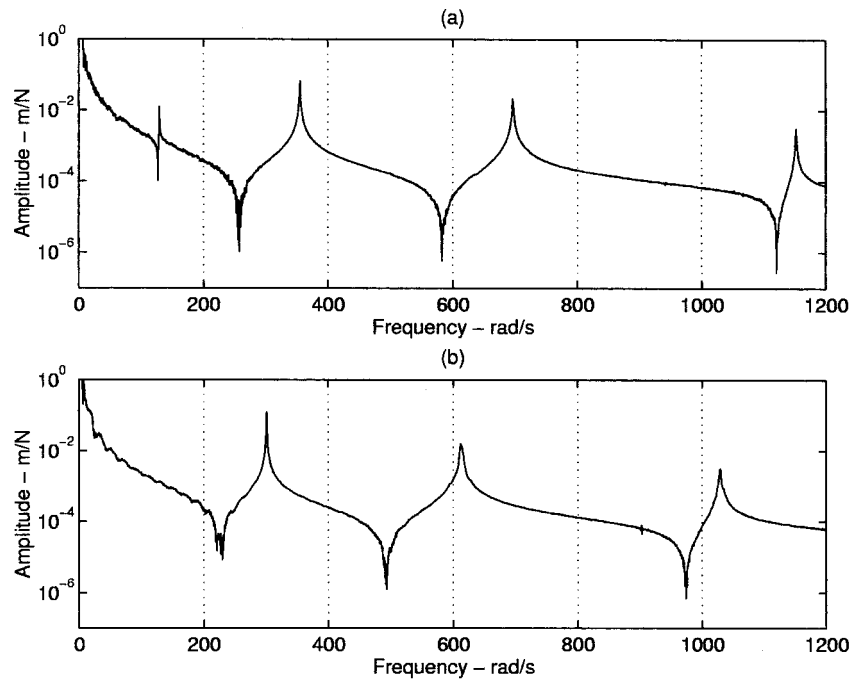


Fig. 10 Measured receptances  $h_{13,13}$  from the free-free beam; (a) before, modification, (b) after modification

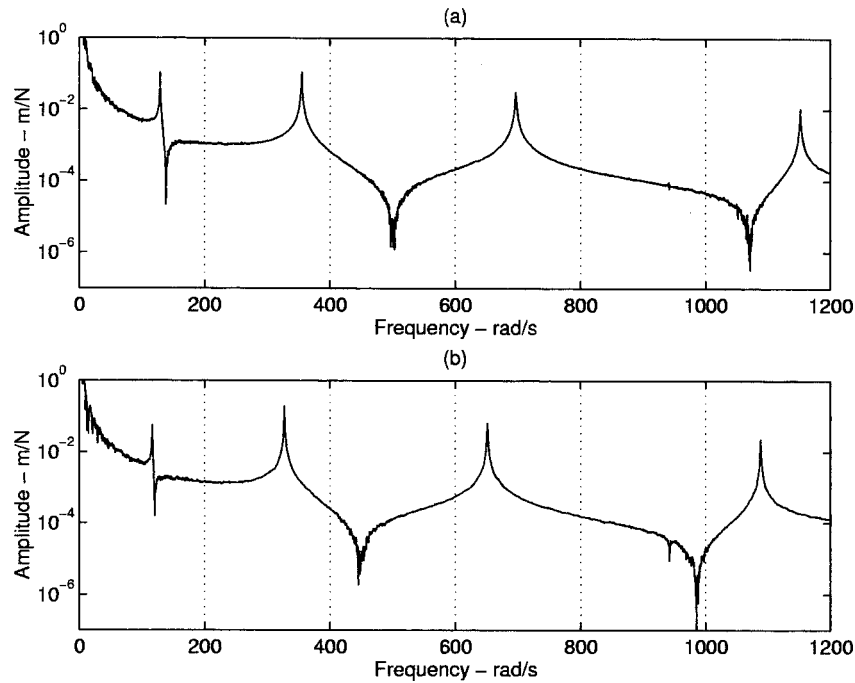
fied system is shown in Fig. 10(a) where it is seen that the zero closest to the desired 500 rad/s occurs at around 590 rad/sec. It is clear from Eqs. (22), (29), and (30) that measurements of  $h_{1,1}$ ,  $h_{1,13}$  and  $h_{13,13}$  at 500 rad/sec are needed in order to determine the magnitude of the modification. Of course, the measurements contained noise which was significant in the imaginary parts of the receptances because of very small damping present in the beam. A

fourth-order polynomial was fitted to the measured frequency responses over a range of frequencies from 400–550 rad/s which resulted in the following estimates at 500 rad/s:

$$h_{1,1} = -1.5477E-05 - 2.2974E-07i,$$

$$h_{1,13} = -8.7217E-06 - 7.2866E-08i,$$





**Fig. 11 Measured receptances  $h_{17,13}$  from the free-free beam; (a) before modification, (b) after modification**

$$h_{13,13} = -4.0974E-06 - 5.0385E-08i.$$

A modification consisting of a mass of 1.2926 kg together with a dashpot of  $-30.9432$  Ns/m was determined for the assignment of the zero to 500 rad/s. The calculated mass was then attached at the end of the beam. No attempt was made to apply the negative damper, so that the zero of the modified system would be expected to lie very slightly away from the imaginary axis of the eigenvalue plane. The point receptance  $h_{13,13}$  following mass modification of the system is shown in Fig. 10(b) where it is indeed clear that the zero has been successfully relocated close to 500 rad/s. High resolution of the frequency data achieved by “zooming” shows the zero to be at 494 rad/s.

In a second study using the free-free configuration a zero of  $h_{17,13}$  was assigned to 450 rad/s by means of a mass modification at coordinate 1. Coordinate 17 was located at the other end of the beam. The following measurements were obtained from smoothed data at 450 rad/s:

$$h_{1,1} = -2.5655E-05 - 4.0148E-07i,$$

$$h_{1,13} = -1.1752E-05 - 1.3706E-07i,$$

$$h_{1,17} = 1.7008E-05 + 2.6705E-07i,$$

$$h_{17,13} = 3.9734E-06 + 1.0998E-07i,$$

The computed mass and damping modifications were 0.2004 kg and  $-1.5260$  Ns/m, and a physical mass modification of 0.2 kg was applied. Figure 11 shows the cross receptance  $h_{17,13}$  before and after application of the added mass. In the initial (unmodified) condition the zero is at around 500 rad/sec. High-resolution frequency data shows the zero at about 446 rad/s following modification by the added mass.

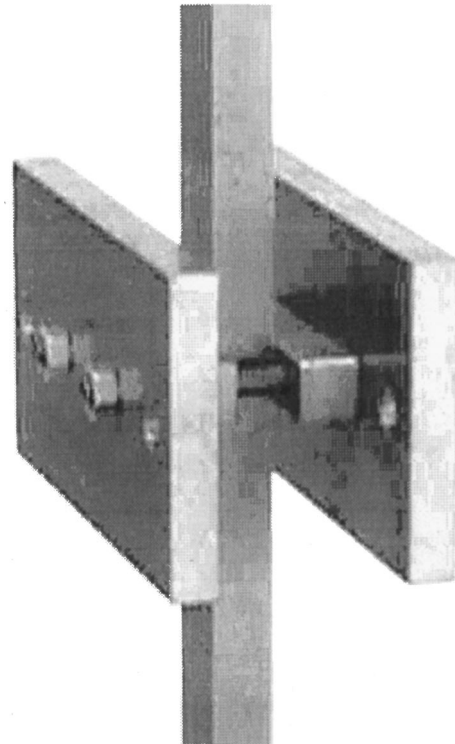
**Clamped-Free Beam Experiment.** A zero of  $h_{13,13}$  was assigned to 500 rad/s by means of a mass modification at coordinate 9 (at midspan). The following smoothed measurements were obtained at 500 rad/s:

$$h_{9,9} = -3.2156E-07 - 1.2307E-08i,$$

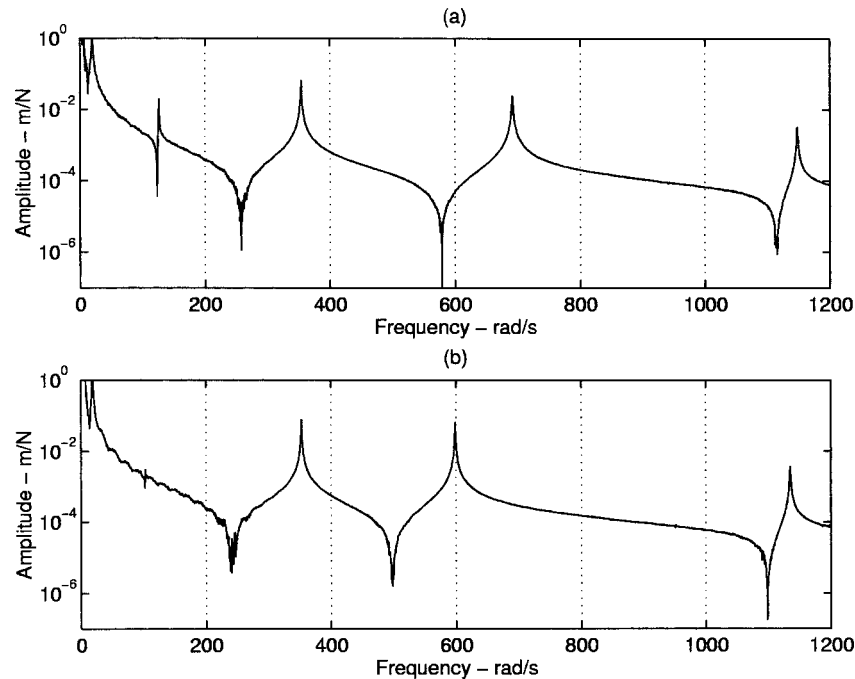
$$h_{9,13} = -4.8480E-06 - 1.1843E-08i,$$

$$h_{13,13} = -3.8839E-06 - 3.3760E-08i.$$

The computed mass and damping modifications were 0.6981 kg and  $-2.1528$  Ns/m. A mass of 0.698 kg was applied and its at-



**Fig. 12 Beam with added mass**



**Fig. 13 Measured receptances,  $h_{13,13}$  from the clamped-free beam; (a) before modification, (b) after modification**

tachment to the beam is shown in Fig. 12. A small rotary inertia is introduced by the physical application of the mass but it has a very small effect as shown in Fig. 13 where the zero, initially at around 585 rad/s, is moved to 498 rad/sec (from high resolution measurements) by the application of the added mass.

## Conclusions

A method is proposed for the assignment of zeros in point and cross receptances by passive stiffness, damping, and mass modifications. The frequency responses of the zeros eigenvalue problem are determined from the frequency responses of the complete system and used to determine the magnitude of the modification. The method is demonstrated in simulated experiments to assign the zeros of point and cross receptances in a physical beam by adding point masses.

## Acknowledgment

Mr. Simon James carried out the modal tests on the beam.

## References

- [1] Baldwin, J. F., and Hutton, S. G., 1985, "Natural modes of modified structures," *AIAA J.*, **23**, pp. 1737–1743.
- [2] Ram, Y. M., 2000, "Dynamic Structural Modification," *Sound Vib Dig.*, **32**, No. 1, pp. 11–18.
- [3] Weissenburger, J. T., 1966, "The Effect of Local Modifications on the Eigenvalues and Eigenvectors of Linear Systems," Ph.D. thesis, Washington University.
- [4] Pomazal, R. J., and Snyder, V. C., 1971, "Local Modifications of Damped Linear Systems," *AIAA J.*, **9**, pp. 2216–2221.
- [5] Mottershead, J. E., and Lallement, G., 1999, "Vibration Nodes, and the Cancellation of Poles and Zeros by Unit-Rank Modifications to Structures," *J. Sound Vib.*, **222**, No. 5, pp. 833–851.
- [6] Bucher, I., and Braun, S., 1993, "The Structural Modification Inverse Problem: An Exact Solution," *Mech. Syst. Signal Process.*, **7**, No. 3, pp. 217–238.
- [7] Bucher, I., and Braun, S., 1997, "Left Eigenvectors: Extraction From Measurements and Physical Interpretation," *ASME J. Appl. Mech.*, **64**, pp. 97–105.
- [8] Tsuei, Y. G., and Yee, E. K. L., 1989, "A Method for Modifying Dynamic Properties of Undamped Mechanical Systems," *ASME J. Dyn. Syst., Meas., Control*, **111**, pp. 403–408.
- [9] Li, T., He, J., and Sek, M., 2001, "Local and Global Pole-Zero Cancellation of Mass-Spring Systems," *Mech. Syst. Signal Process.*, **15**, No. 1, pp. 121–127.
- [10] Mottershead, J. E., Mares, C., and Friswell, M. I., 2001, "An Inverse Method for the Assignment of Vibration Nodes," *Mech. Syst. Signal Process.*, **15**, No. 1, pp. 87–100.
- [11] Inman, D. J., 1993, *Engineering Vibrations*, Prentice-Hall, Englewood Cliffs, NJ.
- [12] Ram, Y. M., and Elhay, S., 1996, "The Theory of a Multi-Degree-of-Freedom Vibration Absorber," *J. Sound Vib.*, **195**, No. 4, pp. 607–615.
- [13] Sun, J. Q., Jolly, M. R., and Norris, M. A., 1995, "Passive, Adaptive and Active Tuned Vibration Absorbers—A Survey," *ASME J. Vib. Acoust.*, **117**, No. 8, pp. 234–242.
- [14] Mottershead, J. E., 1998, "On the Zeros of Structural Frequency Response Functions and Their Sensitivities," *Mech. Syst. Signal Process.*, **12**, pp. 591–598.
- [15] Mottershead, J. E., 2001, "Complex and Defective Zeros in Cross Receptances," *J. Sound Vib.*, in press.
- [16] Kailath, T., 1980, *Linear Systems*, Prentice-Hall, Englewood Cliffs, NJ.
- [17] Zhang, Q., and Lallement, G., 1989, "Selective Structural Modifications: Applications to the Problems of Eigensolution Sensitivity and Model Adjustment," *Mech. Syst. Signal Process.*, **3**, pp. 55–69.
- [18] Singh, K. V., and Ram, Y. M., 2000, "Dynamic Absorption by Passive and Active Control," *ASME J. Vib. Acoust.*, **122**, No. 4, pp. 429–433.

# On Some Issues in Shakedown Analysis

G. Maier

Professor,

Department of Structural Engineering,  
Technical University (Politecnico),

Piazza L. da Vinci, 32,  
I-20133 Milan, Italy

Fellow ASME

*Shakedown analysis, and its more classical special case of limit analysis, basically consists of "direct" (as distinct from time-stepping) methods apt to assess safety factors for variable repeated external actions and procedures which provide upper bounds on history-dependent quantities. The issues reviewed and briefly discussed herein are: some recent engineering-oriented and cost-effective methods resting on Koiter's kinematic theorem and applied to periodic heterogeneous media; recent extensions (after the earlier ones to dynamics and creep) to another area characterized by time derivatives, namely poroplasticity of fluid-saturated porous media. Links with some classical or more consolidated direct methods are pointed out. [DOI: 10.1115/1.1379368]*

## 1 Introduction

The main lasting contributions of Warner Koiter to mechanics are his elastic instability theory which explained post-buckling behavior and imperfection sensitivity, and the kinematic theorems concerning the "adaptation" or "shakedown" of elastic-plastic solids under fluctuating external actions. The two preceding Koiter Lectures have been in various ways related to the former, more monumental and elaborate area of Koiter's contributions. This one is related to, and inspired by, the latter area of his work.

The inelastic behavior of structures subjected to repeated variable loading, and possible consequent structural failures, have attracted the attention of researchers in engineering mechanics such as Hans Bleich and Ernst Melan, before plastic collapse of ductile structures (under loads assumed as monotonically amplified in time) was given a definitive theoretical and methodological basis by Gvozdev, Prager, Drucker, Hill, and other founders of structural plasticity. The pioneering works appeared in the 1930s and those developed in the 1950s by Neal, Symonds, Koiter, and others were primarily aimed at establishing methods apt to provide the critical "live load" multiplier, usually called "safety factor," beyond which failure occurs by either incremental collapse or alternating plasticity, and below which the structure "shakes down," in the sense that plastic yielding ceases after a while (or, more precisely, the dissipated energy is bounded in time). These methods of shakedown analysis, like the more classical ones of rigid-plastic limit analysis (with respect to plastic collapse under monotonic loading), are frequently called "direct" or "simplified" methods, as distinct from the step-by-step procedures of inelastic analysis along the whole time-history of loading.

Solution techniques marching in time have been for decades, and to some extent still are, in the main stream of research in computational mechanics of solids and structures. However, in favor of direct methods from a practical standpoint there are meaningful circumstances like the following ones.

(a) In many engineering cases, reliable information on variable external actions concern not their time-history, but merely ranges within which they vary and are expected to show up again and again.

(b) Numerical solutions to practical nonlinear initial boundary value problems of inelastic structural analysis were prohibitive before the advent of high-performance computers and profes-

sional software, and are still expensive and discouraging in a number of engineering situations, especially for repeated parametric studies often needed in preliminary structural design.

(c) Direct methods of shakedown analysis conceptually (though not historically) represent generalizations of limit analysis, which is rooted in the remote origin of structural mechanics and engineering. This generalization, to a large extent, preserves the main features of limit analysis, namely its simplicity, elegance, and mathematical formalism (optimality properties, duality pattern) and its practical appeal of providing information on the carrying capacity which are essential for a physical insight into the structural system and for its design.

Despite the significant developments in the methodology and computational tools of time-stepping inelastic analysis occurred in the last few decades, in view of the above motivations (a)–(c), direct methods for ultimate limit-state analysis can still be regarded as competitive in a variety of engineering situations. Some of the original severe limitations (such as the restrictions to associative perfect plasticity and to infinitesimal deformations) have been, and are further being, relaxed. Remarkable progresses have been made also in the area of those direct methods which are intended to provide bounds on history-dependent quantities. All these developments have been fostered by, and fruitfully combined with, parallel developments in discretization techniques (particularly by finite elements and boundary elements) and in inequality-constrained mathematical optimization in vector spaces (i.e., in mathematical programming).

The theories and methods for direct analysis of structures under repeated variable external actions (or, as a special case, under monotonic loading, as dealt with by limit analysis), and their origin and growth, have been surveyed at diverse stages in various publications after Koiter's 1960 essay ([1]): e.g., in the "ad hoc" textbooks by König [2] and Kamenjarzh [3] and in a number of general and specific publications on plasticity, (e.g., [4–6]). An up-to-date abundant, though incomplete, bibliography can be found in a recent survey article ([7]).

Here only a subdomain of the shakedown analysis area will be considered and briefly discussed. This subdomain is believed to be representative of several current research topics in the area. Sections 3 and 4 are devoted to brief discussions of shakedown theory and bounding techniques, respectively, generalized to poroplasticity; these generalizations exhibit both similar and distinct features with respect to the extensions started long ago to dynamics, namely to other situations characterized by the presence and essential role of time derivatives in the governing field equations. Preliminarily, as a homage to Warner Koiter, some recent applications of his kinematic shakedown theorem will be reviewed in Section 2. In the writer's opinion these three topics are at present

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, February 28, 2001. Editor: L. T. Wheeler. Discussion on the paper should be addressed to the 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.

among those open to and worthy of further research, with expected new results of use in structural and materials engineering.

## 2 Application of Koiter Theorem to Heterogeneous Periodic Solids

The following citations (verbatim, to within slight notation changes) from the conspectus of plasticity theory W. T. Koiter published in 1960 ([1]) may provide a concise, still valuable basis to shakedown (SD) analysis by a kinematic approach ( $\tau$  denotes time,  $V$  volume of the body,  $S_i$  its unconstrained boundary).

(I) "The concept of an admissible plastic strain rate cycle  $\dot{\varepsilon}_{ij0}^p(\tau)$  is defined by its characteristic property that the increments of the plastic strains in such a cycle for some time interval  $T$  constitute a kinematically admissible strain distribution."

(II) "The body will not shakedown, i.e., it will fail ultimately by cyclic plastic deformations, if any admissible plastic strain rate cycle  $\dot{\varepsilon}_{ij0}^p(\tau)$  and any external loads  $b_i(\tau), t_i(\tau)$  within the prescribed limit can be found for which

$$\int_T d\tau \left\{ \int_V b_i \dot{u}_{i0}^p dV + \int_{S_i} t_i \dot{u}_{i0}^p dS \right\} > \int_0^T d\tau \int_V F(\dot{\varepsilon}_{ij0}^p) dV \quad (2.1)$$

where  $F(\dot{\varepsilon}_{ij0}^p)$  is the plastic energy dissipation function in the strain rate cycle  $\dot{\varepsilon}_{ij0}^p(\tau)$ ."

(III) "The structure will shakedown if a number  $\mu > 1$  can be found with the property that, for all admissible plastic strain rate cycles  $\dot{\varepsilon}_{ij0}^p(\tau)$  and all external loads  $b_i(\tau), t_i(\tau)$  within the prescribed limits,

$$\mu \int_0^T d\tau \left\{ \int_V b_i \dot{u}_{i0}^p dV + \int_{S_i} t_i \dot{u}_{i0}^p dS \right\} < \int_0^T d\tau \int_V F(\dot{\varepsilon}_{ij0}^p) dV. \quad (2.2)$$

The upper bound of such number  $\mu$  is then obviously the factor of safety with respect to shakedown."

(IV) "An initial hope that the second shakedown theorem might facilitate shakedown analysis has not been confirmed by practical applications" (p. 108 of [1]).

The above "second theorem" consists of statements (II) and (III), both based on the concept defined at point (I). Koiter had formulated this theorem in a note ([8]) communicated by C. B. Biezeno, his mentor, to the Dutch Royal Academy at the meeting of Dec. 17, 1955, defining it as "a new theorem, complementary to Melan's" and apt to "be combined with Melan's in order to obtain upper and lower bounds for the allowable load variations for shakedown." A nontrivial historical scrutiny (far beyond the present purposes) might trace its close roots in the work done by Neal and Symonds at Brown University in the early 1950s (as pointed out by a recent paper by Symonds in [6]), and its remote roots in the kinematic ("unsafe," "upper bound," "dual") limit theorem as a special case, and hence, to Galileo's anticipation of it in his last "Dialogue."

Ductile periodic systems such as perforated steel plates in powerplants and metal-matrix composites adopted in aerospace, and other, engineering situations, have been investigated from a limit and SD analysis standpoint in several publications (referred to, e.g., in [7]). Noteworthy applications and extensions of Koiter theorems have been studied (e.g., in [9,10]) for temperature-dependent material models in view of high temperature industrial situations, which have fostered also direct methods for the creep range, e.g., [11,12]. Here focus is on application of the above statements (I)–(III), by a procedure which, in the writer's experience ([13–16]), did recently confirm the "initial hope" (IV) expressed by W. Koiter 40 years ago. This particular procedure is briefly outlined below.

(i) In the space of average (or "macroscopic") stresses  $\Sigma_{ij}$ , which are regarded as "loads," the given region  $\Omega$  where they fluctuate, and which represents the "loading domain," is approxi-

mated by a (hyper)polyhedron  $\bar{\Omega}$ . For the loading conditions corresponding to all the  $v$  vertices of this  $\sigma_{ij}^{ek}$  (say  $\Sigma_{ij}^k, k=1 \dots v$ ) polyhedron, the relevant linear-elastic response in terms of local "microscopic" stress field  $\sigma_{ij}^{ek}(x_r)$  over the representative volume  $V$  are preliminarily computed by a conventional homogenization technique and finite element procedure ( $x_r$  being the Cartesian coordinates).

(ii) The kinematic admissibility required by the definition (I) in the presence of texture periodicity, implies that the cumulative plastic strain field  $\varepsilon_{ij0}^p(x_r)$  at the end of any cycle must be enforced to satisfy the conditions

$$\varepsilon_{ij0}^p(x_r) = \varepsilon_{ij0}^p + \frac{1}{2} \left( \frac{\partial}{\partial x_i} \bar{u}_{j0} + \frac{\partial}{\partial x_j} \bar{u}_{i0} \right), \quad x_r \in V, \quad \bar{u}_{i0}(x_r) \text{ periodic on } S \quad (2.3)$$

where subscript 0 marks variables belonging to an admissible plastic strain rate cycle (briefly "admissible cycle" AC) in the sense specified by (I);  $\Delta \bar{u}_i$  represents a displacement field over the representative volume  $V$  obeying periodicity conditions to be imposed on the boundary  $S$  of  $V$ ;  $\Delta \varepsilon_{ij}^p$  are average (macroscopic) strains constant over  $V$ .

(iii) The sufficient (II) and necessary (III) inadaptation conditions combined, as usual in SD analysis, reduce the search for the SD limit or safety factor  $s$  to the minimization of the dissipated energy, say  $D$ , cumulative in space and time along the admissible cycle, after the normalization which sets equal to 1 the external work done by the loads because of the AC considered. This work can be expressed in two forms:

$$D = \int_0^T d\tau \int_V \sigma_{ij}^{ek} \dot{\varepsilon}_{ij0}^p dV = \int_V \sum_{k=1}^v \sigma_{ij}^{ek}(x_r) \Delta \varepsilon_{ij}^{pk}(x_r) dV = 1. \quad (2.4)$$

Because of the virtual work principle, the former expression (where the average stress input is reflected by the linear elastic stress response to it,  $\sigma_{ij}^{ek}$ , over  $V$ ) equals the external works due to the AC. The latter expression, where the time integral becomes a sum, is justified by a SD theorem which ensures that shakedown occurs under a (hyper) polyhedral load domain if it does under the sequence of loads corresponding to its  $v$  vertices ([17]) and, hence, permits to materialize any admissible cycle in the sense of (a) as a sequence of increments  $\Delta \varepsilon_{ij}^{pk}(x_r), k=1 \dots v$ . Consistently with the adopted finite element model, Gauss integration reduces to a sum the space integrals in Eq. (2.4).

(iv) The minimization of the dissipated energy  $D$  has to be performed over the set of all admissible cycles, namely with respect to: the displacement vector  $\bar{U}$  governing the modeled field  $\bar{u}_{i0}(x_r)$ , account taken of the periodicity on  $S$ ; the overall strain tensor  $\varepsilon_{ij0}^p$ ; the cycle increments  $\Delta \varepsilon_{ij0}^{pk}(x_r^h), k=1 \dots v$ , the new index  $h$  running over the set of Gauss points in the whole volume  $V$ . The constraints to be enforced result from Eqs. (2.3) and (2.4), suitably algebraized according to the space discretization adopted.

(v) The mathematical features of the minimization problem arrived at depend on the constitutive models locally attributed to the (individually homogeneous) constituents in the representative volume  $V$ . If Mises-Huber associative perfect plasticity is assumed for all the phases, then the global dissipated energy turns out to be a convex function, nondifferentiable ("nonsmooth") in the origin; the plastic incompressibility (entailed by the normality in flow rules) is expressed by a linear equality constraints in every Gauss point  $h$  and for all steps  $k$  in the admissible cycle.

(vi) The mathematical programming problem formulated by the above path, can be numerically solved by a variety of algorithms now available as commercial software. The one "ad hoc" originated by Y. G. Zhang [18] and adopted in [13–15] rests on penalization for plastic incompressibility and on Lagrangian multipliers for the other constraints. This algorithm solves the nonlin-



ear stationarity equations for the Lagrangian by an iterative procedure, at each step of which linear equations are solved by using results from the preceding iteration and user-available factors for fixing denominators proportional to the current local dissipation.

The above outlined procedure can provide various kinds of technical information, some examples of which are shown in Figs. 1 and 2, for a perforated steel plate interpreted as a plane stress system and for a ductile composite, respectively, with Mises-Huber perfectly plastic materials (in the latter case the yield stress being  $\sigma_0^m$  and  $\sigma_0^f = 8.7\sigma_0^m$ , for matrix and fibers, respectively). Although it rests on Koiter's classical statements (II) and (III), this direct method exhibits various aspects (pointed out below) which appear to be worth noting at the present development stage of SD analysis and are in parts distinct from their counterparts in other kinematic approaches to similar kinds of engineering problems (e.g., see, [19–22]).

(a) The heterogeneity periodicity, dealt with in the spirit of homogenization theory, is accounted for both in the preliminary linear elastic computations of  $\sigma_{ij}^{ek}$  and in the search for the SD limit  $s$ . It plays a twofold role: first, in terms of the boundary conditions on the representative volume; second, through Eq. (3), at the end of the admissible cycles according to definition (I). Computational savings are obviously implied by the formulation of the SD analysis on the (minimum) representative volume, singled out on a statistical basis or in view of the strict microstructural periodicity assumed here.

(b) The penalty approach to the enforcement of the plastic incompressibility constraint (at each step  $k$  of any admissible cycle) can be regarded primarily as a provision apt to avoid or attenuate “locking” manifestations. From this standpoint, it represents an alternative to reduced integration, and to mixed finite element modeling (see, e.g., [16]) and, in particular, to multifield modeling in Prager's generalized variables considered herein in Section 3. Locking due to the combination of constitutive and modeling kinematic constraints, and its unconservative consequences, were met early in computational limit analysis, but, clearly, are expected to occur in shakedown analysis as well. Penalization can merely be interpreted as a way to enforce constitutive kinematics

in a “soft” average way. The pros are: compatibility with traditional simple FE methods centered on displacement modeling; no increase in the optimization variables (contrary to the implications of the Lagrange multiplier approach). The main contra of the penalty method is the empiricism underlying the choice of the penalty factors or factor  $f$  (a single factor for an overall violation norm turned out to be suitable, in the writer's experience). As expected and as visualized in Fig. 3, a very high  $f$  leads to either locking manifestations or numerical instability, a very low  $f$  to kinematic relaxation and consequent conservative but erroneous SD limits ([13]). The range of penalty factors leading to reasonably accurate solutions can be determined only empirically, by comparative numerical tests, and depends primarily on the tackled problem and the adopted material model (e.g., for the same Prandtl's indentation problem, it turned out to cover various orders of magnitude with Mises models, much less with Drucker-Prager's and Pande-Zienkiewicz's models for frictional materials). Fortunately for industrial applications, the suitable range can be estimated once for all for each kind of problems (e.g., hexagonal representative volume of fiber-reinforced composites ([14,15]) defective pipelines ([13])), since it is fairly insensitive to details such as geometry and finite element mesh.

(c) The iterative solution technique above mentioned at stage (vi) involves (at least two) further parameters to be assumed on empirical basis (one for selecting at each iteration the yielding Gauss points, the other for replacing a vanishing denominator in the nonyielding points). After tuning the procedure also as for these tolerances, fast convergence was observed, though not theoretically corroborated by proof.

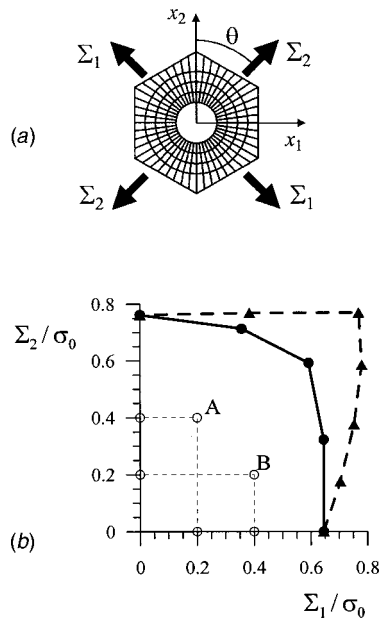
(d) The validation of the particular kinematic method consisting of phases (i)–(vi) has been achieved by comparisons with earlier results by diverse approaches (e.g., [23–25]) and with computations carried out step by step by means of a commercial nonlinear finite element code, with reference to defective pipelines ([13,26]), and to heterogeneous solids like perforated plane-stress plates and fiber-reinforced composites (see [14,27] and illustrative examples of Figs. 1–3).

(e) Computing costs of the direct kinematic method, with respect to those of time-marching inelastic analysis, were found to be reduced by a small factor (say 2 or 3) for the special case of limit analysis, but by one or two orders of magnitude for SD analysis, in view of the trial-and-error strategy required to evaluate SD limits by step-by-step computations. The above particular kinematic SD analysis method was ascertained to be cost-effective in engineering practice, but not necessarily superior to other recently proposed direct methods. In particular, e.g., the direct iterative method consisting of repeated linear analyses with suitable modified elastic moduli (see, e.g., [28,29]) exhibits the theoretical appeal of a convergence proof and the practical advantage of using commercial general-purpose computer codes.

As a conclusion of this section, it can probably be said that W. Koiter's hope expressed in (IV) is confirmed nowadays, at least in view of the direct method outlined in what precedes, but probably also in view of others, among the various methods devised in the last decade.

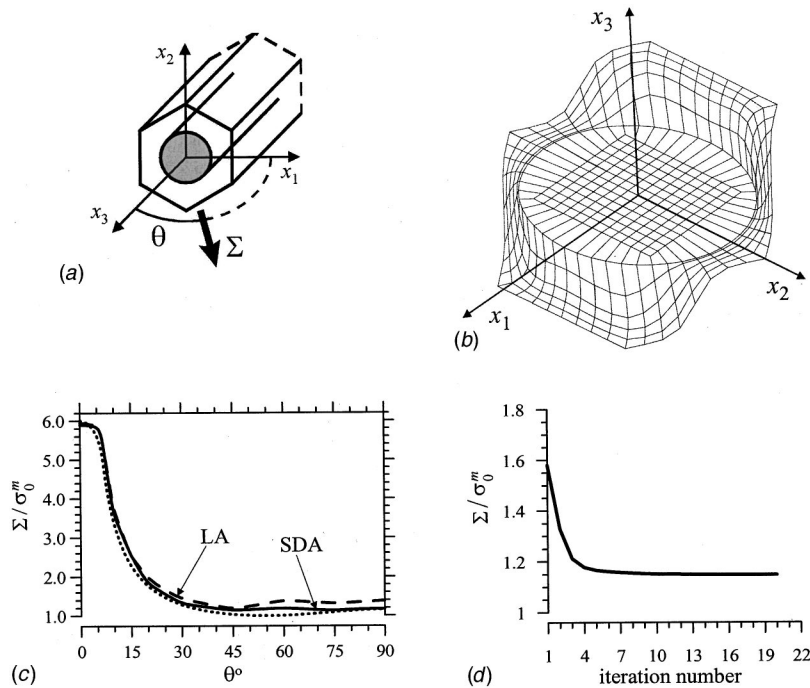
The spreading applications of SD concepts to heterogeneous solids and in particular to periodic composites as seen above, might be beneficial to other areas of materials mechanics, where multiscale approaches have become fashionable and promising, “in primis” to fatigue criteria.

Daniel Drucker wrote in 1963 ([30]): “When applied to the microstructure there is a hope that the concepts of endurance limit and shakedown are related, and that fatigue failure can be related to energy dissipated in idealized material when shakedown does not occur.” Recent research works by Dang Van and others [31] seem to confirm this far-sighted thought. In fact, connections have been established between fatigue crack advancement at the mac-



**Fig. 1 Shakedown analysis of a perforated plate: (a) representative volume and finite element mesh; (b) shakedown limit locus (solid line) in the average stress plane for rectangular loading domains like those defined by points A and B; for comparison, the plastic collapse locus in dashed lines ( $\sigma_0$  being the material yield stress)**





**Fig. 2 Shakedown analysis of a metal-matrix composite subjected to uniaxial average stress  $\Sigma$ : (a) representative volume; (b) incremental collapse mechanism for  $\theta=30$  deg; (c) shakedown limit (solid line SDA) versus plastic collapse limit evaluated by the present kinematic method (dashed line LA) and by the static method in [71] (dotted line) ( $\sigma_0^m$  being the matrix yield stress); (d) convergence on the shakedown limit in the former procedure with penalty factor  $10^6$**

rosopic level on one side and lack of shakedown at the microscopic level on the other, through appropriate microstructural modeling of polycrystal metals.

Material ratchetting, in particular of metals under severe cyclic thermal loading, is also being successfully interpreted in the light of shakedown theory at the microscale ([32,33]). This can be regarded as another recent case of the scale interaction advocated about 40 years ago in a further, more general sentence by Drucker: “Phenomenological theory never really answers a question

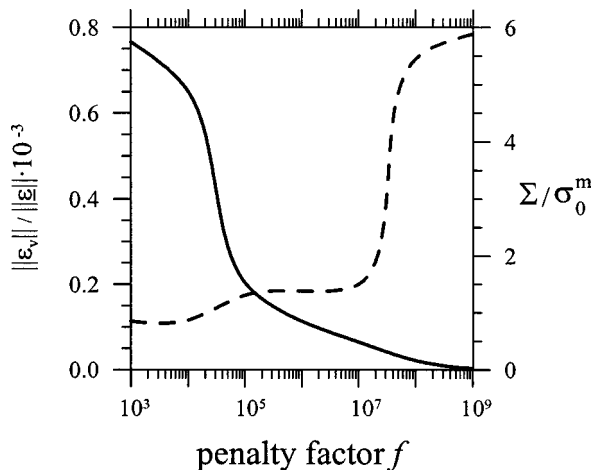
“why”? The microscopic and the macroscopic views are mutually dependent and should interact at all levels” ([30]).

### 3 Shakedown Analysis in Poroplasticity

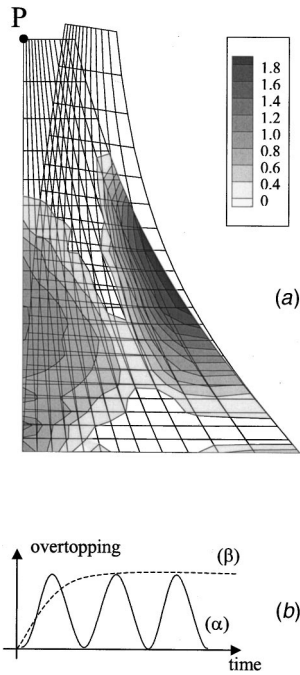
After almost three decades since their extension from quasi-static to dynamic plasticity, in the last few years direct methods have been studied in another context where field equations involve time derivatives, namely in the mechanics of deformable porous media saturated by a viscous liquid, the filtration motion of which interacts with the deformations of the solid skeleton.

The classical and lasting theoretical foundations of coupled (fluid-solid) problems are represented by Biot’s linear poroelasticity theory shaped mainly in the 1940s. Almost two decades earlier pioneering works due to Terzaghi, Fillunger, and others were motivated by geotechnical and dam engineering problems. At present, motivations for further developments in this area arise from a number of diverse technologies such as slope stability in geotechnical engineering; subsidence control in environmental engineering; biomechanics of bones and soft tissues; exploitation of oil and gas deposits; design and rehabilitation of new and existing earth dams; and, in the presence of diffused cracks, also masonry and concrete dams (the illustrative examples of Figs. 4–6 concern two simplified models of gravity dams). Comprehensive treatments of the accumulated knowledge concerning fluid-saturated (fully or partially) linear, and nonlinear inelastic porous media can be found, e.g., in [34], together with time-marching procedures for numerical solutions of poroelastic and poroplastic initial boundary value problems.

Direct methods of poroplastic analysis have been developed in the last few years ([35–37]), so far under the restrictive hypotheses which follow: (a) linear kinematics; (b) quasi-static regime;



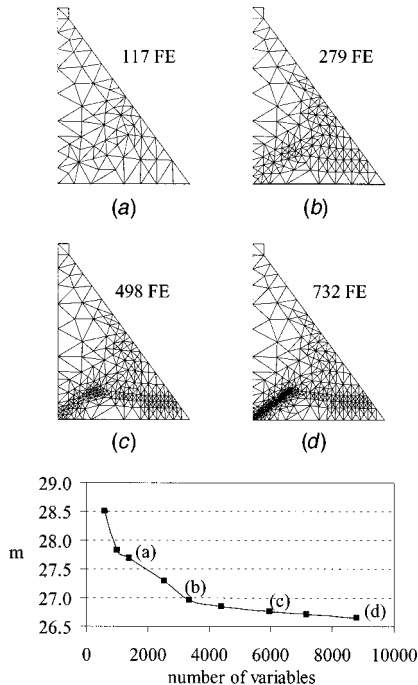
**Fig. 3 Influence of the penalty factor on the plastic collapse limit (dashed line) and on the violation (measured by a norm of the plastic volumetric strain field) of the normality constraint (solid line)**



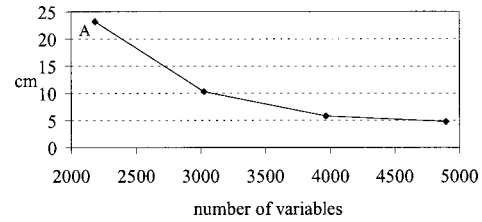
**Fig. 4 Incremental collapse mechanism with relevant Mises-equivalent plastic strain rate field (a) for an idealized gravity dam interpreted as a poroplastic system under periodic live load (bα)**

(c) full saturation of solid skeleton by a single viscous liquid; (d) permeability constant in time; and (e) material stability in Drucker sense, implying associativity of flow rules.

The field equations express equilibrium of “total” stresses  $\sigma_{ij}$  (total in the sense that they concern the two-phase, “bulk” solid), geometric compatibility between strains  $\varepsilon_{ij}$  and displacements  $u_i$



**Fig. 5 Self-adaptive limit analysis governed by a normalized measure of the plastic strain rate density of the collapse mechanism in piecewise-linearized plasticity**



**Fig. 6 Upper bounds on the residual displacement at the top of the idealized poroplastic dam model of Fig. 4, as a function of the mesh refinement (point A corresponds to the finite element mesh shown in Fig. 4)**

of the solid skeleton, conservation of the fluid mass, and Darcy’s filtration law. These field equations involve the variation of fluid content  $\zeta$  and the flux  $q_i$  (defined as accumulated liquid volume per unit bulk volume and, respectively, per unit time and unit crossed area orthogonal to axis  $x_i$ ). Darcy law relates the flux to the gradient of pressure  $p$  through the permeability tensor  $k_{ij} = k_{ji}$ . The external actions include: data on the boundary  $S$  (given displacement on  $S_u$  and traction on  $S_t$ , with  $S_t \cup S_u = S$  and  $S_t \cap S_u = \{\emptyset\}$ ; given pressure on  $S_p$  and flux on  $S_q$ , with  $S_q \cup S_p = S$  and  $S_q \cap S_p = \{\emptyset\}$ ; bulk body forces; and specific weight of the fluid (per unit liquid volume). The initial conditions can merely concern the pressure field over the volume  $V$  of the considered system.

The constitutive law at the macroscopic, phenomenological level, to be combined with the preceding field equations, relates static quantities  $\sigma_{ij}$  and  $p$  to their work-conjugate kinematic ones  $\varepsilon_{ij}$  and  $\zeta$ . Like in elastoplasticity, each one of the latter variables is assumed to be the sum of a reversible, poroelastic addend and an irreversible, poroplastic addend (marked by superscripts  $e$  and  $p$ , respectively). The poroelastic addends are linearly related to the static variables according to Biot’s model, namely  $\{\varepsilon_{ij}^e, \zeta^e\} = C\{\sigma_{ij}, p\}$ ,  $C$  being a (positive-definite, linear, symmetric) algebraic operator which, for isotropic materials, is governed by four material parameters.

The irreversible (“plastic”) addends  $\varepsilon_{ij}^p$  and  $\zeta^p$  are governed by nonholonomic relationships in rates. Like in structural elastoplasticity, see, e.g., [38–40], in order to simplify theoretical and numerical developments for SD analysis, the poroplastic model can be given a “piecewise linear” (PWL) approximation as follows:

$$\varphi_\alpha = N_{ij\alpha}^\sigma \sigma_{ij} + N_{\alpha p}^p p - Y_\alpha \leq 0, \quad Y_\alpha = Y_\alpha^0 + H_{\alpha\beta} \lambda_\beta \quad (3.1)$$

$$\dot{\varepsilon}_{ij}^p = N_{ij\alpha}^\sigma \dot{\lambda}_\alpha, \quad \dot{\zeta}^p = N_{\alpha p}^p \dot{\lambda}_\alpha \quad (3.2)$$

$$\dot{\lambda}_\alpha \geq 0, \quad \varphi_\alpha \dot{\lambda}_\alpha = 0, \quad (\alpha, \beta = 1 \dots n_y) \quad (3.3)$$

where the quantities denoted by  $N$ ,  $Y^0$ , and  $H$  are constants and  $Y_\alpha^0$  represents the (positive) initial “yield limit” of the  $\alpha$ th mode. Interpreted in the  $\{\sigma_{ij}, p\}$  and  $\{\varepsilon_{ij}^p, \zeta^p\}$  spaces superposed, the poroelastic domain is the convex (hyper) polyhedron consisting of the intersection of the  $n_y$  half-spaces  $\varphi_\alpha \leq 0$ . Each yield plane  $\varphi_\alpha = 0$  ( $\alpha = 1 \dots n_y$ ) is defined by its unit normal  $\{N_{ij\alpha}^\sigma, N_{\alpha p}^p\}$  and by its distance from the origin, namely by the current yield limit  $Y_\alpha$ . Generally, the  $\alpha$ th yield plane translates at yielding of mode  $\beta = 1 \dots n_y$  (forwards or backwards, depending on the sign of  $H_{\alpha\beta}$ ); its possible contribution to plastic “deformations”  $\{\varepsilon_{ij}^p, \zeta^p\}$  is directed as the (fixed) gradient  $\{N_{ij\alpha}^\sigma, N_{\alpha p}^p\}$  of the relevant yield function and plastic potential  $\varphi_\alpha$ .

The symmetry of the “hardening matrix”  $H_{\alpha\beta}$  is entailed by the adoption of a “locked-in, free energy” potential for the internal variables, which in the above PWL models identify with the plastic multipliers  $\lambda_\alpha$ . However, a nonsymmetric hardening matrix may be directly introduced in the definition of  $\varphi_\alpha$ , (Eq. (1a)). Like in plasticity ([39]), matrix  $H_{\alpha\beta}$  is apt to represent diverse

types of hardening (kinematic, isotropic and mixed, with or without saturation); its positive semi-definiteness, together with normality, implies material stability in Drucker's sense.

In the present PWL context, under the practically weak assumption that  $Y_\alpha^0$  and  $V$  are bounded, the energy characterization of shakedown is readily seen to be provided by the boundedness in time of the plastic multipliers everywhere in  $V$ . On this basis (restricted for convenience, but endowed with all essential features of more general approaches), the two statements which follow can be proven ([35,36]). They represent a generalization to poroplasticity of Melan's static SD theorems, which in turn can be regarded as generalizations of the static theorems of limit analysis.

(I) Shakedown does occur if internal variables (in PWL models  $\lambda_\alpha$ ,  $\alpha = 1 \dots n_y$ ) and a steady-state poroelastic response to fictitious imposed constant strains  $\varepsilon_{ij}^p$  (specifically, consequent constant self-stresses  $\sigma_{ij}^s$ ) exist such that the superposition of it and of the fictitious poroelastic response (marked by superscript  $E$ :  $\sigma_{ij}^E, p^E$ ) to the external actions, after a time  $\tau^*$ , strictly satisfies the constitutive yield inequalities.

(II) If the poroplastic system shakes down in the sense that dissipative inelastic phenomena cease after a time  $\tau^*$ , then there are internal variables and time-independent self-stresses such that the superposition of the latter on the poroelastic response (superscript  $E$ ) after a time  $\tau^*$  satisfies the yield inequalities (strictly with arbitrarily slight increases of the yield limits).

These statements are formally similar to those which extend to hardening (e.g., [41–45]) Ceradini's generalization of the "static" Melan's theorems to dynamics in classical perfect plasticity ([46]). A substantial similarity might be noted between the kinematic poroplastic SD theorems (dual to I and II, see [36]), not discussed herein, and the generalizations to dynamics of Koiter's theorems (Section 2) established in [47,48] for traditional plasticity and further developed, e.g., in [41,45,49,50]. In fact, the fluid turns out to influence the behavior of the system only through the fictitious poroelastic response to the variable loads (like inertia and damping in dynamics). Under the given loading history, captured in that fictitious response, SD arises from the capacity of residual stresses, generated by inelastic strains in the skeleton, to eventually become time-constant self-stresses apt to stop the yielding processes; the time-independence of imposed strains  $\varepsilon_{ij}^p$  and of irreversible changes in fluid content  $\zeta^p$  prevents any influence of them on the fluid motion (provided the permeability is not affected by them).

Like for Koiter's theorems of Section 2, the weak assumption (Lyapunov stability) that infinitesimal perturbations of yield limits have consequences of the same order on the SD limits, legitimates the transition from strict to weak inequalities in the sufficient SD condition and the consequent identification of this with the necessary condition in the operative formulations of SD analysis.

Space discretization of the system is a natural first step toward numerical solutions. Among various approaches available now in computational mechanics, the one used in the first implementation of SD analysis in poroplasticity was mixed finite element modeling "in Prager's generalized variables" (i.e., such that in each pair of work-conjugate fields, the interpolation are related to each other in such a way that the scalar product is preserved with its possible energy meaning). This approach, resting on its variational basis expounded in [51], exhibits the combined appeals of avoiding locking (see Section 2) and of conferring to the discretized description the essential features (symmetry and sign-definiteness of key operators) of the continuum formulations. The latter property legitimates theoretical developments in the space-discrete context.

After the finite element discretization, the search for the SD limits based on the static theorems (I) and (II) can be formulated as follows ([35,36]):

$$\begin{aligned} s = \max_{\mu, \sigma^s, \lambda, \tau^*} \{ \mu \}, \quad \text{subject to:} \\ [\mathbf{N}_\sigma^T \sigma^E(\tau) + \mathbf{N}_p^T \mathbf{p}^E(\tau)] \mu + \mathbf{N}_\sigma^T \sigma^s \leq \mathbf{Y}_0 + \mathbf{H} \lambda, \\ \mathbf{C}^T \sigma^s = \mathbf{0} \quad \forall \tau \geq \tau^*. \end{aligned} \quad (3.4)$$

Here boldface symbols represent matrices and vectors concerning the whole finite element model, account taken of the boundary conditions; the vectors  $\sigma$ ,  $\mathbf{p}$ ,  $\lambda$ , and  $\mathbf{Y}_0$  govern the space distribution of stress, pressure, plastic multipliers, and original yield limits, respectively; superscripts  $E$  and  $S$  denote poroplastic response and self-equilibrium, respectively, the latter being expressed by the equality constraints in problem (3.4). The inequality constraints reflect the yield conditions (3.1) throughout the model. The load factor  $\mu$  concerns the fluctuating live loads, being understood that possible "dead" external actions constant in time, after a transient supposed to be immaterial, merely contribute a time-constant addend to the vector  $\mathbf{Y}_0$  of the generalized original yield limits.

Let the "envelope vector"  $\mathbf{M} \equiv \{ \dots M_\alpha \dots \}^T$  be defined as follows, index  $\alpha$  running over the set of all  $n_y$  yield modes in the whole discrete model:

$$M_\alpha(\tau^*) = \max_{\tau \geq \tau^*} \{ (\mathbf{N}_\sigma^T \sigma^E(\tau) + (\mathbf{N}_p^T \mathbf{p}^E(\tau)))^T \}. \quad (3.5)$$

If a time  $\tau^*$  is fixed and no longer considered an optimization variable, vector  $\mathbf{M}$  defined by Eq. (3.5) becomes constant; then the optimization problem (3.4) reduces to linear programming (LP) and, clearly, provides a lower bound on the safety factor  $s$ :

$$s \geq s^*(\tau^*) = \max_{\mu, \sigma^s, \lambda} \{ \mu \}, \quad \text{subject to:}$$

$$\mu \mathbf{M}(\tau^*) + \mathbf{N}_\sigma^T \sigma^s \leq \mathbf{Y}_0 + \mathbf{H} \lambda, \quad \mathbf{C}^T \sigma^s = \mathbf{0}. \quad (3.6)$$

The following remarks may elucidate potentialities and limitations of the above formulated static SD analysis in poroplasticity, and are also intended to survey some peculiar features of its classical counterpart in plasticity, as a special case of it.

(A) A theoretically and practically crucial question on Eq. (3.6) is how to choose the instant  $\tau^*$  such that the LP problem yields the SD limit (i.e.,  $s^* = s$ ). This purpose can easily be achieved for the following kinds of loading histories.

(i) Periodic external actions: after a transient which depends on the initial conditions and becomes negligible after a few periods, the fictitious poroelastic response becomes periodic (and can be computed efficiently in closed form ([36])). As for SD the transient is immaterial and, therefore, the periodic poroelastic response above can be used in Eq. (3.5) with  $\tau^* = 0$  and leads to  $s^* = s$  in Eq. (3.6).

(ii) Intermittent actions, which (like earthquakes) materialize in a set of known loading histories with relatively brief duration, separated by the same steady-state (like the initial one) under dead loads only. If the poroelastic responses to all those histories in sequence are employed in Eq. (3.5),  $M_\alpha$  computed for  $\tau^* = 0$  becomes independent from the sequential order and, hence, so becomes  $s^*$ , which coincides with  $s$ . In elastic-plastic dynamics intermittent excitations have been studied by Polizzotto and co-workers, with the denomination of "unrestricted" SD ([44,45]).

(B) Hardening material models involve internal variables in the SD criteria and, hence, in the search for the safety factor by mathematical optimization. In the PWL formulation (3.6), the internal variables coincide with the  $\lambda_\alpha$  and these govern through Eq. (3.2a) the plastic strains, which in turn can be conceived as the physical sources of self-stresses  $\sigma^s$ . Therefore,  $\sigma^s$  can be expressed as linear consequences of vector  $\lambda$ :  $\sigma^s = \mathbf{Z} \mathbf{N}_\sigma \lambda$ , denoting by  $\mathbf{Z}$  the (symmetric, negative semi-definite) influence matrix. However, the effort of computing  $\mathbf{Z}$  and the possible algorithmic difficulties caused by numerical noise in its many zero eigenvalues, may more than balance the gains due to the reduction of



variables. An obvious alternative rests on expressing  $\sigma^s$  in terms of redundants through the equilibrium equations in (3.6). Both the above variable reductions in large-size LP problems seem to exhibit an algorithm-dependent, often balanced set of pros and cons, at least in perfect plasticity (i.e., for  $\mathbf{H}=\mathbf{0}$ ). As well known and expected in all contexts of SD analysis, hardening may strongly influence both the SD limit and the collapse mechanism (e.g., kinematic hardening without saturation prevents ratchetting and entails alternating plasticity). A number of results achieved on nonlinear hardening in plasticity (see, e.g., [41–43,52]) can easily be transferred to poroplasticity. The same can be said for recent results concerning damage ([53–56] and cracks [57,58]).

(C) The solid skeleton is a frictional material in most engineering situations amenable to poroplasticity. This implies that the yield functions  $\varphi_\alpha$  depend on the first stress invariant and are distinct from the plastic potentials (say  $\varphi_\alpha$ ) primarily to avoid excessive dilatancy. The latter consequence, nonassociativity, can easily be allowed for PWL approximations in poroplasticity ([36]) like in plasticity, see, e.g., [38,42,49,59]. In fact, two “fictitious associative materials” can be generated (i) by ignoring the deviation from normality and (ii) by assuming as poroelastic deformation the maximum polyhedron contained in the actual domain and having as outward normals the gradients of the plastic potentials. The linear programs, formulated like in Eq. (3.6) but on the basis of model (i) or (ii), lead to (unfortunately often loose) upper and lower bound on  $s$ , respectively.

(D) External actions interpreted as growing slowly and monotonically up to values constant in time, after the extinction of a transient regime, give rise to a time-independent poroelastic state. If the relevant (constant in time) stress and pressure vectors are introduced in Eq. (3.5), problem (3.6) with  $\mathbf{H}=\mathbf{0}$  (perfect plasticity) yields the safety factor  $s$  with respect to plastic collapse, i.e., it defines the carrying capacity of the modeled poroplastic system asymptotically in time. In other terms, the present SD analysis specializes to limit analysis by a static approach generalized to PWL poroplasticity. A remarkable feature of limit analysis in the above sense is that the relevant safety factor may turn out to be less than the safety factor with respect to live loads, fluctuating between zero and the same peak value, an impossible occurrence in plasticity. In other terms, the motion of the fluid in it may increase the carrying capacity of a poroplastic system; moreover, the growth of its frequency may reduce the SD limit in other circumstances ([36]).

(E) From the computational standpoint direct methods have been recently enriched with various valuable contributions, e.g., see ([21,22,60]) and references in [7]. In the present context, first it is worth noting that the generation of PWL constitutive laws entails a computing cost which is generally more than compensated by the savings due to the transition from nonlinear to linear programming. Problem (3.4) may substantially be reduced in size by tentatively ignoring every yield mode  $\alpha$  for which  $Y_\alpha - M_\alpha$  exceeds a pre-assigned tolerance (i.e., whose maximum projection  $M_\alpha$  is sufficiently far from the relevant yield plane, so that this is unlikely to be active). The neglected yield inequalities which turn out to be violated by the stresses resulting from the first trial solution are reconsidered for another solution, see, e.g., [61]. Self-adaptivity as an effective source of computational economy was only recently investigated in the area of direct methods (see, e.g., [60,62,63]). In our experience, satisfactory results are achieved in two-dimensional cases by the simple technique of mesh refinement, specifically by bisection of the longest side of triangular finite elements governed by a suitable norm of plastic strains belonging to the collapse mechanism provided by the solution of the dual to problem (3.6).

The illustrative example of Fig. 4 corroborates remark (D). The drastically idealized plane-strain dam model shown in Fig. 4(a) is formulated with material data, boundary conditions and mixed FE discretization adopted in [36], in particular: Drucker-Prager associative poroplasticity in “effective” stresses; PWL-approximation by 8 yield planes; impervious foundation; 25 m “overtopping”

live load (additional to self-weight and hydrostatic pressure) represented by uniform pressure on the upstream wall. Under fluctuating live load (Fig. 4(b)- $\alpha$ ) with period of three days, the safety factor is  $s=2.57$ . The incremental collapse mechanism (obtained from the solution of the dual LP) is visualized in Fig. 4(a). Under monotonically increasing live load (Fig. 4(b)- $\beta$ ) up to the same overtopping level (of 25 m), the safety factor turns out to be  $s=2.47$ .

Remark (E) is illustrated by Fig. 5 where the simple self-adaptivity mentioned in (E) with “longest side” remeshing strategy, is applied to the limit analysis (as a special case of SD analysis) of another dam model. The four meshes shown there (5(a)–(d)), are correlated in the graph to the growth of the number of optimization variables  $\sigma^s$ , in the solution of problem (3.6) with  $\mathbf{H}=\mathbf{0}$ , and to the improvement (decrease) of the critical overtopping height, i.e., of the safety factor.

The above brief review of recent developments in poroplastic SD analysis has given rise to remarks mostly applicable to the whole area of direct methods. Peculiar of this new subarea of poroplasticity are its needs for research (now in progress) at least on the following issues: relaxation of the restrictive assumptions of full saturation and of constant permeability and provisions for limited ductility of the solid skeleton. As for the last requirement (which of course arises for direct methods in plasticity as well), a promising approach is a combination of limit and deformation holonomic analysis: with constitutive piecewise linearization, even in the presence of softening (i.e., for any matrix  $H_{\alpha\beta}$ ), this approach leads to optimization under complementarity constraints, a fashionable topic of current research in mathematical programming.

## 4 Upper Bounds

Shakedown analysis may be nonconservative when some of the idealizations on which it rests are violated. Let us consider in particular the following hypotheses: (a) “small” deformation and serviceability within SD ranges; (b) unlimited ductility; and (c) in poroplasticity, linear diffusion equation, i.e., permeability constant in time. Remedies to the above possible unconservativeness, as a consequence of invalidation of some basic assumption, may be found still in the area of direct methods, precisely in methods apt to provide upper bounds on the post-SD values of meaningful quantities which depend on the whole history of the inelastic system. Upper bounds on meaningful residual displacements, plastic strains and volumetric plastic strains in representative crucial points, if they turn out to be lower than critical thresholds, guarantee that the above hypotheses (a) (b) (c), respectively, are actually acceptable in the engineering situation considered; otherwise they may give guidelines for suitable adjusted reformulations and iterated solutions of the SD analysis problem, or they may induce recourse to costly time-stepping analyses.

A variety of bounding techniques in the above sense were established since the early 1970s, especially in dynamics and in viscoplasticity and creep, e.g., ([11,12,64–70]). In fact, the time-dependence of the system reduces the applicability of SD concept (e.g., SD in the present sense clearly occurs under impact loads, as it would do with unrealistic isotropic hardening without saturation). Beyond the present purposes would be a survey of bounds, from those based on “dummy loads” to their unifying generalizations using “perturbation parameters” (e.g., [65]), to those emerging from “ad hoc” energy considerations, and, finally, to the bilateral ones like in [66].

In the framework of the space discretization and constitutive piecewise linearization adopted in Section 3 for poroplastic SD analysis, the actual post-shakedown value of a meaningful displacement or plastic strain turns out to be a linear function of the post-shakedown generalized plastic multiplier vector  $\bar{\lambda}$ .

In this context several upper bounds on these quantities have recently been established using the condensation of variables mentioned at (B) in Section 3. After the SD analysis by LP leading

to the safety factor  $s$ , the solution of a min-max problem in the space of vector  $\{\lambda^*, \lambda\}$ , provides the most stringent upper bound. The lower bound which is likely to often attain the best compromise between the requirements of accuracy and computer economy, encompasses two phases: convex quadratic programming in  $\lambda^*$ ; linear programming in  $\lambda$ . For the residual horizontal displacement of the crest point P in the dam model of Fig. 4(a), under periodic overtopping load (Fig. 4(b)-a) with amplitude equal to half the SD limit (but beyond the poroelastic range), the bound resulting from the above procedure is plotted in Fig. 6 versus the dimensionality of vector  $\lambda$  (and  $\lambda^*$ ). It is worth noting that the mesh refinement (here without self-adaptivity like in Fig. 5) significantly improves the bound, which converges on 4.75 cm, while the "actual" displacement computed by time-marching analysis amounts to almost 2.6 cm.

## 5 Conclusions

"Research in engineering science cannot be considered to be completed until its full numerical evaluation has been obtained, and the results have been presented in a form accessible to other engineers." This sentence of Warner Koiter in the "Festschrift" for his 65th anniversary suggests a criterion for assessing the maturity of recent developments in shakedown analysis, at least in areas related to the issues considered in this paper. Nowadays available numerical evaluations are probably sufficient to corroborate the competitiveness and, in some cases, superiority of selected direct methods (like the ones considered in Section 2) with respect to time-stepping methods in engineering analysis of ductile, elastic-plastic, and stable (in Drucker's sense) structures and microstructures exposed to persistent yielding under variable repeated external actions. This remark holds not only in the original classical context of quasi-static perfect plasticity, but also for various generalizations scattered along the last three-four decades, "in primis" extensions to dynamics, hardening, geometric effects, and creep. Time is ripe for implementations into practical engineering tools "presented in a form accessible to other engineers." Such presentation advocated by Koiter should cover also carefully selected modern algorithms of mathematical programming, unusual so far in commercial software for structural and mechanical engineering (though they might be in common with software for optimum design and parameter identification). The practicalness requires putting to work contributions dispersed in the literature such as those concerning self-adaptivity and remeshing, reduction of yield mode number, discontinuities at finite element interfaces, and anti-locking provisions.

Research is certainly not completed in the new area of coupled problems, where the severe limitations of the extensions surveyed in Section 3 and 4 should (and are being) overcome in view of reliable engineering applications. A priori shakedown criteria of real practical interest are still to be found, despite many valuable research results for situations characterized by finite strains, damage, cracks, and softening behavior.

In the case of shakedown analysis, and more generally of direct methods for ultimate state analysis, the appealing challenge in the years to come appears to be the generation, on a unifying theoretical basis, of practical computing tools, versatile enough to provide both physical insight and quantitative essential information on the load-carrying capacity with respect to incremental and plastic collapse, low and high-cycle fatigue and possible wear as well, of diverse systems, such as infrastructures and buildings in seismic areas, structural components of power plants, dams, offshore structures, pavements and rails, and textures of materials at the microscale.

## Acknowledgments

The recent results concerning periodic heterogeneous media (Section 2) and poroplastic systems (Sections 3 and 4) have been obtained in the framework of "CNR Project 99.01818.PF34" on "Special materials" and of MURST Project 98.08160.821 on

"Large Dams," respectively. The cooperation with various colleagues and with my former students R. Ardito, V. Carvelli, and G. Cocchetti is gratefully acknowledged.

## References

- [1] Koiter, W. T., 1960, "General Theorems for Elastic-Plastic Solids," *Progress in Solid Mechanics*, Vol. 1, J. N. Sneddon and R. Hill, eds., North-Holland, Amsterdam, pp. 165–221.
- [2] König, J. A., 1987, *Shakedown of Elastic-Plastic Structures*, Elsevier, Amsterdam.
- [3] Kamenjarsch, J. A., 1996, *Limit Analysis of Solids and Structures*, CRC Press, Boca Raton, FL.
- [4] Lloyd Smith, D., ed., 1990, *Mathematical Programming Methods in Structural Plasticity*, Springer-Verlag, New York.
- [5] Mróz, Z., Weichert, D., and Dorosz, S., eds., 1995, *Inelastic Behavior of Structures Under Variable Loads*, Kluwer, Dordrecht, The Netherlands.
- [6] Weichert, D., and Maier, G., eds., 2000, *Inelastic Analysis of Structures Under Variable Repeated Loads*, Kluwer, Dordrecht, The Netherlands.
- [7] Maier, G., Carvelli, V., and Cocchetti, G., 2000, "On Direct Methods for Shakedown and Limit Analysis," *Eur. J. Mech. A/Solids (Special Issue)*, **19**, pp. S79–S100.
- [8] Koiter, W. T., 1956, "A New General Theorem on Shakedown of Elastic-Plastic Structures," *Proc. K. Ned. Akad. Wet.*, **B59**, pp. 24–34.
- [9] Karadeniz, S., and Ponter, A. R. S., 1984, "A Linear Programming Upper Bound Approach to the Shakedown Limit of Thin Shells Subjected to Variable Thermal Loading," *J. Strain Anal.*, **19**, pp. 221–230.
- [10] Yan, A., and Nguyen-Dang, H., 2001, "Kinematical Shakedown Analysis With Temperature-Dependent Yield Stress," *Int. J. Numer. Methods Eng.*, **50**, pp. 1145–1168.
- [11] Polizzotto, C., 1984, "Deformation Bounds for Elastic Plastic Solids Within and Out of the Creep Range," *Nucl. Eng. Des.*, **83**, pp. 293–301.
- [12] Ponter, A. R. S., 1972, "Deformation, Displacement and Work Bounds for Structures in a State of Creep and Subject to Variable Loading," *ASME J. Appl. Mech.*, **39**, pp. 953–959.
- [13] Carvelli, V., Cen, Z., Liu, Y., and Maier, G., 1999, "Shakedown Analysis of Defective Pressure Vessels by a Kinematic Approach," *Arch. Appl. Mech.*, **69**, pp. 751–764.
- [14] Carvelli, V., Maier, G., and Taliercio, A., 1999, "Shakedown Analysis of Periodic Heterogeneous Materials by a Kinematic Approach," *Mech. Eng. (Strojnicky Casopis)*, **50**, No. 4, pp. 229–240.
- [15] Carvelli, V., Maier, G., and Taliercio, A., 2000, "Kinematic Limit Analysis of Periodic Heterogeneous Media," *Comp. Meth. Eng. Sci.*, **1**, pp. 15–26.
- [16] Casciaro, R., and Cascini, L., 1982, "A Mixed Formulation and Mixed Finite Elements for Limit Analysis," *Int. J. Numer. Methods Eng.*, **18**, pp. 211–243.
- [17] König, J. A., and Kleiber, M., 1978, "On a New Method of Shakedown Analysis," *Bull. Acad. Pol. Sci., Ser. Sci. Tech.*, **26**, pp. 165–171.
- [18] Zhang, Y. G., 1995, "An Iterative Algorithm for Kinematic Shakedown Analysis," *Comput. Methods Appl. Mech. Eng.*, **127**, pp. 217–226.
- [19] Kamenjarsch, J. A., and Merzljakov, A., 1994, "On Kinematic Method in Shakedown Theory; I. Duality of Extremum Problems; II. Modified Kinetic Method," *Int. J. Plast.*, **10**, pp. 363–392.
- [20] Kamenjarsch, J. A., and Weichert, D., 1992, "On Kinematic Upper Bounds for the Safety Factor in Shakedown Theory," *Int. J. Plast.*, **8**, pp. 827–837.
- [21] Sloan, S. W., and Kleeman, P. W., 1995, "Upper Bound Limit Analysis Using Discontinuous Velocity Fields," *Comput. Methods Appl. Mech. Eng.*, **127**, pp. 293–314.
- [22] Teixeira de Freitas, J. A., 1991, "A Kinematic Model for Plastic Limit Analysis of Solids by the Boundary Element Method," *Comput. Methods Appl. Mech. Eng.*, **88**, pp. 189–205.
- [23] Dvorak, G. J., Lagoudas, D. C., Huang, C. M., 1994, "Fatigue Damage and Shakedown in Metal Matrix Composite Laminates," *Mech. Compos. Mat. Struct.*, **1**, pp. 171–202.
- [24] Francescato, P., and Pastor, J., 1997, "Lower and Upper Numerical Bounds to the Off-Axis Strength of Unidirectional Fiber-Reinforced Composites by Limit Analysis Methods," *Eur. J. Mech. A/Solids*, **16**, pp. 213–234.
- [25] Weichert, D., Hachemi, A., and Schwabe, F., 1999, "Shakedown Analysis of Composites," *Mech. Res. Commun.*, **26**, pp. 309–318.
- [26] Liu, Y. H., Cen, Z. Z., and Xu, B. Y., 1995, "A Numerical Method for Plastic Limit Analysis of 3-D Structures," *Int. J. Solids Struct.*, **32**, pp. 1645–1658.
- [27] Weichert, D., Hachemi, A., and Schwabe, F., 1999, "Application of Shakedown Analysis to the Plastic Design of Composites," *Arch. Appl. Mech.*, **69**, pp. 623–633.
- [28] Hamilton, R., Boyle, J. T., Shi, J., and Mackenzie, D., 1996, "A Simple Upper-Bound Method for Calculating Approximate Shakedown Loads," *ASME J. Pressure Vessel Technol.*, **120**, pp. 195–199.
- [29] Ponter, A. R. S., and Carter, K. F., 1997, "Shakedown State Simulation Techniques Based on Linear Elastic Solutions," *Comput. Methods Appl. Mech. Eng.*, **140**, pp. 259–279.
- [30] Drucker, D. C., 1963, "On the Macroscopic Theory of Inelastic Stress-Strain-Time-Temperature Behavior," *Advances in Materials Research in the NATO Nations (AGAR Dograph 62)*, Pergamon Press, New York, pp. 193–221.
- [31] Dang Van, K., and Papadopoulos, I. V., 1999, *High-Cycle Metal Fatigue From Theory to Applications*, CISM, Springer-Verlag, New York.
- [32] Ponter, A. R. S., and Leckie, F. A., 1998, "Bounding Properties of Metal-



- Matrix Composites Subjected to Cyclic Loading," *J. Mech. Phys. Solids*, **46**, pp. 697–717.
- [33] Silberschmidt, V. V., Rammerstorfer, F. G., Werner, E. A., Fischer, F. D., and Uggowitzer, P. J., 1999, "On Material Immanent Ratchetting of Two-Phase Materials Under Cyclic Purely Thermal Loading," *Arch. Appl. Mech.*, **69**, pp. 727–750.
- [34] Lewis R. W., and Schrefler B. A., 1998, *The Finite Element Method in the Static and Dynamic Deformation and Consolidation of Porous Media*, John Wiley and Sons, Chichester.
- [35] Cocchetti, G., and Maier, G., 1998, "Static Shakedown Theorems in Piecewise Linearized Poroplasticity," *Arch. Appl. Mech.*, **68**, pp. 651–661.
- [36] Cocchetti, G., and Maier, G., 2000, "Shakedown Analysis in Poroplasticity by Linear Programming," *Int. J. Numer. Methods Eng.*, **47**, No. 1–3, pp. 141–168.
- [37] Cocchetti, G., and Maier, G., 2000, "Upper Bounds on Post-Shakedown Quantities in Poroplasticity," *Inelastic Analysis of Structures Under Variable Repeated Loads*, D. Weichert and G. Maier, eds., Kluwer, Dordrecht, The Netherlands, pp. 289–314.
- [38] Maier, G., 1969, "Shakedown Theory in Perfect Elastoplasticity With Associated and Nonassociated Flow-Laws: A Finite Element, Linear Programming Approach," *Meccanica*, **4**, pp. 250–260.
- [39] Maier, G., 1970, "A Matrix Structural Theory of Piecewise-Linear Plasticity With Interacting Yield Planes," *Meccanica*, **5**, pp. 55–66.
- [40] Tin-Loi, F., 1990, "A Yield Surface Linearization Procedure in Limit Analysis," *Mech. Struct. Mach.*, **18**, pp. 135–149.
- [41] Comi, C., and Corigliano, A., 1991, "Dynamic Shakedown in Elastoplastic Structures With General Internal Variable Constitutive Laws," *Int. J. Plast.*, **7**, pp. 679–692.
- [42] Corigliano, A., Maier, G., and Pycko, S., 1995, "Dynamic Shakedown Analysis and Bounds for Elastoplastic Structures With Nonassociative, Internal Variable Constitutive Laws," *Int. J. Solids Struct.*, **32**, pp. 3145–3166.
- [43] Du, S. T., Xu, B. Y., and Dong, Y. F., 1993, "Dynamic Shakedown Theory of Elastoplastic Work-Hardening Structures Allowing for Second-Order Geometric Effects," *Acta Mech. Solidica Sinica*, **6**, pp. 15–26.
- [44] Polizzotto, C., 1984, "On Shakedown of Structures Under Dynamic Agencies," *Inelastic Analysis Under Variable Loads*, A. Sawczuk and C. Polizzotto, eds., Cograss, Palermo, pp. 5–29.
- [45] Polizzotto, C., Borino, G., Caddemi, S., and Fuschi, P., 1993, "Theorems of Restricted Dynamic Shakedown," *Int. J. Mech. Sci.*, **35**, pp. 787–801.
- [46] Ceradini, G., 1969, "Sull'adattamento dei corpi elastoplastici soggetti ad azioni dinamiche," *Gior. Genio Civile*, **415**, pp. 239–258.
- [47] Corradi, L., and Maier, G., 1973, "Inadaptation Theorems in the Dynamics of Elastic-Work Hardening Structures," *Ing. Arch.*, **43**, pp. 44–57.
- [48] Corradi, L., and Maier, G., 1974, "Dynamic Non-Shakedown Theorem for Elastic Perfectly-Plastic Continua," *J. Mech. Phys. Solids*, **22**, pp. 401–413.
- [49] Corigliano, A., Maier, G., and Pycko, S., 1995, "Kinematic Criteria of Dynamic Shakedown Extended to Nonassociative Constitutive Laws With Saturation Hardening," *Rend. Acc. Naz. Lincei. Sci., Ser. IX*, **VI**, pp. 55–64.
- [50] Pham, D. C., 1996, "Dynamic Shakedown and a Reduced Kinematic Theorem," *Int. J. Plast.*, **12**, pp. 1055–1068.
- [51] Maier, G., and Comi, C., 1997, "Variational Finite Element Modelling in Poroplasticity," *Recent Developments in Computational and Applied Mechanics*, B. D. Reddy, ed., CIMNE, Barcelona, pp. 180–199.
- [52] Maier, G., and Novati, G., 1990, "Dynamic Shakedown and Bounding Theory for a Class of Nonlinear Hardening Discrete Structural Models," *Int. J. Plast.*, **6**, pp. 551–572.
- [53] Druyanov, B., and Roman, I., 1999, "Conditions for Shakedown of Damaged Elastic Plastic bodies," *Eur. J. Mech. A/Solids*, **18**, pp. 641–651.
- [54] Dvorak, G. J., Lagoudas, D. C., and Huang, C. M., 1994, "Fatigue Damage and Shakedown in Metal Matrix Composite Laminates," *Mech. Compos. Mat. Struct.*, **1**, pp. 171–202.
- [55] Feng, X. Q., and Yu, S. W., 1995, "Damage and Shakedown Analysis of Structures With Strain-Hardening," *Int. J. Plast.*, **11**, pp. 237–249.
- [56] Hachemi, A., and Weichert, D., 1997, "Application of Shakedown Theory to Damaging Inelastic Material Under Mechanical and Thermal Loads," *Int. J. Mech. Sci.*, **39**, pp. 1067–1076.
- [57] Huang, Y., and Stein, E., 1996, "Shakedown of a Cracked Body Consisting of Kinematic Hardening Material," *Eng. Fract. Mech.*, **54**, pp. 107–112.
- [58] Yan, A. M., and Nguyen, D. H., 1999, "Limit Analysis of Cracked Structures by Mathematical Programming and Finite Element Technique," *Comput. Mech.*, **24**, pp. 319–333.
- [59] Nayroles, B., and Weichert, D., 1993, "La notion de sanctuaire d'élasticité et d'adaptation des structures," *C. R. Acad. Sci., Ser. II: Mec., Phys., Chim., Sci. Terre Univers*, **316**, pp. 1493–1498.
- [60] Christiansen, E., and Andersen, K. D., 1999, "Computation of Collapse States With von Mises Type Yield Condition," *Int. J. Numer. Methods Eng.*, **46**, pp. 1185–1202.
- [61] Tin-Loi, F., 1989, "A Constraint Selection Technique in Limit Analysis," *Appl. Math. Model.*, **13**, pp. 442–446.
- [62] Borges, L. A., Feijóo, R. A., and Zouain, N., 1999, "A Directional Error Estimator for Adaptive Limit Analysis," *Mech. Res. Commun.*, **26**, pp. 555–563.
- [63] Franco, J. R. Q., Oden, J. T., Ponter, A. R. S., and Barros, F. B., 1997, "A Posteriori Error Estimator and Adaptive Procedures for Computation of Shakedown and Limit Loads on Pressure Vessels," *Comput. Methods Appl. Mech. Eng.*, **150**, pp. 155–171.
- [64] Cocks, A. C. F., and Leckie, F. A., 1988, "Deformation Bounds for Cyclically Loaded Shell Structures Operating Under Creep Condition," *ASME J. Appl. Mech.*, **55**, pp. 509–516.
- [65] Polizzotto, C., 1982, "A Unified Treatment of Shakedown Theory and Related Bounding Techniques," *Solid. Mech. Arch.*, **7**, pp. 19–75.
- [66] Genna, F., 1991, "Bilateral Bounds for Structures Under Dynamic Shakedown Conditions," *Meccanica*, **26**, pp. 37–46.
- [67] Capurso, M., 1979, "Some Upper Bound Principles for Plastic Strains in Dynamic Shakedown of Elastoplastic Structures," *J. Struct. Mech.*, **7**, pp. 1–20.
- [68] Corradi, L., 1976, "Mathematical Programming Methods for Displacement Bounds in Elastoplastic Dynamics," *Nucl. Eng. Des.*, **37**, pp. 161–177.
- [69] Maier, G., 1973, "Upper Bounds on Deformations of Elastic-Workhardening Structures in the Presence of Dynamic and Second-Order Effects," *J. Struct. Mech.*, **2**, pp. 265–280.
- [70] Ponter, A. R. S., 1975, "General Displacement and Work Bounds for Dynamically Loaded Bodies," *J. Mech. Phys. Solids*, **23**, pp. 151–163.
- [71] Taliercio, A., 1992, "Lower and Upper Bounds to the Macroscopic Strength Domain of a Fiber-Reinforced Composite Material," *Int. J. Plast.*, **8**, pp. 741–762.

## Warner T. Koiter Medal Recipient

Conferral at the Koiter Lecture,  
2000 International Mechanical Engineering Congress and Exposition

THE WARNER T. KOITER MEDAL was established in 1996 to recognize distinguished contributions to the field of solid mechanics with emphasis on the effective blending of theoretical and applied elements, and on a high degree of leadership in the international solid mechanics community.

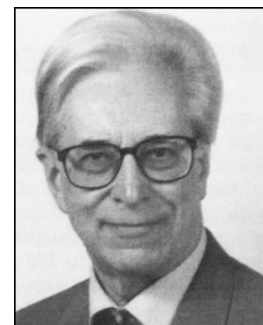
The medal honors the late Dr. Warner T. Koiter (1914–1997), world renowned authority in the field of solid mechanics, and it commemorates his vast contributions as research engineer and teacher.

GIULIO MAIER, professor, Technical University of Milan, Italy, for fundamental research in solid and structural mechanics; for pioneering analytical and computational work on structural and theoretical plasticity, computational mechanics, mathematical programming methods, structural optimization, and the application of this research to engineering; and for his outstanding leadership in the applied mechanics community.

Dr. Maier is an outstanding researcher, a leader in the applied mechanics community and a key link between European and U.S. applied mechanicians.

Since 1961, Maier has been at the Technical University (Politecnico) of Milan, Italy, where in 1970 he assumed the position of professor of structural mechanics and engineering (Scienza delle Costruzioni). He served as department head, coordinator of continuing education and coordinator of the doctoral school in structural engineering.

Previous academic appointments include research associate at



Giulio Maier

Brown University (Providence, R.I.) and Cambridge University, England, in 1964 and 1967, respectively; and visiting professor of civil engineering at the University of Illinois at Urbana-Champaign and the University of Minnesota, Minneapolis, in 1970 and 1997, respectively.

Maier's research contributions include: various extremum theorems in plasticity; effects of constitutive instability (softening, nonassociativity) on structural responses to loads; criteria for overall stability, bifurcation, algorithmic stability in dynamics; shakedown theorems and upper bounds in dynamics, poroplasticity and with geometric effects; mathematical programming methods for structural analysis and optimum design; parameter identification methods in materials and structural mechanics; symmetric Galerkin boundary element methods; and various practical procedures for analysis and design of tension structures and offshore structures and pipelines.

As a consultant, Maier, with his co-workers, contributed to significant projects, such as the Trans-Mediterranean Pipeline, where he solved structural problems in deep water laying operations across the Sicily Channel between Tunisia and Italy; the subsea tunnels crossing the Messina Straits, where he conducted explosive damage simulations; and safety studies of off-shore pipelines and nuclear power plant components.

Maier is the author/co-author of 230 publications. He has been associate editor of the *European Journal of Mechanics/Solids* and

editor of *Meccanica*, and is at present a member of the editorial boards of 15 scientific journals.

A member of ASME, Maier is a Fellow and Life Member of the American Society of Civil Engineers; a Fellow of the Italian National Academy (Lincei) in Rome, the Italian National Academy of Sciences (dei XL) in Rome and the Academy of Lombardy (Istituto Lombardo) in Milan; a member and former president of the Italian Association of Theoretical and Applied Mechanics; a member and former council member of the European Mechanics Society; and a Fellow of the International Association of Computational Mechanics. He is also a foreign member of the Polish Academy of Sciences, the Hungarian Academy of Sciences, the Russian Academy of Engineering, the Royal Society of South Africa and the New York Academy of Sciences.

Among his other honors is the University Medal from the University of Colorado, Boulder (1998); the Copernicus Medal from the Polish Academy of Sciences, Warsaw (1984); and the Feltrinelli Prize from the Italian National Academy (1981).

Maier received his master's degree (laurea) in mechanical engineering at the University of Trieste, Italy, in 1955. He earned his "specializzazione" (doctoral degree) in aerospace engineering at the University of Rome, Italy, in 1958. He received honorary doctorates from the Aristotle University (Thessaloniki, Greece) and the Faculté Polytechnique de Mons, Belgium, in 1985 and 1992, respectively. He is a chartered civil engineer in Milan, Italy.

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 Editor of the JOURNAL OF APPLIED MECHANICS. Discussions on the Brief Notes should be addressed to the Editorial Department, ASME International, Three Park Avenue, New York, NY 10016-5990, or to the 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 Existence of a Solution for a Solid Circular Plate Bilaterally Supported Along Two Antipodal Boundary Arcs and Loaded by a Central Transverse Concentrated Force

**G. Monegato**

Politecnico di Torino, Corso Duca degli Abruzzi 24,  
10129 Torino, Italy

**A. Strozzi**

Faculty of Engineering, Modena and Reggio Emilia  
University, Via Vignolese 905, 41100 Modena, Italy

*A purely flexural mechanical analysis is presented for a thin, solid, circular plate, deflected by a central transverse concentrated force, and bilaterally supported along two antipodal periphery arcs, the remaining part of the boundary being free. This problem is modeled in terms of a singular integral equation of the Prandtl type, which possesses a unique solution expressed in terms of a reaction force containing a factor exhibiting square root endpoint singularities. This solution is then shown not to respect the requested boundary constraints. It is therefore concluded that, within the framework of the purely flexural plate theory, the title problem cannot admit the weighted  $L^2$  solution here examined. It cannot, however, be excluded that a solution to the title problem exists, which possesses stronger endpoint singularities than those examined in this paper, or is of a more general form than the one considered here. [DOI: 10.1115/1.1379037]*

### 1 Introduction

Circular plates axisymmetrically loaded but constrained nonaxisymmetrically have repeatedly attracted the attention of the researchers. The analytical-numerical papers addressing such problems may be classified in two main categories, where the first group comprises analytical solutions based upon an integral representation, for example, Sherman [1], whereas the second set encompasses studies of differential fashion, for instance, Samodurov and Tikhomirov [2].

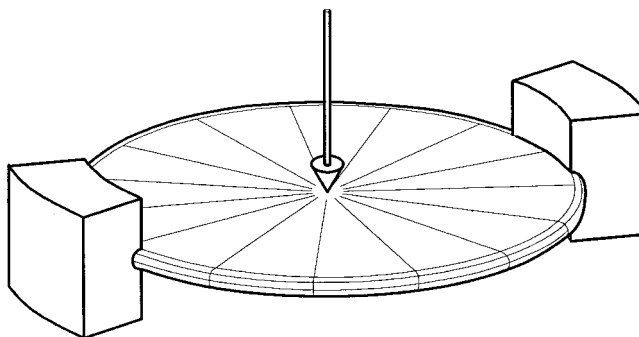
In this paper a thin, solid circular plate, bilaterally supported along two antipodal periphery arcs and deflected by a transverse central force is considered (see Fig. 1). Within the framework of the purely flexural plate theory, it is demonstrated that this problem admits no solution exhibiting certain types of endpoint algebraic/logarithmic singularities, a results which justifies this return to a classical problem. This negative conclusion is parallel to that obtained by Grigolyuk and Tolkachev [3], p. 381, with regard to an infinite plate resting on a line support.

2 Formulation of the Integral Equation for the Plate

Within the framework of the classical, purely flexural theory of elasticity for plates, the expression for the edge deflection  $w(\theta)$  in a thin, solid, circular plate loaded by a transverse, concentrated, central force  $P$ , and by two antipodal border forces  $P/2$  equilibrating the central load is [4]

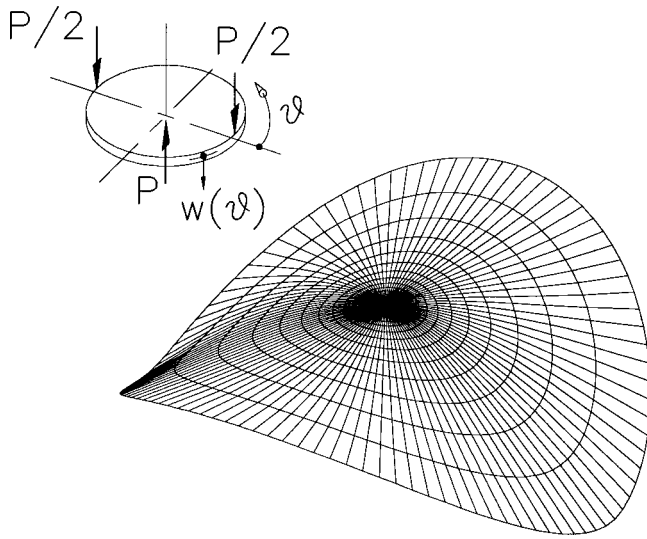
$$w(\theta) = \frac{48\pi D(3+\nu)(1-\nu^2)}{Pr_o^2} \left[ \ln|2\sin\theta| - \cos\theta \ln\left|\tan\frac{\theta}{2}\right| \right] + 12(1+\nu)^2[\pi(\theta - \sin\theta) - \theta^2] - 2\pi^2(1+\nu)^2 + 3(1-\nu)^3$$

$$0 \leq \theta \leq \pi \quad (2.1)$$



**Fig. 1 A thin, solid, circular plate, bilaterally supported along two antipodal periphery arcs and deflected by a transverse central force**

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, Dec. 18, 2000. Associate Editor: R. C. Benson.



**Fig. 2 The deflection  $w(\theta)$  in a circular plate loaded by a transverse, concentrated, central force  $P$ , and by two antipodal border forces  $P/2$**

where  $P$  represents the intensity of the central load,  $D$  denotes the flexural rigidity of the plate,  $\nu$  indicates the Poisson's ratio,  $r_o$  the plate outer radius, and  $\theta$  the angular coordinate, whose origin coincides with the point of application of one of the two boundary loads (Fig. 2). The boundary deflection  $w(\theta)$  is expressed with respect to the plate center. In the following, Eq. (2.1) is symbolized as

$$w(\theta) = PK(\theta). \quad (2.2)$$

In a solid circular plate of radius  $r_o$ , loaded by the same distributed force  $F$  applied along two antipodal boundary arcs of angular width  $2\alpha$  and by an equilibrating central force  $P$ , the integral representation of the deflection of the loaded plate border with respect to the plate center is

$$w(\theta) = 2 \int_{-\alpha}^{+\alpha} K(|\theta - \omega|) F(\omega) r_o d\omega \quad (2.3)$$

where the presence of number 2 in Eq. (2.3) derives from the observation that a unity periphery load is connected to a central force of intensity 2 (see Fig. 2). The origin of angle  $\theta$  coincides with the center of one of the loaded arcs, while angle  $\omega$  is an integration device whose origin coincides with that of  $\theta$ .

When a circular plate is considered which is loaded by a central transverse force  $P$  and is bilaterally sustained along two antipodal edge arcs of angular width  $2\alpha$ , the supported arcs exhibit a constant deflection with respect to the plate center, since the support is assumed to be rigid in comparison to the plate flexibility. The integral equation describing the title problem may thus be obtained by equalling to a constant the integral expression (2.3) of the plate deflection. It is also noted that, since the plate supported arcs must remain flat after deflection, the derivatives with respect to  $\theta$  of the plate deflection  $w$  along the supported arc must vanish. In particular, a remarkably simple integral equation is achieved by annulling the sum of the first and third derivatives

$$\frac{dw(\theta)}{d\theta} + \frac{d^3w(\theta)}{d\theta^3} = 0 \Rightarrow \int_{-\alpha}^{+\alpha} \left( \frac{\partial K}{\partial \theta} + \frac{\partial^3 K}{\partial \theta^3} \right) F(\omega) d\omega = 0. \quad (2.4)$$

By exploiting the symmetry of the reaction force  $F$  and the condition (2.5), after lengthy passages the integral Eq. (2.4) takes the particularly compact Prandtl-type form

$$\begin{aligned} & \int_{-\alpha}^{+\alpha} \frac{1}{\tan(\theta - \omega)} F(\omega) d\omega + \frac{(1 + \nu)\pi}{2} \int_{-\alpha}^{\theta} F(\omega) d\omega \\ &= \frac{(1 + \nu)}{4r_o} P \left( \theta + \frac{\pi}{2} \right) \end{aligned} \quad (2.5)$$

which is a Fredholm integral equation of the first kind with Hilbert-type kernel, (Mikhlin [5]).

The integral Eq. (2.5) has to be coupled with an additional condition, expressing the fact that each plate support sustains half the central load  $P$  whose intensity is considered as known,

$$\int_{-\alpha}^{+\alpha} F(\omega) r_o d\omega = \frac{P}{2}. \quad (2.6)$$

Concerning the derivatives of  $w$  with respect to  $\theta$  employed in (2.5), it is noted that the integral representation of the first derivative of  $w(\theta)$  is well defined when, for example,  $F \in L^1(-\alpha, \alpha)$ , or even when  $F$  is a Dirac delta function  $\delta(\omega - \omega_0)$ . The corresponding formal representation of the third derivative still holds, although the derivative itself must be interpreted in a proper weak sense, when  $F \in L^1(-\alpha, \alpha)$ . This derivative is also defined when  $F(\omega) = \delta(\omega \pm \alpha)$ , since in this case it belongs to a weighted  $L^1$  space. Therefore, when  $F(\omega)$  is a  $L^1$  function, eventually plus a linear combination of  $\delta(\omega \pm \alpha)$ , nothing is lost by considering the Eq. (2.5).

It is observed that a reaction force which solves the title problem is also a solution to the new Eq. (2.5). On the other side, Eq. (2.5) models a problem which is more general than that expressed in the title, since the first of Eqs. (2.4) implies that  $w$  be of the form

$$w(\theta) = C_1 + C_2 \cos \theta \quad (2.7)$$

where, due to the problem symmetry, only the deflection terms which are symmetrical with respect to  $\theta$  have been considered in (2.7). Equation (2.7) shows that a solution to the integral Eq. (2.5) coupled with condition (2.6) produces a plate deflection which is the sum of a constant and of a cosinusoidal term. A constant deflection is consistent with the problem expressed in the title, since rigid supports would produce constant deflections of the supported border arcs with respect to the plate center whereas the cosinusoidal term is undesired. The two constants  $C_1$  and  $C_2$  are not generic, since they depend upon the angular width  $2\alpha$  of the supports and upon the selected  $F(\omega)$ . The expressions for constants  $C_1$  and  $C_2$  can be derived from (2.3) and (2.7),

$$C_2 = - \frac{d^2 w(\theta)}{d\theta^2} \Big|_{\theta=0}; \quad C_1 = w(\theta) \Big|_{\theta=0} - C_2, \quad (2.8)$$

where  $w(\theta)$  in (2.8) depends on the selected  $F(\omega)$ . Since the cosinusoidal deflection is not consistent with the modeling of the title problem, it is explored in the following section whether constant  $C_2$  vanishes.

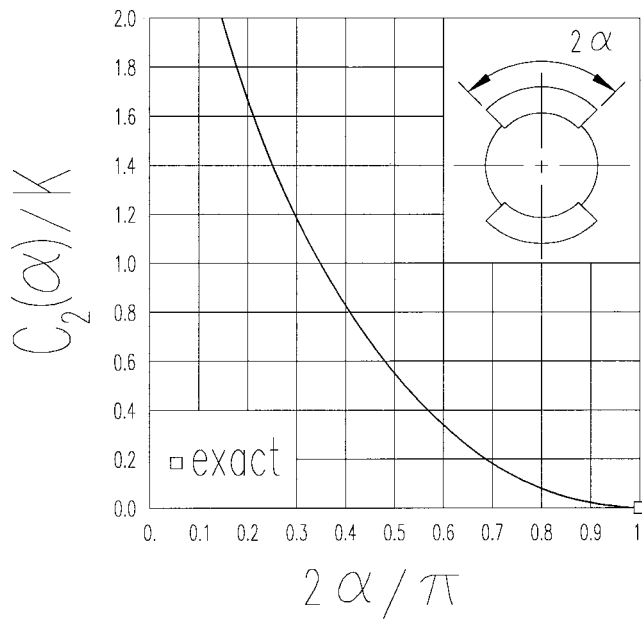
### 3 Evaluation of the Reaction Force Distribution and of the Plate Deflections for the Solution of the Integral Eq. (2.5)

As suggested by the form of the analytical solution to the integral Eq. (2.5) for vanishing Volterra term (Ling [6], p. 306), a solution to the complete Eq. (2.5) is expressed in the form (Gladwell [7])

$$F(\omega) = \frac{\cos u\omega \left( \frac{\sin \omega}{\sin \alpha} \right)}{\sqrt{\sin^2 \alpha - \sin^2 \omega}} \quad (3.1)$$

with  $u(x)$  even. This is a standard procedure when solving singular integral equations. Setting





**Fig. 3** Coefficient  $C_2$  normalized over  $K$ , versus  $2\alpha/\pi$ , for  $\nu=0.3$

$$x = \frac{\sin \omega}{\sin \alpha}, \quad t = \frac{\sin \theta}{\sin \alpha} \quad (3.2)$$

where in the following it is assumed that  $0 < \alpha < \pi/2$ , Eq. (2.5) becomes

$$\frac{\sqrt{1-t^2 \sin^2 \alpha}}{\sin \alpha} \frac{1}{\pi} \int_{-1}^{+1} \frac{1}{\sqrt{1-x^2}} \frac{u(x)}{(t-x)} dx + \frac{(1+\nu)}{2} \int_{-1}^t \frac{u(x)}{\sqrt{1-x^2}} dx = f(t) \quad (3.3)$$

where

$$f(t) = \frac{(1+\nu)}{4\pi r_o} P \left[ a \sin(t \sin \alpha) + \frac{\pi}{2} \right] \quad (3.4)$$

whereas condition (2.6) assumes the new form

$$\int_{-1}^{+1} \frac{u(x)}{\sqrt{1-x^2}} dx = \frac{P}{2r_o}. \quad (3.5)$$

It may be proved that the Prandtl integral Eq. (3.3) coupled with condition (3.5) possesses a unique solution for the whole physical range of  $\nu$ , with  $u(x)$  respecting the condition

$$\int_{-1}^1 \frac{|u(x)|^2}{\sqrt{1-x^2}} dx < \infty, \quad (3.6)$$

that is, with  $F(\omega)$  belonging to the weighted  $L^2$  space defined by the inequality

$$\int_{-\alpha}^{+\alpha} \sqrt{\sin^2 \alpha - \sin^2 \omega} |F(\omega)|^2 d\omega < \infty. \quad (3.7)$$

This proof is omitted for brevity, where details may be found in Monegato and Strozzi [8].

The reaction force  $F(\omega)$  expressed by (3.1) which solves the integral Eq. (3.3) is numerically determined to the requested accuracy. It is then shown that this reaction force produces a non-vanishing constant  $C_2$  for all angular widths  $2\alpha$  of the supports,  $0 < 2\alpha < \pi$ . Approximating  $u(x)$  by  $u_N(x)$

$$u_N(x) = \frac{P}{2\pi r_o} + \sum_{i=1}^N c_{2i} T_{2i}(x) \quad (3.8)$$

where  $T_n$  are the classical Chebyshev polynomials of the first kind, all integrations in (3.3) may be exactly expressed in terms of the Chebyshev polynomials of the second kind. Consequently, Eq.

(3.3) has been numerically solved with the collocation method, by adopting the  $N$  positive zeroes of the Chebyshev polynomials of the second kind of degree  $2N$ . It can be shown that this collocation method is convergent in the considered weighted  $L^2$  space for variable  $F(\omega)$ , which in turn means that the  $k$ th derivative, with  $k=0,1,2$ , of the approximant of  $w(\theta)$  is uniformly convergent to the corresponding derivative of  $w(\theta)$ . This justifies the numerical approach employed to determine constants  $C_1$  and  $C_2$ , hence to prove the nonexistence results. The omitted details may be found in Monegato and Strozzi [8]. The choice  $N=4$  in (3.8) appears to be sufficient to guarantee three to four significant digits.

The numerically determined reaction force  $F(\omega)$  remains reasonably constant along the supported arcs, apart from the infinite peaks at the support extremities. Unfortunately, for  $0 < 2\alpha < \pi$  the plate supported edge arcs, when deflected by the above reaction force distribution, do not remain flat, as it would be requested by the title problem. In fact, constant  $C_2$  of Eq. (2.7), when computed from (2.8) by employing the unique reaction force  $F(\omega)$  which solves Eq. (2.5), does not vanish, but it remains positive and decreasing with  $\alpha$  for  $0 < 2\alpha < \pi$ . Consequently,  $w(\theta)$  does not stay constant for  $0 \leq \theta \leq \alpha$ , but it decreases with  $\theta$ . Figure 3 displays the value of coefficient  $C_2$  normalized over  $K = Pr_o^2/(2\pi D)$ , versus  $2\alpha/\pi$ , for  $\nu=0.3$ . For  $\alpha \rightarrow 0$ , coefficient  $C_2 \rightarrow \infty$  as  $\log \alpha$ , (Strozzi et al. [9]), whereas  $C_2$  vanishes only for  $2\alpha = \pi$ . As a consequence, for  $0 < 2\alpha < \pi$  the plate supported edge arcs, when deflected by the reaction force, exhibit a cosinusoidal profile.

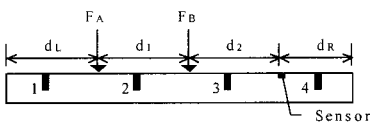
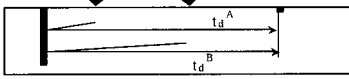
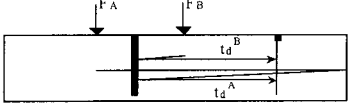
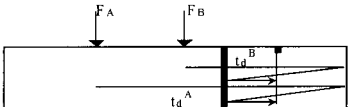
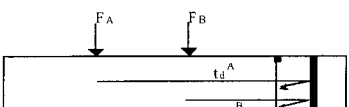
The above numerical results prove that reaction force distributions do exist which cause a cosinusoidal deflection of the plate loaded edge arcs. Conversely, such results demonstrate that no reaction force profile  $F(\omega)$  of form (3.1) and belonging to the weighted  $L^2$  space defined by (3.7) exists which keeps the plate supported edge arcs flat, for any angular width of the supports comprised in the interval  $0 < 2\alpha < \pi$ . Forms of the unknown function  $F(\omega)$  belonging to the weighted  $L^2$  space defined by (3.7) include, for example, functions possessing an endpoint singularity bounded by  $C(1-x^2)^\beta$ , where  $\beta > -3/4$ , and where variable  $x$  is defined by (3.2). Consequently, reaction forces with the above singularities do not constitute a solution to the title problem. For instance, a reaction force distribution  $F(\omega)$  exhibiting the classical square root singularities at the contact extremities,  $\beta = -1/2 > -3/4$ , encountered when a rigid rectangular punch indents a deformable half-space; Gladwell [7], cannot solve this plate problem. Similarly, algebraic/logarithmic singularities of type, say,  $(1-x^2)^{-1/2} \log(1-x^2)$  are unacceptable. This negative result is corroborated by an analogous conclusion reached by Grigolyuk and Tolkachev ([3], p. 381) with regard to an infinite plate resting on a line support. On the other side, functions  $F(\omega)$  possessing stronger singularities,  $\beta \leq -3/4$ , do not verify condition (3.7) and, consequently, are not considered in this study. It cannot therefore be excluded that a solution to the title problem might exist which exhibits such endpoint singularities. It cannot equally be excluded a priori that a solution  $F(\omega)$  of the equation  $w = C$  derived from (2.3) exists, which is of a more general form than the one considered here. For instance, Grigolyuk and Tolkachev ([3], p. 381) propose that the distributed reaction be formed by a distributed force superposed to a distributed couple. Notice also that for the case  $2\alpha = \pi$ , for which the plate is axisymmetrically supported, a solution to the title problem does exist, which consists of a constant reaction force. The same conclusions hold for any axisymmetric, non-self-equilibrated loading, for instance for a plate simply supported along two antipodal arcs and deflected by a uniform transverse pressure.


It is finally shown that the reaction force can neither be constituted by a distributed force and by two concentrated loads at the support extremities. In other words, the reaction force cannot be of the form  $F_0(\omega) + C[\delta(\omega - \alpha) + \delta(\omega + \alpha)]$ , where  $F_0$  is of type (3.1),  $\delta$  is the Dirac delta function, and where the presence of constant  $C$  aims at annulling constant  $C_2$  in (2.7). Indeed, such a function cannot satisfy Eq. (2.5) since

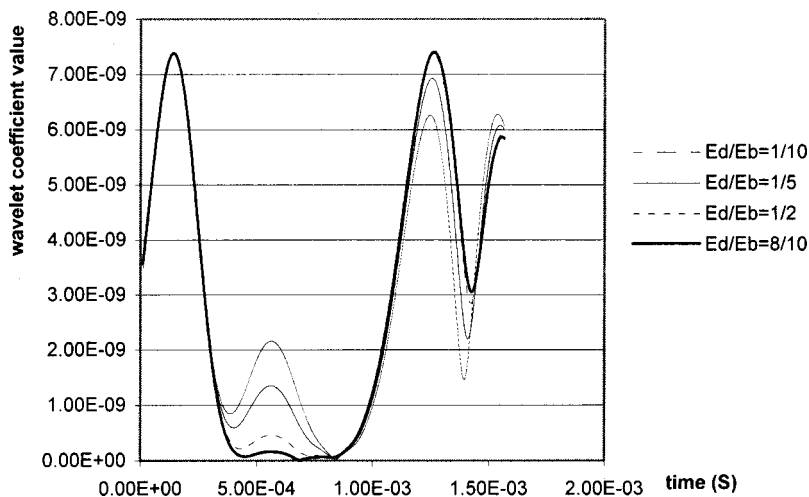


while all other terms in (2.5) give rise to smoother functions at  $\omega = \pm \alpha$ .

**Table 2** Detecting damage using signals from two loads (four possible damage positions)\*

Case		$t_d^A - t_d^B$	$t_d^A - t_d^B$	$t_R^A > t_d^A \& t_R^B > t_d^B$
1		$<0$	—	—
2		$>0$	$\frac{d_1 + 2d_2}{V_{h1}}$	—
3		$>0$	$\frac{d_1}{V_{h1}}$	No
4		$>0$	$\frac{d_1}{V_{h1}}$	Yes

\*Note:  is damage location



**Fig. 2** Distribution of Gabor wavelet coefficients at scale 13 for different degree of damage for sensor position A

$$x_d = \frac{x_A}{2t_{A1}} (t_{A1} + t_{A2}) \quad (2)$$

where  $t_{A1}$  and  $t_{A2}$  denote the time the direct and reflected wave reaches sensor A. For the case where the damage is of significant width, two additional times based on wave reflections are needed in order to estimate the width of the damage zone. Similar equations can be written to locate the damage location if the sensor is at either position B or C.

To obtain the arrival times of a wave at an appropriate frequency, Gabor wavelets ([3]) are used to transform the response

time history captured by the sensor. Since higher frequency waves are more suitable for damage detection using wave propagation analysis, the distributions of the wavelet coefficients at high scales are considered. The appropriate scale is chosen which has adequate peaks to signify the arrival times of waves at the sensor position. Based on experience, the first three to five peaks give reasonable results, corresponding to scale 13 for the examples presented in this note.

A fixed ended beam with a crack was numerically studied where the response data was computed from finite element analysis. The pertinent data of the beam are: width  $B = 20$  mm, thick-

ness  $H = 10$  mm, length  $L = 2$  m, modulus  $E = 206$  GPa, and actual damage location is 0.5 m. The results are shown in Table 1 indicating that the approach yields good results for the three sensor positions.

In practice, often no prior knowledge of the approximate position of the damage relative to the sensor and loading positions is available. In this case, either additional time values caused by the wave reflections from the damage are used or a second load  $F_B$  is considered. The former is not preferred due to complications arising from multiple wave reflections. Using  $F_B$  and based on simple wave propagation, it is shown in Table 2 that the location of the damage can be estimated, where the average velocity of wave propagation can be computed as

$$V_{h1} = \frac{d_1 + d_2}{t_1^A} = \frac{d_2}{t_1^B} \quad (3)$$

Once the relative position of the damage is known, the distance between the damage and load can be computed as described earlier. The same procedure can be easily extended to locate the loading position if the latter is not known.

Numerical simulations were performed to study the effect of damage height  $c$  on detectability. Results for  $c/H = 0.4, 0.6, 0.8$ , and 1.0 indicate that the damage locations can be estimated to within a one percent error.

The damage above was simulated by setting the Young's modulus of the damage zone to one-tenth that of the undamaged beam. The wavelet coefficients at scale 13 are presented in Fig. 2 with the damage zone having Young's moduli of one-tenth, one-fifth, half, and eight-tenths of the undamaged beam. As the severity of damage increases, the lower peak in the wavelet coefficient curve corresponding to the damage becomes more significant. The computed positions for the four cases are not significantly different, ranging from 0.505 m for the one-tenth case to 0.500 m for the eight-tenth case.

The detectability is compared between simply supported (SS) and fixed ended (FF) boundary conditions. The SS case yields a 4.2 percent error in the computed damage position compared to 0.2 percent for the FF case. The probable reason is due to the vertical boundaries at the end where, for the FF case, it is uniformly constrained. For the SS case, only the lower corner is constrained and the other nodes are free. Hence, the reflected wave in the latter case is a composite wave but computationally the averaged wave speed is used.

### 3 Concluding Remarks

The applicability of the wavelet transform to damage detection has been examined where the Gabor wavelet is used to analyze the time history data caused by thin damage. For the case where the damage position is completely unknown, the use of two load positions is suggested. The results of the sensitivity of the method in terms of the height of damage, severity of the damage and boundary conditions indicate the promise of this approach.

### References

- [1] Surace, C., and Ruotolo, R., 1994, "Crack Detection of a Beam Using the Wavelet Transform," *Proceedings of the 12th International Modal Analysis Conference*, Honolulu, pp. 1141–1147.
- [2] Wang, Q., and Deng, X., 1999, "Damage Detection With Spatial Wavelets," *Int. J. Solids Struct.*, **36**, pp. 3443–3468.
- [3] Kishimoto, K., Inoue, H., Hamada, M., and Shibuya, T., 1995, "Time Frequency Analysis of Dispersive Waves by Means of Wavelet Transform," *ASME J. Appl. Mech.*, **62**, pp. 841–846.
- [4] Graff, K. F., 1975, *Wave Motion in Elastic Solids*, Oxford University Press, London.

## A Novel Technique in the Solution of Axisymmetric Large Deflection Analysis of a Circular Plate

L. S. Ramachandra<sup>1</sup> and D. Roy

Department of Civil Engineering, Indian Institute of Technology, Kharagpur 721 302, India

*In the present paper a new linearization technique referred to as the locally transversal linearization (LTL) is used for large deflection analyses of axisymmetric circular plates. The LTL procedure, where solution manifolds of linearized equations are made to intersect transversally those of the nonlinear ordinary differential equations, reduces the given set of nonlinear ordinary differential equations to a set of nonlinear algebraic equations in terms of a discretized set of unknown response vectors.*

[DOI: 10.1115/1.1379039]

### 1 Introduction

It is usually not possible to replace the nonlinear governing differential equations in terms of some linearized equations even over small step sizes, as the tangent spaces of the nonlinear differential equations are space dependent. A way out of this problem is, however, found in the proposed locally transversal linearization (LTL) method, wherein solution manifolds of the linearized equations are made to intersect transversally those of the given nonlinear equations at a set of pre-selected points along the independent axis where the solution vectors need to be determined. The method, thus, attempts to satisfy exactly the governing nonlinear ordinary differential equations at these pre-selected points. The LTL-based solution may not follow the actual path in between any two such successive points of intersections. The method finally reduces the given non-linear ordinary differential equations to a collection of coupled sets of nonlinear algebraic equations in terms of the unknown solution vectors at each of the chosen points. In present study, a boundary value problem is treated as a constrained initial value problem, where one of the boundaries (say, the left boundary) of the domain is treated as the initial point to start integrating the ordinary differential equations; however, all the state variables at the left boundary are not known a priori. To start with, only a few of these conditions are known at the left boundary while the rest are specified at the right boundary of the domain. In this study, the unknown initial conditions at the left boundary are treated as unknown variables to be determined such that the resulting solution satisfies prescribed boundary conditions at the right boundary. The method is adequately described elsewhere ([1]) and, hence, not reproduced here. The large deflection axisymmetric analysis of circular plate is carried out using the proposed LTL technique. Comparisons are made with the power series solutions and other approximate solutions and in the process, high numerical accuracy of the proposed method is readily brought out.

### 2 Numerical Examples

The equations relevant to the axisymmetric large deflection analysis of circular plate may be written in terms of in-plane ( $u$ ) and out-of-plane ( $w$ ) displacements as ([2])

<sup>1</sup>To whom correspondence should be addressed. e-mail: lsr@civil.iitkgp.ernet.in.

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. 1, 2000; final revision, Mar. 13, 2001. Associate Editor: A. K. Mal.

$$\frac{d^2 u}{dr^2} + \frac{1}{r} \frac{du}{dr} - \frac{u}{r^2} = -\frac{(1-\nu)}{2r} \left( \frac{dw}{dr} \right)^2 - \frac{dw}{dr} \frac{d^2 w}{dr^2} \quad (1)$$

$$\begin{aligned} \frac{d^3 w}{dr^3} + \frac{1}{r} \frac{d^2 w}{dr^2} - \frac{1}{r^2} \frac{dw}{dr} = \frac{12}{h^2} \frac{dw}{dr} \left[ \frac{du}{dr} + \nu \frac{u}{r} + \frac{1}{2} \left( \frac{dw}{dr} \right)^2 \right] \\ + \frac{1}{Dr} \int_0^r q r dr \end{aligned} \quad (2)$$

where, “ $h$ ” is the thickness,  $\nu$  the Poisson’s ratio,  $D$  the flexural rigidity, “ $q$ ” intensity of lateral load, and “ $a$ ” the radius of the plate. The associated boundary conditions are

$$u(r=0)=0$$

and

$$\frac{dw}{dr}(r=0)=0$$

$$u(r=a)=0; \quad w(r=a)=0;$$

and

$$\frac{dw}{dr}(r=a)=0 \quad \text{for clamped} \quad (3)$$

$$u(r=a)=0; \quad w(r=a)=0;$$

and

$$M_r(r=a)=0 \quad \text{for simply supported}$$

To solve the above boundary value problem, the following LTL equations over the  $i$ th subdomain  $D_i$  and with initial conditions  $\{\bar{u}_{i-1}^{(i-1)}, d\bar{u}_{i-1}^{(i-1)}/dr=u_{i-1}'^{(i-1)}, \bar{w}_{i-1}^{(i-1)}, d\bar{w}_{i-1}^{(i-1)}/dr=\bar{w}_{i-1}'^{(i-1)}, d^2\bar{w}_{i-1}^{(i-1)}/dr^2=\bar{w}_{i-1}''^{(i-1)}\}$  are employed (superscripts stand for the segment number and subscripts denote the node number).

$$\frac{d^2 \bar{u}^{(i)}}{dr^2} = \beta_i \quad (4)$$

$$\frac{d^3 \bar{w}^{(i)}}{dr^3} = \gamma_i \quad (5)$$

A complete solution of Eqs. (4) and (5) may be easily written and the arbitrary constants of integration are evaluated via initial conditions:

$$\bar{u}_{i-1} = u_{i-1}; \quad \bar{u}_{i-1}' = u_{i-1}'; \quad \bar{w}_{i-1} = w_{i-1}; \quad \bar{w}_{i-1}' = w_{i-1}'$$

and

$$\bar{w}_{i-1}'' = w_{i-1}'' \quad (6)$$

After obtaining the expressions for  $\bar{u}(r)$  and  $\bar{w}(r)$ , the analytical expressions for  $\bar{u}'(r), \bar{u}''(r), \bar{w}'(r), \bar{w}''(r)$ , and  $\bar{w}'''(r)$  in terms of the unknown solution parameters  $\beta_i$  and  $\gamma_i$  may be obtained. Substituting  $\bar{u}'(r), \bar{u}''(r), \bar{w}'(r), \bar{w}''(r)$ , and  $\bar{w}'''(r)$  in the original Eqs. (1) and (2) at  $r=r_i$ , two coupled nonlinear algebraic equations in  $\beta_i$  and  $\gamma_i$  are obtained. In a similar way one may continue obtaining nonlinear equations for the conditionally constant parameters  $\beta_j$  and  $\gamma_j$  ( $j=1, 2, \dots, n$ ) at all the intersection points. This results in “ $2n$ ” nonlinear algebraic equations in  $2n+3$  unknown parameters  $\beta_j$  and  $\gamma_j$  ( $j=1, 2, \dots, n$ ) and  $u'(r=0), w(r=0), w''(r=0)$ . It may, however, be observed that for the present problem  $u, w$ , and  $w'$  or  $M_r$  are known at the right end, i.e.,  $r=a$ . Thus, in the conditional initial value problem approach, the unknowns  $u'(r=0), w(r=0), w''(r=0)$  are to be so determined as to result in a constrained dynamical trajectory which satisfies the boundary conditions at  $r=a$ . This constraint condition leads to the required additional equations via the analytical expression available for  $\bar{u}^{(n)}(r), \bar{w}^{(n)}(r)$  and  $\bar{w}'^{(n)}(r)$ . Since equations are coupled, one has to solve all the nonlinear

algebraic equations simultaneously in order to obtain the equilibrium path of the structure. In the present study, a globally convergent nonlinear equation solver based on line searches and backtracking along the Newton directions (see [3],) has been adopted for solving the system of the nonlinear algebraic equations.

### 3 Results and Discussion

The aforementioned method has been used to obtain the elastic, large deflection response of the uniformly loaded circular plate with various boundary conditions. In the present case, to avoid singularity at the center, a central small hole (the ratio of hole to plate radii is 0.005) is considered. While the deflections are evaluated at the edge of the hole, the moments and stress resultants are calculated slightly inside the plate, i.e., at a distance of “ $0.01a$ ” from the center. Good convergence of deflection at the free edge has been observed and they are not presented here for the sake of brevity. The results are presented taking 40 divisions along the radius.

Way [4] has obtained exact results for the large deflection analysis of a clamped circular plate. The present results are compared with that of Way’s results in Figs. 1 and 2.

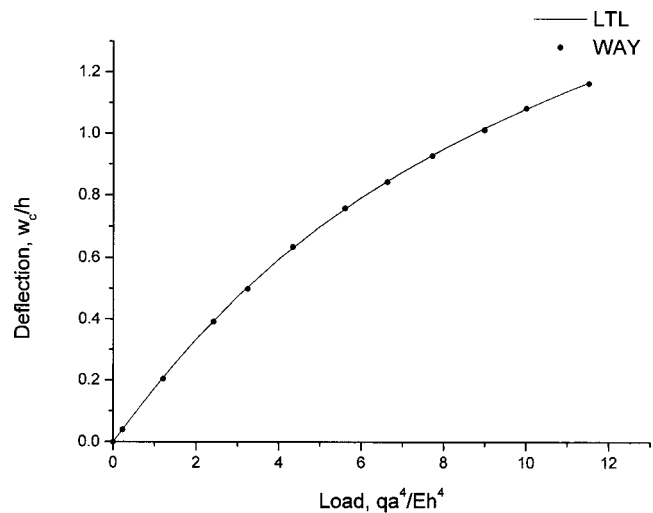


Fig. 1 Load-deflection curve for the uniformly loaded circular clamped plate

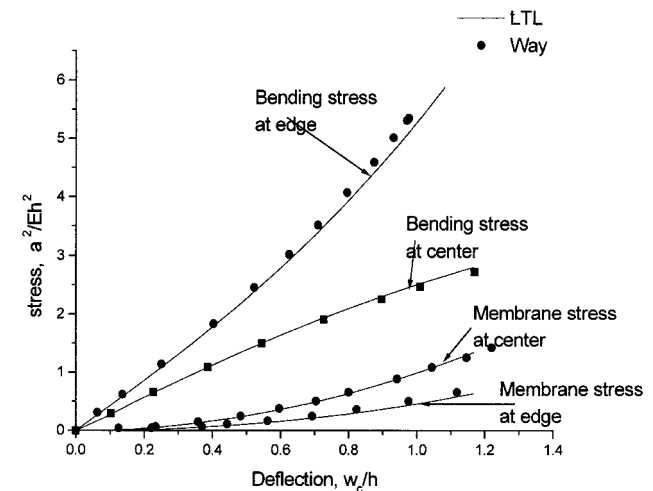


Fig. 2 Load-membrane/bending stresses curves for the uniformly loaded circular clamped plate



**Table 1 Large deflection of uniformly loaded simply supported circular plate ( $\nu=0.25$ ): comparison of locally transversal linearization (LTL) and exact results**

Results obtained from LTL				Federhofer and Egger		
$qa^4/Eh^4$	$w_c/h$	$\sigma_r^m a^2/Eh^2$ (Edge)	$\sigma_t^b a^2/Eh^2$ (Edge)	$w_c/h$	$\sigma_r^m a^2/Eh^2$ (Edge)	$\sigma_t^b a^2/Eh^2$ (Edge)
3.07	0.884	0.478	0.715	0.882	0.469	0.714
7.18	1.247	0.984	1.052	1.245	0.967	1.051
25.66	1.966	2.595	1.780	1.965	2.544	1.780
102.64	3.132	6.948	3.046	3.136	6.798	3.042
205.28	3.952	11.083	3.954	3.953	10.943	3.933
307.92	4.522	14.625	4.593	4.526	14.432	4.557

$q$ —is the intensity of uniformly distributed lateral load

$w_c$ —central deflection

$E$ —Young's modulus;  $\sigma_r^m$ —radial membrane stress;  $\sigma_t^b$ —tangential bending stress

It is observed that the results compare well. The LTL results are compared with the exact solutions given by Federhofer and Egger [5] for the case of simply supported plate in Table 1. Even for very high nonlinearity, the deflections obtained from the present method compare well with the Federhofer and Egger. However, in the case of stresses and moments the comparison is not that good. This may be due to the reason that the exact method employ insufficient terms in the representation of the displacements. However, in the LTL method such problems do not arise as the nonlinear equations are exactly satisfied.

## References

- [1] Ramachandra, L. S., and Roy, D., 2000, "A New Method for Non-linear Two-Point Boundary Value Problems in Solid Mechanics," ASME J. Appl. Mech., in press.
- [2] Timoshenko, S. P., and Woinowsky-Krieger, S., 1959, *Theory of Elastic Plates and Shells*, 2nd Ed., McGraw-Hill, New York.
- [3] Ortega, J. M., and Rheinboldt, W. C., 1970, *Iterative Solutions of Non-Linear Equations in Several Variables*, Academic Press, San Diego, CA.
- [4] Way, S., 1934, "Bending of Circular Plates with Large Deflections," ASME J. Appl. Mech., **56**, pp. 627–636.
- [5] Federhofer, K., and Egger, H., 1946, "Berechnung der dünnen Kreisplatte mit grosser Ausbiegung," Sitzungsber. Akad. Wiss. Wien, Math.-Naturwiss. Kl., Abt. 2B, **155**, No. 2a, pp. 15–43.

## On Characteristic Times in Generalized Thermoelasticity

### D. V. Strunin

Department of Mathematics and Computing,  
University of Southern Queensland,  
Toowoomba, QLD 4350, Australia

*The model of Green and Lindsay is a popular generalization of the theory of thermoelasticity incorporating second sound. Within the model the second sound is intimately linked to a presence of two characteristic times,  $t_1$  and  $t_2$ , constrained by an inequality  $t_2 \leq t_1$ . We present a modification of the theory where no constraints on the times arise.* [DOI: 10.1115/1.1386696]

The classical Fourier law of heat conduction leads to instantaneous propagation of heat to infinitely remote areas of space. This paradox is traditionally surmounted by describing the heat signal as a wave, called second sound. One of the most popular gener-

alizations of the thermoelasticity theory allowing for the second sound is the model of Green and Lindsay (GL) ([1]). It contains two characteristic times: one,  $t_1$ , in the momentum equation and the other, directly responsible for the second sound time,  $t_2$ , in the energy equation. The model asserts (see also [2]) that

$$t_2 \leq t_1. \quad (1)$$

Due to lack of measurements, theoretical knowledge of  $t_1$  and  $t_2$  is highly desirable. So far condition (1) was exploited, e.g., in [3], to estimate a lower limit for  $t_1$  on the basis of an independent theoretical estimate for  $t_2$ . Inequality (1) also has qualitative consequence: It is prohibitive for spontaneous thermomechanical instability. This implies that in an isolated body which is not subjected to external forces and heat sources, small deformations and small temperature gradients cannot grow.

Here we formulate a modified version of the GL model, where the condition (1) does not appear. First, consider an entropy production inequality ([4,5]) which underlies the GL model.

$$\begin{aligned} & -\rho_0 \left( \frac{\partial \psi}{\partial \theta} + \eta \frac{\partial \phi}{\partial \theta} \right) \dot{\theta} - \rho_0 \left( \frac{\partial \psi}{\partial \theta} + \eta \frac{\partial \phi}{\partial \theta} \right) \ddot{\theta} - \rho_0 \left( \frac{\partial \psi}{\partial \theta_{,A}} + \eta \frac{\partial \phi}{\partial \theta_{,A}} \right) \dot{\theta}_{,A} \\ & + \left[ s_{AB} - \frac{1}{2} \rho_0 \left( \frac{\partial \psi}{\partial e_{AB}} + \frac{\partial \psi}{\partial e_{BA}} \right) - \frac{1}{2} \rho_0 \eta \left( \frac{\partial \phi}{\partial e_{AB}} + \frac{\partial \phi}{\partial e_{BA}} \right) \right] \dot{e}_{AB} \\ & - \frac{Q_A}{\phi} \left[ \frac{\partial \phi}{\partial \theta} \theta_{,A} + \frac{\partial \phi}{\partial \theta_{,B}} \theta_{,BA} + \frac{\partial \phi}{\partial \dot{\theta}} \dot{\theta}_{,A} + \frac{1}{2} \left( \frac{\partial \phi}{\partial e_{MN}} + \frac{\partial \phi}{\partial e_{NM}} \right) \right. \\ & \left. \times e_{MN,A} + \frac{\partial \phi}{\partial \rho_0} \rho_{0,A} + \left( \frac{\partial \phi}{\partial X_A} \right)' \right] \geq 0, \end{aligned} \quad (2)$$

where  $X_A$  are the reference coordinates;  $\theta$  is the deviation of temperature from some basis temperature,  $\theta_0$ ;  $\rho_0$  is the density; a superposed dot denotes differentiation with respect to time,  $t$ ; prime denotes partial differentiation holding all independent variables except  $X_A$  fixed;  $_{,A}$  denotes partial differentiation with respect to  $X_A$ ;  $Q_A$  is the heat flux per unit area of the  $X_A$ -plane; tensor  $s_{AB}$  is represented through the Cauchy stress tensor  $t_{ik}$  by  $(\det x_{i,A}) t_{ik} = x_i x_k s_{AB}$ ;  $x_i$  are spatial coordinates; and  $e_{AB} = (1/2)(x_{r,A} x_{r,B} - \delta_{AB})$ . The value  $Q_A$  can be expressed through the usual heat flux,  $q_i$ , by  $(\det x_{i,A}) q_i = x_{i,A} Q_A$ . Further, the function  $\psi$  is related to  $s_{AB}$  via  $s_{AB} = (\rho_0/2)(\partial \psi / \partial e_{AB} + \partial \psi / \partial e_{BA})$ ; and the function  $\phi$  is connected to  $\psi$ , specific internal energy,  $\varepsilon$ , and specific entropy,  $\eta$ , via  $\psi = \varepsilon - \eta \phi$ . To satisfy (2) Green and Lindsay chose to set to zero several terms in (2), leading to six equations, one of which is

$$\rho_0 \left( \frac{\partial \psi}{\partial \theta} + \eta \frac{\partial \phi}{\partial \theta} \right) \ddot{\theta} = 0, \quad (3)$$

so that a sum of only a few terms of (2) is non-negative:

$$-\rho_0 \left( \frac{\partial \psi}{\partial \theta} + \eta \frac{\partial \phi}{\partial \theta} \right) \dot{\theta} - \frac{Q_A}{\phi} \left[ \frac{\partial \phi}{\partial \theta} \theta_{,A} + \left( \frac{\partial \phi}{\partial X_A} \right)' \right] \geq 0. \quad (4)$$

Consider small temperature deviations  $\theta$  and displacements  $u_i$  ( $u_i = x_i - X_i$ ). Then, assuming the function  $\psi$  to have the form  $\psi = \psi(\theta, \dot{\theta}, \theta_{,i}, e_{ij})$ , where  $e_{ij} = (1/2)(u_{i,k} + u_{k,i})$ , and the function  $\phi$  to have the form  $\phi = \phi(\theta, \dot{\theta})$  and going over from the variables  $X_A$  to the variables  $x_i$ , one can expand these functions into the Taylor series:

$$\begin{aligned} \rho_0 \psi = & -a\theta - b\dot{\theta} - \frac{1}{2}d\theta^2 - e\theta\dot{\theta} - \frac{1}{2}f\dot{\theta}^2 + a_{ijk}\theta_{,i}\theta_{,j}\theta_{,k} \\ & + a_{ik}e_{ik}\theta + b_{ik}e_{ik}\dot{\theta} + \frac{1}{2}\alpha k_{rs}\theta_{,r}\theta_{,s} + \frac{1}{2}k_{ikrs}e_{ik}e_{rs}, \end{aligned} \quad (5)$$

$$\phi = \theta_0 + \theta + \alpha\dot{\theta} + \beta\theta\dot{\theta} + \frac{1}{2}\gamma\dot{\theta}^2.$$

Using (3), (4), and (5), Green and Lindsay eventually reduced (we omit details) the entropy inequality to the form

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 7, 2000; final revision, May 1, 2001. Associate Editor: D. A. Siginer.



$$(d\alpha - h)\ddot{\theta}^2 + k_{ij}\theta_i\theta_j \geq 0, \quad (6)$$

where by definition  $h = \alpha f - b\gamma/\alpha^2$ . Equation (6) must hold for any physically possible temperature field. Therefore, taking into account that the thermal conductivity tensor  $k_{ij}$  is positive-definite, (6) is met provided that  $h/d \leq \alpha$  which is just the relation (1) where  $t_2 = h/d$ ,  $t_1 = \alpha$ .

Let us consider different way of satisfying the entropy inequality by exchanging the terms between (3) and (4) as follows:

$$\rho_0 \left( \frac{\partial \psi}{\partial \theta} + \eta \frac{\partial \phi}{\partial \theta} \right) \dot{\theta} = 0, \quad (7)$$

$$-\rho_0 \left( \frac{\partial \psi}{\partial \theta} + \eta \frac{\partial \phi}{\partial \theta} \right) \ddot{\theta} - \frac{Q_A}{\phi} \left[ \frac{\partial \phi}{\partial \theta} \theta_{,A} + \left( \frac{\partial \phi}{\partial X_A} \right)' \right] \geq 0. \quad (8)$$

The rest of the five equations are identical to those in the original GL model ([1]). Using (7) and (8) we transform (2) to

$$\begin{aligned} (b - a\alpha)\ddot{\theta} + (e - d\alpha - a\beta)\ddot{\theta}\dot{\theta} + (f - e\alpha - a\gamma)\ddot{\theta}\dot{\theta} \\ + a_i\alpha\ddot{\theta}\theta_{,i} + (-b_{ik} + a_{ik}\alpha)\ddot{\theta}e_{ik} + \frac{a_i}{\alpha}\theta\theta_{,i} \\ + \frac{a_{rsi}}{\alpha}e_{rs}\theta_{,i} + k_{ij}\theta_{,i}\theta_{,j} \geq 0. \end{aligned} \quad (9)$$

Equation (9) must hold for all physically possible temperature fields and thus similarly to [1] we require

$$\begin{aligned} b = a\alpha, \quad d = e - a\beta, \quad f = e\alpha + a\gamma, \quad a_i = 0, \\ b_{ik} = a_{ik}\alpha, \quad a_{rsi} = 0. \end{aligned} \quad (10)$$

The entropy inequality (9) then becomes

$$k_{ij}\theta_{,i}\theta_{,j} \geq 0. \quad (11)$$

Relation (11) imposes no restriction on the times,  $t_1$  and  $t_2$ , as  $k_{ij}$  is positive-definite. It can be shown that energy and momentum equations have virtually the same form as in the original model with the only difference being another representation of the coefficient of the thermal relaxation term  $\ddot{\theta}$ .

Consider a thermomechanical system in the form of a closed wire (ring). Supposing a curvature small and neglecting its affects on dynamics write the governing momentum and energy equations for longitudinal thermomechanical disturbances in nondimensional form

$$\begin{aligned} \ddot{u} - u_{xx} - \beta\theta_x - \beta\tau_1\dot{\theta}_x = 0, \\ \tau_2\ddot{\theta} + \dot{\theta} - \chi\dot{u}_x - \theta_{xx} = 0, \end{aligned} \quad (12)$$

where  $\beta$  and  $\chi$  are positive thermomechanical constants,  $\tau_1$  and  $\tau_2$  are the nondimensionalized times,  $t_1$  and  $t_2$ , respectively. As the model is linear, it will suffice to study a behavior of harmonic disturbances,  $u = U \exp(\omega t + ikx)$  and  $\theta = \Theta \exp(\omega t + ikx)$ , where the wave number,  $k = 2\pi/L$  ( $L$  is the perimeter of the wire), is a real number and the frequency,  $\omega$ , is a complex-valued function of  $k$ . Although straightforward our linear analysis is different from the recent analysis ([6]) performed in the spirit of [7] where  $k$  was assumed to be a complex function of real  $\omega$ . Equations (12) result in the characteristic equation

$$\tau_2\omega^4 + \omega^3 + (\tau_2 + 1 + \beta\chi\tau_1)k^2\omega^2 + (1 + \beta\chi)k^2\omega + k^4 = 0. \quad (13)$$

Analysis ([8]) of (13) shows that if  $S \equiv (\tau_2 - \tau_1)(1 + \chi\beta) < 1$  then  $\text{Re } \omega < 0$ , that is the disturbances decay and the reverse if  $S > 1$  then  $\text{Re } \omega > 0$ , so that the disturbances grow. Within the original GL theory  $\tau_2 \leq \tau_1$ ,  $S < 0 < 1$ , and hence the growth of displacement and temperature is impossible. In the modified model the value of  $S$  in principle is not restricted (although due to smallness of  $\tau_1$  and  $\tau_2$  it may still be less than 1 even when  $\tau_2 > \tau_1$ ). From theoretical point of view the possibility of instability presents considerable interest. Remark that balance equation for total energy, following from equations of motion and thermal energy transfer, is met regardless of specific form of solution, be it growing or decaying.

The above consideration reveals a possibility of new formulation of the GL model. Which of the two models (the original or modified) is preferable is not clear so far. We view this question as a part of wider issue concerning the selection of the most realistic model of generalized thermoelasticity. Alongside with the GL model there are other thermoelasticity models allowing for the second sound, for instance the known Lord-Schulman (LS) model ([9]). Unlike the GL theory it contains only one characteristic time, namely the thermal relaxation time which is analogous to  $t_2$ . The analysis ([8]) showed that in the LS model spontaneous instability is impossible so that small disturbances decay. But decay rates in the LS and the GL models are different in the general case. It seems that in the final analysis the selection of the most adequate model has to be based on how well it agrees with experiments. The (hypothetical) spontaneous instability could be a tempting target for experimental detection. If detected for some material the effect would provide a convincing argument in favor of our version of the GL model. If not, other experimental evidences will be necessary to support particular model.

## References

- [1] Green, A. E., and Lindsay, K. A., 1972, "Thermoelasticity," J. Elast. **2**, pp. 1–7.
- [2] Chandrasekharaiah, D. S., 1986, "Thermoelasticity With Second Sound: A Review," Appl. Mech. Rev., **39**, pp. 355–376.
- [3] Prevost, J. H., and Tao, D., 1983, "Finite Element Analysis of Dynamic Coupled Thermoelasticity Problems With Relaxation Times," J. Appl. Mech., **50**, pp. 817–822.
- [4] Müller, I., 1971, "The Coldness, A Universal Function in Thermoelastic Bodies," Arch. Ration. Mech. Anal., **41**, pp. 319–332.
- [5] Green, A. E., and Laws, N., 1972, "On The Entropy Production Inequality," Arch. Ration. Mech. Anal., **45**, pp. 47–53.
- [6] Suh, C. S., and Burger, C. P., 1998, "Effects of Thermomechanical Coupling and Relaxation Times on Wave Spectrum in Dynamic Theory of Generalized Thermoelasticity," J. Appl. Mech., **65**, pp. 605–613.
- [7] Nowacki, W., 1962, *Thermoelasticity*, Pergamon, Oxford, UK.
- [8] Strunin, D. V., Melnik, R. V. N., and Roberts, A. J., 2001, "Coupled Thermo-mechanical Waves in Hyperbolic Thermoelasticity," J. Therm. Stresses, **24**, pp. 121–140.
- [9] Lord, H. W., and Schulman, Y., 1967, "A Generalized Dynamic Theory of Thermoelasticity," J. Mech. Phys. Solids, **15**, pp. 299–309.

## Closed-Form Solutions for the Stress Fields Induced by Blunt Wedge-Shaped Indenters in Elastic Half-Planes

C. E. Truman<sup>1</sup>

Department of Mechanical Engineering, University of Bristol, Queen's Building, University Walk, Bristol BS8 1TR, UK. Mem. ASME

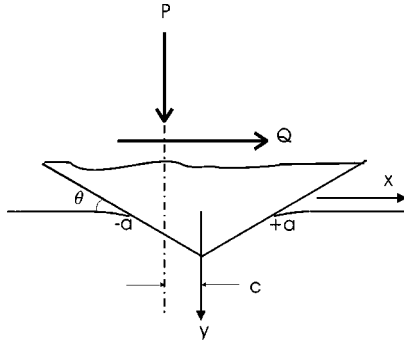
A. Sackfield

Department of Mathematics, Nottingham Trent University, Burton Street, Nottingham, NG1 4BU, UK

*Closed-form expressions are given for the Muskhelishvili potentials created by wedge-shaped indenters contacting elastic half-planes. The potentials are given for normal and sliding contact of both similar and dissimilar materials. Surface values of the tension  $\sigma_{xx}^0$  are also presented.* [DOI: 10.1115/1.1386697]

<sup>1</sup>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, January 1, 2001; final revision, April 17, 2001. Associate Editor: J. R. Barber.



**Fig. 1 The geometry of the problem. A wedge loaded by a normal force,  $P$ , and a tangential force,  $Q$ . When the materials are similar  $c=0$ .**

## 1 Introduction

This paper concludes a series of work which examined the state of stress caused by wedge-shaped indenters penetrating elastic half-planes ([1–3]). Several reasons motivated the study. For power-law indenters which have profiles of the form  $r^\lambda$ , solutions are well known for  $\lambda=0, 2$  and higher even values. The solution for  $\lambda=1$ , to be discussed in this note, is currently only available in specialized conditions which will be mentioned. As well as being a fundamental geometry there has been a need for the stress fields created by such indenters for use in fretting fatigue tests ([4]). Traditionally Hertzian indenters ( $\lambda=2$ ) have been used, but in order to validate crack initiation theories and wear models other geometries have also been used in order to compare results. The simplest such geometry is the wedge, depicted in Fig. 1, where the load is shown applied with a small offset  $c$  from the center of the indenter. The case  $c=0$  corresponds to the case when the indenter and substrate are made of similar materials or the coefficient of friction is zero.

This basic configuration can also be used to model asperities ([5]) or study the motion of styli over disks, although it should be pointed out the analysis uses first-order theory so only applies to wedge angles of greater than about 160 deg.

## 2 Formulation

The basic equations, linking the stresses  $\sigma_{ij}$  with the Muskhelishvili potential  $\Phi$  are ([6,7])

$$\sigma_{xx} + \sigma_{yy} = 2[\Phi(z) + \overline{\Phi(\bar{z})}]$$

$$\sigma_{yy} - \sigma_{xx} + 2i\sigma_{xy} = 2[(\bar{z}-z)\Phi'(z) - \bar{\Phi}(z) - \Phi(z)] \quad (1)$$

where

$$\Phi(z) = \frac{1}{2\pi i} \int_{-1}^1 \frac{p(x) - iq(x)}{x-z} dx. \quad (2)$$

$p(x)$  is the contact pressure distribution,  $q(x)$  is the shear traction distribution which, for the sliding case,  $q(x)=fp(x)$ , with  $f$  being the coefficient of friction and an overbar representing complex conjugation. In what follows, all variables are considered normalized with respect to the contact half-width  $a$ .

## 3 Interior Stress Field

**3.1 Similar Materials.** For the wedge indenter, shown in Fig. 1, it can be shown that for similar materials (or, more specifically, if any shear tractions do not influence the normal pressure distribution) ([7])

$$p(x) = p_0 \cosh^{-1}(1/|x|), \quad |x| \leq 1 \quad (3)$$

where  $p_0 = 2\theta/\pi A$ , and  $A$  is the composite compliance, given by

$$A = \frac{\kappa_1 + 1}{4\mu_1} + \frac{\kappa_2 + 1}{4\mu_2}$$

with  $\kappa_i = 3 - 4\nu_i$  in plane strain and  $\kappa_i = (3 - \nu_i)/(1 + \nu_i)$  in plane stress and  $\nu_i$  and  $\mu_i$  representing Poisson's ratio of body  $i$  and the modulus of rigidity of body  $i$ , respectively. By substituting Eq. (3) into Eq. (2) it is found

$$\Phi(z) = \frac{ip_0}{2} (1 - if) \sin^{-1}\left(\frac{1}{z}\right). \quad (4)$$

This can be compared with the result obtained in an earlier study ([1]), namely

$$\Phi(z) = ip_0(1 - if) \sum_{n=1,2}^{\infty} \frac{(-1)^{(n-1)/2}}{n} (z - \sqrt{z^2 - 1})^n$$

where the notation  $n=1, 2$  implies that summation is over odd values of  $n$  only. It is straightforward to sum this series and show agreement with Eq. (4). A more general result for the Muskhelishvili potential for stick-slip of similar materials has been presented by Jäger [8]. The removal of the summation sign greatly increases the usefulness of the solution. As described in [6,7] it is now routine to determine the stress components from Eqs. (1).

**3.1.1 Surface Values of Stress.** The tension in the surface,  $\sigma_{xx}^0$ , is a useful quantity to know as it is one parameter which governs the growth of surface cracks, caused for instance by fretting. From Eq. (4) and Eqs. (1) this is determined as

$$\begin{aligned} \sigma_{xx}^0 &= p_0 \left\{ f\pi \operatorname{sgn}(x) + \ln \left( \frac{1 + \sqrt{1 - x^2}}{|x|} \right) \right\}, \quad |x| < 1 \\ &= fp_0 \{ \pi - 2 \tan^{-1}(\sqrt{x^2 - 1}) \}, \quad x > 1 \end{aligned} \quad (5)$$

where it can be seen that the stress field is logarithmically singular.

## 3.2 Dissimilar Materials

**3.2.1 Normal Contact.** In the section above, covering the case of similar materials, the normal and sliding cases were not treated separately. This was because, for similar materials, under normal indentation only, no shear tractions arise. However, for dissimilar materials, even under purely normal indentation, shear tractions arise because particles in each surface undergo differing amounts of tangential displacement depending upon their material constants. In this section the stress fields generated in the full slip regime will be found, i.e., it will be assumed that the shear traction is everywhere limited by Coulomb's law but is an odd function of position, i.e.,  $q(x) = -fp(x)$  for  $x < 0$  and  $q(x) = fp(x)$  for  $x > 0$ . It should be remarked that Spence [9] showed that there will always be a finite stick zone, unless the coefficient of friction is identically zero, but the solution to be obtained is a good approximation for low values of friction. For the case of normal indentation only, the shear traction is assumed not to influence the normal traction, an often-used approximation associated with Goodman [10]. In other words the problem is assumed uncoupled in the normal and shear tractions. This assumption is relaxed for the sliding case treated later.

With these assumptions an extra term to be combined with Eq. (4) evaluated with  $f=0$  is found to be

$$\Phi_d(z) = \frac{fp_0}{2\pi} \left\{ \left( \sin^{-1}\left(\frac{1}{z}\right) \right)^2 \right\} \quad (6)$$

which can be compared with the result shown in [2], which is omitted here for brevity, and the simplicity of the new result is obvious.

**Surface values of stress.** By combining Eqs. (4) and (6) the Muskhelishvili potential is found to be

$$\Phi(z) = \frac{ip_0}{2} \left[ \sin^{-1} \left( \frac{1}{z} \right) - \frac{if}{\pi} \left\{ \sin^{-1} \left( \frac{1}{z} \right) \right\}^2 \right]. \quad (7)$$

The surface value of the tension is then

$$\begin{aligned} \sigma_{xx}^0 &= \frac{p_0}{\pi} \left\{ \pi \ln \left[ \frac{1 + \sqrt{1-x^2}}{|x|} \right] \right. \\ &\quad \left. + 2f \left( \frac{\pi^2}{4} - \left( \ln \left[ \frac{1 + \sqrt{1-x^2}}{|x|} \right] \right)^2 \right) \right\}, \quad |x| < 1 \\ &= \frac{fp_0}{\pi} \left\{ \sin^{-1} \left( \frac{1}{x} \right) \right\}^2, \quad x > 1 \end{aligned} \quad (8)$$

which, again, is logarithmically singular.

**3.2.2 Sliding Contact.** The final problem consists of determining the Muskhelishvili potential for full sliding contact of two dissimilar materials. The problem is considered fully coupled in the sense that the pressure distribution influences the shear traction and *vice versa*. The pressure distribution arising for this configuration is given by ([3])

$$p(x) = \begin{cases} -\frac{2}{\pi} p_0 \frac{\sin^2(\pi B)}{1-B} \left( \frac{C}{X} \right)^{1-B} {}_2F_1 \left( 1, 1-B, 2-B; \frac{C}{X} \right), & x < c \\ -\frac{2}{\pi} p_0 \frac{\sin^2(\pi B)}{B} \left( \frac{X}{C} \right)^B {}_2F_1 \left( 1, B, 1+B; \frac{X}{C} \right), & x > c \end{cases}$$

where  $c$  is the offset of the indenter from the line of action of the normal load, associated with the elastic mismatch and shown in Fig. 1 and

$$\tan(\pi B) = \frac{1}{\beta f}$$

where

$$\beta = \frac{\Gamma(\kappa_1 + 1) - (\kappa_2 - 1)}{\Gamma(\kappa_1 + 1) + (\kappa_2 + 1)}, \quad \Gamma = \frac{\mu_2}{\mu_1}$$

is the elastic mismatch of the two materials, and

$$C = \frac{(1-c)}{(1+c)}, \quad X = \frac{(1-x)}{(1+x)} \quad \text{and} \quad Z = \frac{(1-z)}{(1+z)}.$$

If we let  $\omega = Z/C$  we find

$$\begin{aligned} \Phi(z) &= \frac{i(1-if)}{\pi^2} p_0 \sin^2(\pi B) \\ &\times \left\{ \frac{1}{B(1-\omega)} {}_3F_2 \left[ \begin{matrix} 1, & 1, & 1 \\ 1+B, & 2, & \end{matrix} ; \frac{1}{\omega-1} \right] \right. \\ &\quad \left. + \frac{\omega}{(1-B)(1-\omega)} {}_3F_2 \left[ \begin{matrix} 1, & 1, & 1 \\ 2-B, & 2, & \end{matrix} ; \frac{\omega}{\omega-1} \right] \right\} \end{aligned}$$

$$\begin{aligned} &- \frac{1-c}{2B} {}_3F_2 \left[ \begin{matrix} 1, & 1, & 1 \\ 1+B, & 2, & \end{matrix} ; -\frac{1}{2}(1-c) \right] \\ &+ \frac{1+c}{2(1-B)} {}_3F_2 \left[ \begin{matrix} 1, & 1, & 1 \\ 2-B, & 2, & \end{matrix} ; \frac{1}{2}(1+c) \right] \end{aligned} \quad (9)$$

where  ${}_3F_2$  is a generalized Hypergeometric function. This is valid provided that

$$|\arg(1-\omega)| < \pi \quad \text{and} \quad \left| \arg \left( 1 - \frac{1}{\omega} \right) \right| < \pi,$$

which corresponds to everywhere except on the line  $z=0$  inside the contact zone.

## 4 Conclusion

Results have been presented in closed form for the Muskhelishvili potential associated with the two-dimensional contact of a wedge and an elastic half-plane. This enables rapid and exact determination of all the surface and interior stress values. Closed-form solutions were also obtained for the surface values of tension  $\sigma_{xx}^0$  for a couple of important cases. This quantity is thought to be responsible for propelling surface cracks, caused either by monotonic or cyclic loading.

The pertinent equations were (4), (7), and (9) and their simplicity over existing solutions (where they existed) was demonstrated. Where previous solutions existed the simplicity of the new results is obvious.

## References

- [1] Truman, C. E., Sackfield, A., and Hills, D. A., 1995, "Contact Mechanics of Wedge and Cone Indenters," *Int. J. Mech. Sci.*, **37**, No. 3, pp. 261–275.
- [2] Truman, C. E., Sackfield, A., and Hills, D. A., 1996, "A Comparison of Plane and Axisymmetric Contacts Between Elastically Dissimilar Bodies," *ASME J. Tribol.*, **118**, pp. 439–442.
- [3] Truman, C. E., Sackfield, A., and Hills, D. A., 1998, "Sliding Contact Between a Blunt Wedge and an Elastically Dissimilar Half-Plane," *J. Strain Anal.*, **33**, No. 6, pp. 469–472.
- [4] Kuno, M., Waterhouse, R. B., Nowell, D., and Hills, D. A., 1989, "Initiation and Growth of Fretting Fatigue Cracks in the Partial Slip Regime," *Fatigue Fract. Eng. Mater. Struct.*, **12**, No. 5, pp. 387.
- [5] Yoshimoto, G., and Tzokizoe, T., 1958, "On the Mechanism of Wear Between Metal Surfaces," *Wear*, **1**, pp. 472–481.
- [6] Muskhelishvili, N. I., 1953, *Some Basic Problems of the Mathematical Theory of Elasticity*, Noordhoff, Groningen.
- [7] Hills, D. A., Nowell, D., and Sackfield, A., 1993, *Mechanics of Elastic Contacts*, Butterworth Heinemann, Oxford.
- [8] Jäger, J., 1997, "Half-Planes Without Coupling Under Contact Loading," *Arch. Appl. Mech.*, **67**, pp. 247–259.
- [9] Spence, D. A., 1973, "An Eigenvalue Problem for Elastic Contact With Finite Friction," *Proc. Cambridge Philos. Soc.*, **73**, pp. 249–268.
- [10] Goodman, L. E., 1962, "Contact Stress Analysis of Normally Loaded Rough Spheres," *ASME J. Appl. Mech.*, **29**, No. 3, pp. 515–522.